Nancy-Université
Université Henri Poincaré

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

Toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

➢ Contact SCD Nancy 1 : theses.sciences@scd.uhp-nancy.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

Nancy-Université

Université Henri Poincaré Nancy 1

Département de formation doctorale en informatique
UFR STMIA

École doctorale IAEM Lorraine

Navigation dans de grands ensembles non structurés de documents visuels

THÈSE

présentée et soutenue publiquement le 7 octobre 2008

pour l'obtention du

Doctorat de l'université Henri Poincaré – Nancy 1

(spécialité informatique)

par

Olivier Christmann

Composition du jury

Président :	M. Daniel R. Mestre, Directeur de Recherche, Institut des Sciences du Mouvement
Rapporteurs :	M. Jean-Daniel Fekete, Directeur de Recherche, INRIA Saclay - Ile-de-France M. Christian Jacquemin, Professeur, Université Paris Sud
Examinateurs :	Mme Noëlle Carbonell, Professeur, Université Henri Poincaré Nancy 1 Mme Laurence Nigay, Professeur, Université Joseph Fourier de Grenoble M. Ye-Qiong Song, Professeur, Institut National Polytechnique de Lorraine

Laboratoire Lorrain de Recherche en Informatique et ses Applications — UMR 7503

Loria

A mes parents

Remerciements

J'exprime tous mes remerciements à Noëlle Carbonell pour m'avoir accueilli dans l'équipe MErLIn au LORIA et grâce à qui j'ai pu mener mes recherches avec une rigueur scientifique et un plaisir permanents. Je la remercie pour la qualité de nos échanges, sa présence de chaque instant, ses précieux conseils et encouragements durant ces 4 années.

Je remercie vivement Jean-Daniel Fekete et Christian Jacquemin, qui ont accepté d'être les rapporteurs de ce travail. Je remercie également Laurence Nigay, Daniel Mestre et Ye-Qiong Song pour leur participation au jury de soutenance.

Je remercie les participants au trois expérimentations menées durant ma thèse.

Je remercie tous les membres de l'équipe MErLIn avec qui j'ai eu la chance de travailler, pour leur gentillesse et leur disponibilité : Daniel Gepner, Suzanne Kieffer, Matjaž Divjak, Jérôme Simonin, Marius Haţegan, Gaëtan Pruvost et Andrei Linţu. Je remercie également les stagiaires de l'équipe : Anne-Sophie Tranchet, Laurent Prost, Yannick Martinet, Stéphane Virte et Alexandre Roy, mon partenaire au Quille-Boule 2007 ! Je remercie également Danièle Marchand et Laurence Benini, les assistantes de projet MErLIn. Une pensée particulière pour Jérôme, Marius et Andrei, qui font partie aujourd'hui de mes meilleurs amis et grâce à qui ces 4 années ont été réellement stimulantes et exaltantes.

Merci à mes amis : Pierre-Antoine, les deux Fabien, Alexandre, Jérémie, Stéphane et Gaëtan.

Je remercie toutes les personnes avec qui j'ai travaillé durant mes activités d'enseignement. En particulier, Jean-Pierre Thomesse, Jean-Claude Derniame, Xavier Rebeuf, Michel Dufner, Leonor Prensa-Nieto et Nicolae Brinzei à l'ENSEM et Brigitte Wrobel-Dautcourt et Joseph Di Martino au Département Informatique de la Faculté des Sciences de Nancy.

Toute ma sympathie à Michel Dufner, qui a été mon premier professeur d'Informatique et qui m'a donné le goût à cette discipline. J'exprime également ma sympathie à Xavier Rebeuf, qui m'a donné l'envie d'enseigner quand j'étais élève puis, lorsqu'il était mon tuteur à l'ENSEM lors de mes années de monitorat, a toujours été présent pour me conseiller et m'aider. Enfin, mes remerciements à Jean-Pierre Thomesse pour son soutien.

Enfin, et c'est presque implicite, je remercie mes parents pour leur aide et leur soutien (mon père pour son *coaching* actif !). Et toute ma famille, particulièrement Anne-Lise, Jeannine, Georgette, Arlette.

Table des matières

Chapitre 1 Présentation générale **1**

Chapitre 2 Motivations de l'étude **5**

2.1 Les documents visuels 6

2.1.1 Définition et usages 6

2.1.2 Caractérisation d'un document visuel 6

2.2 La recherche d'un document visuel 7

2.2.1 La recherche par mots-clés 7

2.2.2 La recherche par le contenu 9

2.2.3 La navigation au sein de collections de documents visuels 10

2.3 Les visualisations 2D : applications aux photographies 11

2.3.1 *Overview + details / focus + context* 12

2.3.2 L'utilisation des données intrinsèques 13

2.3.3 Faible innovation et absence d'études ergonomiques 14

2.4 Les représentations 3D interactives 15

2.4.1 Naissance des représentations 3D, et premières limites 15

2.4.2 Les métaphores, tentatives de se rapprocher du monde réel 17

2.4.3 Les enjeux de la troisième dimension 19

2.5 Objectifs, Motivations et Hypothèses 21

Chapitre 3 Choix d'une représentation 3D interactive **25**

3.1 Une représentation, 2 vues différentes 25

3.1.1 Rappel des contraintes 25

3.1.2 Perspective et fidélité 27
3.1.3 Description des deux vues et implémentation 29
3.1.4 Discussion préliminaire 30
3.2 Des métaphores d'interaction adaptées 32
3.2.1 Le périphérique, un choix important 32
3.2.2 Le déplacement de/dans la représentation 33
3.2.3 Le dispositif de zoom 34
3.2.4 Le dispositif de repérage 34
3.2.5 Bref résumé 35
3.3 Discussion 35
3.4 Applications 37

Chapitre 4 Méthodologie expérimentale **39**
4.1 Tâches de recherche visuelle 39
4.1.1 Généralités 39
4.1.2 Tâche T1 : conception 42
4.1.3 Tâche T2 : conception 43
4.2 Matériel visuel expérimental 43
4.2.1 Choix du matériel visuel 43
4.2.2 Choix des cibles 44
4.2.3 Constitution des collections 45
4.3 Choix des participants 45
4.4 Déroulement des expérimentations 46

Chapitre 5 Etude préliminaire **49**
5.1 Méthodologie 49
5.1.1 Hypothèses testées 50
5.1.2 Conception de la représentation 51
5.1.3 Interaction avec les représentations 51
5.1.4 Représentation du point de vue courant de l'utilisateur 54

5.1.5 Implémentation . . . 55
5.2 Protocole expérimental . . . 56
5.2.1 Généralités . . . 56
5.2.2 Choix des cibles et descriptions textuelles . . . 57
5.2.3 Constitution des collections . . . 58
5.2.4 Conditions initiales . . . 59
5.2.5 Choix des participants . . . 59
5.2.6 Déroulement de l'expérimentation . . . 60
5.3 Méthodes d'analyse . . . 61
5.3.1 Données et mesures . . . 61
5.4 Etude quantitative . . . 62
5.4.1 Temps de recherche . . . 62
5.4.2 Nombre de succès . . . 62
5.4.3 Qualité du repérage spatial - T2 . . . 63
5.5 Etude qualitative . . . 65
5.5.1 Pertinence des cibles choisies et de leurs descriptions textuelles (T1) . . . 65
5.5.2 Locomotion *versus* Manipulation . . . 67
5.6 Conclusions générales . . . 70
5.6.1 Synthèse des résultats . . . 71
5.6.2 Préférences des sujets . . . 71
5.6.3 Respect des métaphores et améliorations . . . 72
5.6.4 Conclusion . . . 73

Chapitre 6 Première étude **75**

6.1 Méthodologie . . . 76
6.1.1 Présentation générale . . . 76
6.1.2 Terminologie . . . 77
6.1.3 Hypothèses à tester . . . 77
6.1.4 Matériel visuel et représentation . . . 78
6.1.5 Interaction avec la représentation . . . 79

6.1.6 Implémentation et développements logiciels 82
6.2 Protocole expérimental 83
6.2.1 Généralités 83
6.2.2 Choix des cibles et descriptions textuelles 83
6.2.3 Constitution des collections 85
6.2.4 Placement des cibles 86
6.2.5 Conditions initiales 87
6.2.6 Choix des participants 87
6.2.7 Déroulement de l'expérimentation 87
6.3 Méthodes d'analyse 88
6.3.1 Données recueillies 88
6.3.2 Variables et mesures 89
6.3.3 Validité interne 89
6.3.4 Validité externe 91
6.3.5 Tests statistiques 91
6.4 Influence du type de vue sur les performances 92
6.4.1 Résultats globaux 92
6.4.2 Discussion 94
6.5 Préférences des sujets 96
6.6 Influence des performances sur les préférences 98
6.6.1 Observations 98
6.6.2 Analyse statistique 99
6.7 Stratégies d'exploration visuelle 99
6.7.1 Détermination des stratégies 99
6.7.2 Caractéristiques statiques et dynamiques 99
6.7.3 Acuité visuelle 100
6.8 Qualité du repérage spatial 101
6.8.1 Examen initial 102
6.8.2 Choix du sens de rotation 102

6.9 Analyses complémentaires 104
6.9.1 Difficulté d'identification des paysages par rapport aux objets 104
6.9.2 Influence de la position verticale des cibles sur le succès de la recherche . . 104
6.10 Commentaires sur l'interface : suggestions d'améliorations 105
6.10.1 La représentation 3D 105
6.10.2 Dispositif d'interaction 105
6.10.3 Dispositif de repérage spatial 106
6.10.4 Eloignement-rapprochement 106
6.10.5 Zoom 106
6.10.6 Discussion 107
6.11 Conclusions générales 107
6.11.1 Synthèse des résultats 108
6.11.2 Conclusion 110

Chapitre 7 Seconde étude **111**

7.1 Méthodologie 112
7.1.1 Présentation générale 112
7.1.2 Terminologie 113
7.1.3 Hypothèses à tester 113
7.1.4 Matériel visuel : les trois représentations 114
7.1.5 Interaction avec les représentations 114
7.1.6 Implémentation et développement logiciel 119
7.2 Protocole expérimental 120
7.2.1 Généralités 120
7.2.2 Choix des cibles et descriptions textuelles 122
7.2.3 Constitution des collections 123
7.2.4 Placement des cibles 123
7.2.5 Conditions initiales 123
7.2.6 Choix des participants 124
7.2.7 Déroulement de l'expérimentation 124

7.2.8 Capture et suivi du regard pendant T2 126
7.3 Méthodes d'analyse 128
7.3.1 Sources de données 128
7.3.2 Variables étudiées 128
7.3.3 Validité interne 129
7.3.4 Validité externe 130
7.3.5 Tests statistiques 131
7.4 Présentation des résultats 131
7.4.1 Préférences des participants 131
7.4.2 Performances 133
7.4.3 Facteurs susceptibles d'influencer les préférences 135
7.4.4 Stratégies d'exploration visuelle et données oculométriques pour T2 136
7.4.5 Mémorisation de la position verticale et horizontale des cibles 138
7.4.6 Mesure du confort visuel pendant T2 142
7.4.7 Les vues 3D, retour sur les résultats de l'étude précédente 143
7.4.8 Comparaison 2D-3D 149
7.5 Commentaires des participants sur l'interface : suggestions d'améliorations 151
7.5.1 Représentations visuelles des collections 151
7.5.2 Interaction avec les visualisations 152
7.5.3 Dispositif de repérage spatial 153
7.5.4 Eloignement-rapprochement des vues 154
7.5.5 Zoom des photographies 155
7.5.6 Discussion 155
7.6 Conclusions générales 155
7.6.1 Synthèse des résultats 157
7.6.2 Conclusion 159

Chapitre 8 Conclusions et perspectives **161**

Annexe A Etude préliminaire **169**

Annexe B Premiere Etude **177**

Annexe C Seconde Etude **189**

Table des figures **209**

Liste des tableaux **213**

Bibliographie **217**

Chapitre 1

Présentation générale

Aujourd'hui, les ordinateurs individuels sont dotés d'une capacité de stockage de plusieurs centaines de giga-octets. Cette mémoire est majoritairement constituée de données personnelles. Un exemple simple : pour un disque dur de 320 giga-octets, le fabricant Iomega[1] annonce la possibilité de stocker jusqu'à 1 280 000 photographies, 5 920 heures de musique ou 480 heures de vidéo. Ces chiffres imposent un constat évident : le principal enjeu n'est plus de disposer d'un espace de stockage suffisant, mais bel et bien de pouvoir retrouver des données particulières dans cette énorme masse de fichiers.

Avec la démocratisation des loisirs numériques, nous trouvons trois grandes familles de fichiers sur les réseaux d'échange et les ordinateurs individuels : musiques, films et documents visuels. Les deux premières présentent des fichiers fortement structurés et indexés, étant principalement issus de grands distributeurs (*Emi*, *Universal*,...). De ce fait, ils sont référencés au sein de bases de données reconnues, comme par exemple *Gracenote*[2] pour la musique. Une simple recherche par mots-clés permet de retrouver facilement un de ces fichiers. Concernant les documents visuels, et principalement les photographies, la situation est totalement différente : ils sont massivement stockés et diffusés par des particuliers et non par des professionnels. Et les particuliers ne prennent pas la peine d'annoter leurs photographies et, bien souvent, seules les méta-données intrinsèques (date, caractéristiques de la prise de vue) sont disponibles, car ajoutées automatiquement par les appareils numériques. Ces méta-données restent insuffisantes pour retrouver quelques photographies particulières parmi plusieurs milliers. Des travaux de recherche ont porté sur cette problématique, avec comme solution l'indexation par le contenu (*Content-Based Image Retrieval*), s'appuyant par exemple sur la couleur, le contraste, la forme [Eakins et Graham, 1999] ou, plus récemment, sur la reconnaissance des visages [Zhao *et al.*, 2003]. Ces systèmes ne sont pas accessibles au grand public, et sont très souvent incapables de produire les résultats attendus. Dès lors, il devient nécessaire d'afficher simultanément plusieurs dizaines ou plusieurs centaines de photographies pour trouver celle(s) que l'on cherche. De nombreux travaux ont porté sur la visualisation de grands ensembles d'informations, la plupart s'intéressant aux données structurées. Dans le domaine de la photographie numérique, l'absence de structure et d'indexation se révèle être un handicap : Rodden et Wood [Rodden et Wood, 2003] confirment

[1] http ://www.iomega-europe.com
[2] http ://www.gracenote.com/music

que peu d'utilisateurs prennent la peine de classer leurs photographies en dossiers et sous-dossiers et qu'ils se contentent bien souvent de recopier le contenu de leur carte mémoire en un seul et même répertoire.

Nous sommes face à un double défi : élaborer de nouvelles représentations et métaphores d'interaction qui puissent permettre de naviguer de manière plus agréable, confortable et intuitive dans de grandes collections de documents visuels, et permettre de retrouver plus rapidement un ou plusieurs éléments particuliers au sein de ces collections. En 2001, D. A. Keim [Keim, 2001] concluait que le but ultime, peut-être atteignable dans les 5 ans, était que chaque machine de bureau dispose d'assez de puissance graphique pour permettre une exploration plus rapide et plus intuitive des grandes ressources de données. C'est aujourd'hui possible, les capacités des cartes graphiques étant largement suffisantes pour aller plus loin que les affichages classiques de type matriciel, comme ceux proposés par les explorateurs de fichiers ou les logiciels de gestion de collections de photographies (*Picasa*[3], *ACDSee*[4] par exemple).

Les logiciels récents, comme les nouvelles versions des systèmes d'exploitation d'ailleurs, exploitent de plus en plus la troisième dimension pour rendre plus réalistes et plus intuitives les interfaces. Mais au delà de l'impact esthétique, non négligeable il est vrai pour rendre les interfaces attrayantes, quel peut-être l'apport effectif de la troisième dimension en termes d'efficacité et de confort d'utilisation ? Depuis l'invention de la métaphore du bureau en 1984 par Apple, peu de choses ont finalement évolué, et nous sommes toujours en présence des interface de type WIMP (*Window Icon Menu Pointer*) ; cela peut s'expliquer en partie par l'impact trop prégnant du coût du changement [Beaudouin-Lafon, 2004]. Nous sommes donc face à un grand paradoxe : bien qu'étant capables de réaliser aujourd'hui ce qui était seulement imaginable il y a une vingtaine d'années, il demeure impossible de produire quelque chose sortant du modèle WIMP, car le temps d'adaptation serait rédhibitoire pour le grand public. C'est pourquoi, lorsque nous parlons de la troisième dimension, nous n'entendons pas réalité virtuelle, ou techniques avancées de visualisation, comme la vision stéréoscopique, par exemple. Nous voyons la troisième dimension à un niveau plus *basique*, permettant d'introduire la profondeur et, avec celle-ci, des dispositions spatiales plus évoluées.

L'objectif de notre recherche est de concevoir de nouvelles représentations et métaphores d'interaction de grandes collections non structurées de photographies, afin de faciliter non seulement la navigation mais aussi la recherche visuelle. Nous nous intéressons particulièrement à l'évaluation de l'efficacité de l'interaction et du confort d'utilisation. Notre but est similaire à celui suivi par Shneiderman [Shneiderman, 2003] : au travers de la conception de nouvelles interfaces et de leur évaluation ergonomique, être en mesure de proposer des recommandations que les concepteurs de logiciels pourront mettre en œuvre avec profit.

Partant d'une représentation générique sous forme d'un cylindre vertical, nous proposons deux représentations 3D destinées à contenir environ un millier de photographies non structurées, l'une favorisant l'immersion (vue interne), et l'autre favorisant la manipulation directe (vue externe). Nous avons adopté une approche expérimentale avec 3 études distinctes, pour nous permettre d'évaluer d'une part ces deux représentations l'une par rapport à l'autre, puis en les opposant à

[3]http ://picasa.google.com
[4]http ://fr.acdsee.com

une disposition classique à deux dimensions d'autre part. Pour chaque étude, nous avons conservé le même type de tâche, à savoir la recherche d'une photographie particulière disposée au sein d'une collection d'un millier d'images environ. Plus précisément, nous nous sommes intéressés à deux tâches courantes, réalistes, de recherche visuelle :

- d'une part, la recherche d'une photographie correspondant à une description textuelle, c'est-à-dire la recherche d'une photographie répondant à un ensemble de critères visuels et sémantiques exprimés verbalement. La photographie, ainsi que sa position dans le cylindre, sont inconnues de l'utilisateur ;
- d'autre part, la recherche d'une photographie familière dont la position dans la collection est connue de manière plus ou moins précise.

Le travail est organisé comme suit. Le chapitre 2 présente nos motivations quant à l'étude des représentations de grandes collections non structurées de documents visuels. Plus précisément, nous parlerons des différents types de visualisations proposées ces 20 dernières années, et qui ont profondément marqué et orienté les recherches sur la visualisation en général. Nous démontrerons en quoi ces visualisations sont peu adaptées aux documents visuels qui se prêtent mal aux techniques de recherche par indexation à partir de mots-clés ou des résultats d'analyses évoluées du contenu. L'intérêt scientifique de l'apport de la troisième dimension sera ensuite discuté, ainsi que l'attention portée aux collections non structurées de photographies.

Nous présenterons ensuite, dans le chapitre 3, la représentation 3D que nous avons choisie et conçue, un cylindre vertical, en précisant les raisons de sa supériorité sur d'autres alternatives que nous avons écartées. Ce sont sur cette représentation que porteront les 3 études expérimentales présentées dans les chapitres suivants. L'intérêt de proposer deux implémentations distinctes de cette représentation (vues interne et externe du cylindre vertical) sera discuté. Après avoir décrit les caractéristiques propres à chaque implémentation, nous présenterons de manière générale les modalités d'interaction et fonctionnalités retenues.

Le chapitre 4 décrit la méthodologie utilisée pour comparer et démontrer l'efficacité et la facilité d'utilisation de notre représentation pour naviguer au sein de grandes collections de photographies, en d'autres termes, pour comparer, dans des situations réalistes, la façon dont les implémentations de notre représentation peuvent être acceptées par les utilisateurs. Nous présenterons les deux tâches de recherche, ainsi que l'élaboration du matériel expérimental, et les aspects communs à l'étude préliminaire et aux deux autres études qui lui ont fait suite.

Dans le chapitre 5 est présentée l'étude préliminaire, conduite auprès de 8 sujets, qui avait pour but d'obtenir un premier *feedback* sur notre représentation et d'évaluer son intérêt potentiel. Nous avons donc comparé les vues interne et externe, pour les deux tâches présentées ci-dessus et décrites dans le chapitre 4. Les participants avaient à effectuer la recherche d'une photographie particulière dans 6 collections distinctes, pour chaque vue et chaque tâche. L'analyse a porté sur des données quantitatives (temps de recherche, nombre d'erreurs) ainsi que sur les données qualitatives issues des entretiens de débriefing et des questionnaires.

L'étude préliminaire ayant permis de montrer l'attrait des utilisateurs pour les deux implémentations de la représentation que nous proposons, nous avons décidé de mener une seconde étude. D'une part, pour tenir compte des enseignements de l'étude préliminaire, en particulier au niveau des corrections à apporter et des améliorations possibles, et ainsi rendre nos repré-

sentations réellement agréables et performantes à utiliser. D'autre part, pour être en mesure de réaliser des tests statistiques et obtenir par ce biais des résultats significatifs, grâce un à nombre de sujets plus important (20). L'objectif majeur de cette étude a été de déterminer quelle vue (interne ou externe) était la plus adaptée à la recherche visuelle, que ce soit en termes de confort ou d'efficacité. Les modifications mises en œuvre au niveau de la représentation, du protocole expérimental, et les résultats, analyses et conclusions sont présentés dans le chapitre 6.

Le chapitre 7 décrit la dernière étude menée, visant à comparer nos représentations 3D à une représentation 2D classique, couramment utilisée par les logiciels de gestion de collections de photographies (affichage matriciel, avec défilement latéral). 20 participants ont testé successivement ces 3 vues. Afin d'étudier plus finement les stratégies de parcours visuel, nous avons utilisé un *eye-tracker* pour enregistrer les parcours et les fixations oculaires. Nous rendons compte d'abord de la conception de l'interface 2D et des adaptations nécessaires du protocole expérimental des études précédentes, puis des résultats quantitatifs et qualitatifs obtenus, et des analyses et conclusions résultantes.

Le dernier chapitre présente les conclusions des évaluations de nos représentations 3D, et leur apport au regard des représentations actuelles les plus communément utilisées, en termes de facilité, d'efficacité et de confort d'utilisation. Nous discutons les résultats obtenus, particulièrement ceux relatifs à la troisième étude expérimentale. Enfin, nous dressons les perspectives envisagées.

Chapitre 2

Motivations de l'étude

Les données visuelles (photographies, dessins, plans, vidéos, pochettes de disques,...) sont devenues les données les plus courantes et les plus diffusées. L'explosion de la photographie numérique n'a fait que renforcer ce phénomène. Cette masse de données n'est plus cantonnée aux disques durs des ordinateurs personnels : en moins d'une dizaine d'années, l'accès pratiquement universel à Internet a changé la donne avec, en particulier, tous les mécanismes d'échanges de données (par exemple *Microsoft MSN*) ou de partage en ligne (par exemple *Yahoo Flickr*[5]). Deux questions majeures, soulevées dans *SemNet* [Fairchild *et al.*, 1999] suscitent notre intérêt : d'une part, savoir comment présenter (numériquement) à l'utilisateur de grandes collections de documents, et d'autre part, imaginer comment explorer, manipuler et modifier ces ensembles d'informations. Ces deux questions sont d'actualité et suscitent de nombreux travaux de recherche. Imaginer, concevoir et promouvoir des représentations simples, performantes et agréables à utiliser, aussi bien pour l'utilisateur *lambda* que pour l'utilisateur expérimenté, représente un défi majeur aujourd'hui. Le public visé, autant que les applications potentielles, sont des plus larges : accès personnel dans le cadre familial ou professionnel, accès universel dans le cadre de locaux publics, comme les bibliothèques, ou encore avec des bornes interactives, par exemple. Aujourd'hui, pour rechercher un ou plusieurs documents particuliers, deux techniques principales existent (souvent utilisées conjointement) :

- la navigation au sein de la collection de documents ;
- la recherche par mots-clés et/ou par le contenu, suivant des critères clairement définis.

Après avoir explicité notre définition d'un document visuel, nous exposerons ces deux techniques et positionnerons ainsi notre étude dans le contexte scientifique. Nous nous attacherons ensuite à présenter les travaux antérieurs concernant les représentations 2D et 3D appliquées aux documents visuels. Nous présenterons enfin nos objectifs, les motivations de notre étude ainsi que nos hypothèses de recherche. Un rapide résumé clôturera cette partie, avec un rappel des différents points abordés.

[5] http ://www.flickr.com

2.1 Les documents visuels

2.1.1 Définition et usages

Selon l'Organisation Internationale de Normalisation (ISO), un document (du latin *documentum*) est un :

> *Ensemble formé par un support et une information, généralement enregistré de façon permanente, et tel qu'il puisse être lu par l'homme et la machine.*

Nous nous intéressons ici aux documents sur support numérique. Il est certain qu'avec les facilités de numérisation, tout autre support peut être facilement numérisé (papier, audiovisuel et microforme) ; notre espace de travail est donc très large. Dans une première acception, un document visuel est composé de formes, géométriques ou non, mais ne contient aucun caractère alphanumérique. On peut citer par exemple le plan, le diagramme, le dessin, la photographie, l'estampe, la carte postale. La liste est longue... Aujourd'hui, cette définition est toutefois incomplète, rendue obsolète par l'utilisation de l'outil informatique. Nous sommes habitués aux aperçus, et ainsi la première page d'un rapport, ou encore l'affiche d'un film, permettent souvent d'identifier le document dans son ensemble, et deviennent de fait semblables à des synecdoques. Même sans être capable de lire le texte, la simple structure et l'organisation spatiale d'une page de document dactylographié peuvent donner des informations essentielles pour la reconnaissance du document dans son ensemble. Par exemple, les sites de vente en ligne de livres proposent souvent une numérisation de la première de couverture. Le champ des documents visuels est donc très large et couvre au final une partie très importante des documents informatiques que nous utilisons quotidiennement, soit localement, soit à distance via des sites web dédiés (*Yahoo Flickr*[6] pour les photographies, *YouTube*[7] pour les vidéos, en sont quelques exemples).

Il est encore une distinction à faire avec les documents visuels, à savoir le degré de familiarité de l'utilisateur qui les manipule. Nous distinguerons donc, en fonction du contexte d'utilisation, les données *familières*, connues, des données *non familières*, non connues. Ces dernières sont plus nombreuses que l'on pourrait penser : collection iconographique d'un musée, base de documents multimédias, par exemple.

2.1.2 Caractérisation d'un document visuel

Pour qualifier les documents visuels, il existe trois grandes classes de méta-données, définies par Drucker *et al.* [Drucker *et al.*, 2004] :

- les méta-données intrinsèques : elles sont principalement ajoutées de manière purement automatique par l'appareil qui « produit » le document visuel. Pour la majeure partie des documents, on trouve ainsi la date, le format, la taille, et un nom. Pour des photographies, les caractéristiques de la prise de vue (ouverture, temps d'exposition, résolution,...) sont également ajoutées par l'appareil photographique ;
- les méta-données extrinsèques : elles sont ajoutées explicitement *a posteriori* par l'utilisateur. Il est ici principalement question des mots-clés ajoutés pour qualifier un document

[6] http ://www.flickr.com
[7] http ://www.youtube.com

visuel. De fait, la nature de ces méta-données est donc plus subjective ;
– les méta-données analysées : elles sont obtenues majoritairement à l'aide de logiciels permettant d'analyser le « contenu » des documents visuels, par exemple, pour faciliter la recherche de visages humains, d'objets, de motifs.

Les deux premières catégories de méta-données sont les plus courantes, la dernière demandant l'emploi de logiciels spécifiques, encore peu répandus actuellement, surtout pour le grand public. La figure 2.1 montre, pour une photographie, les données intrinsèques et extrinsèques les plus courantes.

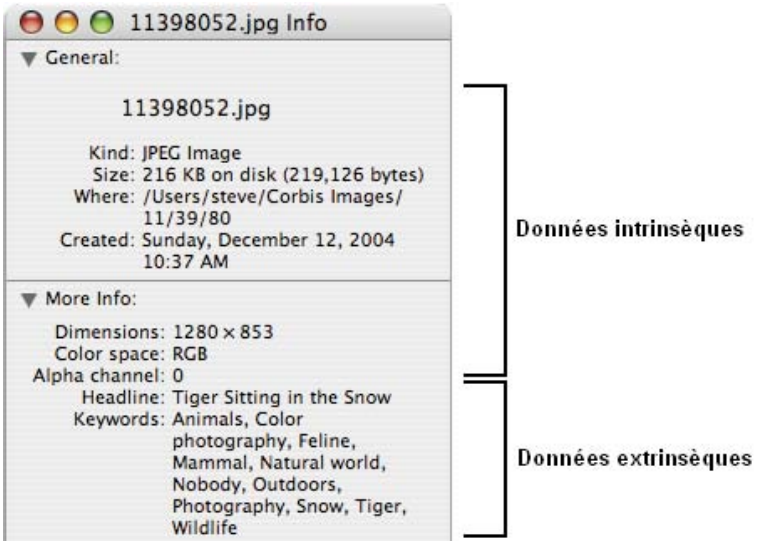


FIG. 2.1 – Données intrinsèques et extrinsèques d'une photographie

2.2 La recherche d'un document visuel

Deux techniques sont majoritairement utilisées, à savoir la recherche par interrogation d'une base de données, avec des critères variés, et la simple navigation, se ramenant à la recherche visuelle d'un item au sein d'une grande collection de documents. Nous allons présenter dans cette partie ces deux types de recherche, avec une mise en valeur de leurs traits distinctifs.

2.2.1 La recherche par mots-clés

La recherche par mots-clés est la plus naturelle, car la plus usuelle. Cela est encore plus vrai aujourd'hui avec les moteurs de recherche sur internet (*Yahoo*, *Google*). Une définition concise de ce type de recherche est donnée dans le rapport technique sur *Apple Spotlight* [Apple, 2005] :

> *It can find something by its text content, filename, or information associated with it, known as metadata*

Le grand public a pris pour habitude de rechercher une information sur Internet en spécifiant simplement quelques mots, censés la caractériser de la manière la plus précise possible. Les résultats

sont relativement fidèles aux attentes car, dans ce cas précis, les documents sont majoritairement textuels : d'un point de vue simpliste, la recherche se réduit à parcourir un document pour vérifier s'il contient les mots recherchés. L'extraction des informations est ici plutôt simple.

Quelles méta-données ?

Ce qui est relativement simple pour un document textuel l'est beaucoup moins pour un document visuel, la recherche par mots-clés étant alors loin d'être aussi immédiate, et dépendant le plus souvent des utilisateurs eux-mêmes, et de leur capacité à produire l'information qui n'est pas directement présente dans les photographies. A la création du document visuel (sous forme numérique), seules les données intrinsèques sont présentes. Les informations basiques, dont la consultation est habituelle pour un fichier informatique, se cantonnent à un nom (souvent abscons), une date de création, un format et une taille. Par ailleurs, la date de création n'est pas nécessairement fiable car, d'une part, l'appareil générant le fichier n'est pas forcément réglé à la bonne date et, d'autre part, la date conservée par l'ordinateur n'est pas obligatoirement celle de création. Dans un usage général, retrouver un document particulier, en ne disposant que de ces seules informations, est quasiment impossible. Pour une photographie, les informations techniques concernant la prise de vue ne sont pas directement exploitables, sauf peut-être pour un professionnel et/ou pour des recherches très ciblées. Les données intrinsèques peuvent être utilisées, pour de petites collections d'images, comme des critères discriminants très spécifiques, mais ne peuvent supporter un passage à l'échelle. C'est l'ajout de mots-clés aux documents (*keywords* ou *tags*) qui permet d'utiliser les techniques usuelles d'indexation afin d'obtenir des bases de données qu'il est possible d'interroger en spécifiant un ou plusieurs mots décrivant le document visuel.

L'annotation

Dès lors, la pertinence de la recherche ne repose plus sur la performance de l'algorithme, mais sur la qualité de l'annotation, et donc du soin apporté à l'association de mots-clés pertinents. L'annotation, quelle que soit sa modalité, se heurte toujours à un problème récurrent : la motivation des utilisateurs. A un retour de vacances, renseigner les mots-clés de plusieurs centaines de photographies est une tâche ingrate que peu de gens veulent assumer. Plus les collections sont grandes, moins les utilisateurs prennent la peine d'annoter leurs photographies [Rodden et Wood, 2003]. De plus, au téléchargement sur l'ordinateur, l'utilisateur, connaissant ses photographies, ne voit pas l'intérêt, à court terme, d'associer des mots-clés à chaque photographie.

Des outils d'assistance à l'annotation ont été proposés, sans grand succès. D'une part, l'usage de la parole dans *ShoeBox* [Mills *et al.*, 2000], censée atténuer le côté laborieux de la tâche d'annotation, avec comme justification qu'il est plus agréable et facile de parler de ses photographies que d'entrer des commentaires textuels. Cette technique se heurte néanmoins à un problème quasi-rédhibitoire : la faiblesse de la précision de la reconnaissance (28,4% dans cette étude), principalement due à l'abondance de noms propres, mal reconnus. D'autres facteurs entrent également en ligne de compte, comme la qualité du microphone, ou encore la diction. Beaucoup d'utilisateurs n'annotent en définitive que les répertoires, par gain de temps, ce qui limite

d'autant l'intérêt de l'annotation, les recherches produisant alors trop de résultats. Une autre approche consiste à se servir de l'émulation entre les personnes pour favoriser l'action d'annotation. Dans *PhotoFinder Kiosk* [Shneiderman *et al.*, 2002], les données dépassent le cadre personnel et suscitent des activités d'annotation à caracrère social : les utilisateurs prennent plaisir à se retrouver sur les photographies, et à les commenter à l'intention de leurs collègues. Se retouver à plusieurs autour d'un écran a un effet non négligeable sur la motivation, ce qui est confirmé par Kules *et al.* [Kules *et al.*, 2004] :

> *The experience of exploring pictures with others (even at separate stations) is qualitatively different than solitary exploration.*

Mais il est certain que la situation évoquée dans [Shneiderman *et al.*, 2002] est relativement peu courante, et il est probable que l'enthousiasme observé ne résisterait pas à l'épreuve du temps.

Certains travaux se sont axés sur la performance de la recherche, plus que sur la facilité d'ajout des mots clés. Par exemple, *PhotoFinder* [Kang et Shneiderman, 2000] permet des recherches booléennes. Mais celles-ci sont mal appréhendées par le public : dans les moteurs de recherche comme *Google*, peu de personnes utilisent la recherche avancée. Avec *Apple Spotlight* [Apple, 2005], il est également possible d'utiliser des termes spécifiques pour combiner différents critères de recherche. Par exemple, pour rechercher des documents de type « image » créés aujourd'hui, on écrit la requête suivante : *kind :image date :today*. Il est nécessaire alors d'apprendre un « langage » spécifique, en dehors de la capacité et des intérêts d'une grande partie des utilisateurs. Pour ces deux techniques, la faiblesse potentielle de la qualité des mots-clés reste donc toujours un obstacle de taille.

2.2.2 La recherche par le contenu

L'obstacle que constitue la définition de mots-clés pertinent peut être franchi grâce à la recherche par le contenu qui produit des méta-données plus riches permettant une recherche plus fine et surtout via un traitement ne nécessitant pas, ou très peu, l'intervention humaine. Ces techniques appelées CBIR (*content-based image retrieval*) ont été explicitées par Kato [Kato, 1992] qui décrit des expériences d'extraction automatique d'informations provenant d'images, basée sur les caractéristiques de forme et de couleur. Un récapitulatif complet de ces techniques peut être trouvé dans [Eakins et Graham, 1999]. Dans *Shoebox* [Mills *et al.*, 2000], il est possible de faire des recherches sur le contenu visuel. Ces recherches s'appuient en fait sur une similarité avec une ou plusieurs images de référence. D'après Rodden et Wood [Rodden et Wood, 2003], il reste néanmoins plus simple d'utiliser des techniques de navigation classiques pour retrouver une photographie familière, que d'essayer de construire une requête visuelle d'après sa mémoire. Cette technique impose de plus de parcourir une partie de la collection pour rechercher une image caractéristique, ce qui est finalement peu simple. Et là encore, les outils actuels sont peu développés pour un usage du grand public, et restent bien souvent des prototypes de recherche. Une des conclusions de Mills *et al.* concernant *Shoebox* est :

> *Visual-based retrieval tools may not be especially valuable.*

2.2.3 La navigation au sein de collections de documents visuels

Nous avons passé en revue deux principales techniques de recherche, s'appuyant sur des mots-clés, ou sur le contenu « visuel » des photographies. Ces deux techniques sont loin d'être idéales, et certaines études [Rodden et Wood, 2003] ont montré que l'intérêt des utilisateurs grand public pour l'indexation automatique et l'annotation ne dépasse pas le stade des premières utilisations, à cause du caractère très contraignant de la saisie des mots-clés, mais également de la faiblesse et de la complexité actuelles des outils d'indexation. Le manque de temps et d'intérêt pour l'annotation est également confirmé par Kusnatovitz et Shneiderman [Kustanowitz et Shneiderman, 2005] :

> *Frequently users do not have time or interest in annotating these pictures.*

C'est également une des conclusions de Kuchinsky *et al.* [Kuchinsky *et al.*, 1999] qui aimeraient pouvoir évaluer la nature de la relation entre la volonté des utilisateurs d'annoter leurs documents, et les bénéfices de cette technique.

Dès lors, l'exploration et la navigation au sein de collections de documents visuels, pour autant qu'elles soient simples, confortables et performantes, peuvent jouer un rôle plus que déterminant dans la recherche de documents. Rodden et Wood [Rodden et Wood, 2003] constatent par ailleurs que :

> *The type of requirement that can only be satisfied with a query is relatively rare ; the most common requirements can easily be satisfied by browsing.*

Dans *ScentTrails*, Olston et Chi [Olston et Chi, 2003] remarquent que la recherche par navigation (*browsing*) est plus adaptée que la recherche par mots-clés dans certains cas précis :

- si l'utilisateur ne sait pas complétement ce qu'il cherche ;
- si certains critères de recherche sont difficilement traduisibles en mots-clés ;
- si la terminologie n'est pas connue.

D. A. Keim [Keim, 2001] confirme que l'exploration de données visuelles permet habituellement une exploration plus rapide, et délivre souvent de meilleurs résultats que les algorithmes de recherche automatique. C'est également une des conclusions de Mills *et al.* [Mills *et al.*, 2000] :

> *The most basic and important feature of any photomanagement software is its support for browsing photographs.*

Il est clair que les besoins des utilisateurs diffèrent selon leur profil : le contexte personnel est totalement différent du contexte professionnel. Dans un contexte personnel, c'est surtout le côté familier et affectif des photographies qui rend leur exploration plus facile, et leur annotation plus laborieuse. *A contrario*, dans un contexte professionnel, les photographies sont souvent stockées dans des bases de données et correctement indexées. Dans ce sens, nous sommes tout à fait d'accord avec l'assertion de Kuchinsky *et al.* :

> *We believe that consumers' information-seking activities differ from those of information retrieval professionals, and that is particulary true when the information involves home media such as photos or videos. In these settings, direct search may be less frequent, whereas riffling and browsing through collections of materials becomes the norm (and serendipity is expected).*

.

Les points faibles des méthodes présentées précédemment affectent surtout l'utilisation grand public : de nombreux logiciels existent pour faciliter la gestion des collections de documents visuels (*Avanquest ACDSee, Apple iPhoto, Google Picasa, ...*), ainsi que des sites web de partage, mais la plus grande majorité des utilisateurs se contentent d'utiliser les outils de base fournis avec les systèmes d'exploitation, l'Explorateur de *Windows* en étant l'exemple le plus représentatif. Le grand public utilise encore peu les logiciels dédiés, et n'est donc pas prêt d'adopter des techniques avancées de recherche. C'est pourquoi nous avons choisi de centrer nos recherches sur les représentations interactives de grandes collections de photographies, ainsi que sur la navigation dans ces collections.

En effet, déjà peu annotées, les photographies « personnelles » sont bien souvent mal classées, les utilisateurs se contentant de recopier le contenu de leur carte mémoire (dans le cas de photographies) dans un seul et même dossier [Rodden et Wood, 2003]. Nous sommes donc face à des ensembles importants et non structurés de données visuelles pour lesquels l'annotation et l'indexation automatique, sans être impuissantes, restent des techniques critiquables avec lesquelles les utilisateurs ne sont pas familiers.

Dans tous les cas, même dans le cadre d'une recherche par mots-clés, le moteur de recherche ne peut produire le résultat unique demandé, car souvent les mots-clés sont relativement généraux, et de ce fait communs à d'autres documents. C'est d'ailleurs toute l'ambiguïté de l'annotation, car des mots-clés précis permettant d'obtenir de bien meilleurs résultats sont plus difficiles à retrouver par la suite : l'équilibre à trouver est donc problématique. En conséquence, les moteurs de recherche génèrent de nombreux résultats, et c'est à l'utilisateur de raffiner ces résultats, par une exploration visuelle. Combs et Bederson [Combs et Bederson, 1999] abondent dans ce sens en proposant que les résultats multiples obtenus par les systèmes de recherche d'images (*image retrieval*) soient affichés pour permettre leur exploration et ainsi le raffinement de la recherche. Il en est de même dans *Periscope* [Wiza *et al.*, 2004], système dans lequel plusieurs représentations sont proposées pour l'affichage des résultats d'une requête, afin d'adapter au mieux représentation et données à afficher.

Quels que soient les outils et les aides, la simple navigation dans une collection de documents visuels est, de fait, incontournable. C'est dans cette direction que les recherches doivent être développées, pour fournir des représentations simples et efficaces, associées à des interacteurs adaptés et intuitifs, favorisant une exploration visuelle agréable, confortable et performante.

2.3 Les visualisations 2D : applications aux photographies

La recherche par navigation (*browsing*) étant supérieure aux autres types de recherche dans la majorité des cas (documents visuels personnels, bases de connaissances destinées au grand public) pour les différentes raisons explicitées précédemment, de nombreuses techniques de visualisation ont été proposées, qui s'appliquent directement à la navigation au sein de grands ensembles de documents visuels.

2.3.1 *Overview + details / focus + context*

La masse de données à représenter grandissant, ainsi que le niveau de détail des représentations à afficher, les limites physiques des dispositifs d'affichage ont rapidement été atteintes, que ce soit en termes de taille ou de résolution. Même si, en une dizaine d'années, la taille moyenne des écrans est passée de 14 à 20 pouces, cette augmentation n'est malheureusement pas suffisante. Ce problème, qui est donc loin d'être nouveau, a donné naissance à de nouvelles techniques de visualisation : le lecteur pourra trouver dans [Plaisant *et al.*, 1995] une taxonomie complète de ces différentes techniques. L'enjeu est de pouvoir agrandir une zone particulière de la visualisation, tout en gardant à l'écran la totalité de la visualisation. Le but recherché est d'accroître les capacités cognitives de l'utilisateur, en augmentant le nombre d'informations disponibles simultanément. Ces techniques sont communément appelées *overview + details* et *focus + context* [Baudisch *et al.*, 2002].

Les visualisations recourant à la technique *overview + details* s'appuient en général sur deux fenêtres de visualisation distinctes : une première affichant l'ensemble ou une grande partie de la visualisation, et une seconde affichant une partie restreinte de la visualisation, sur laquelle l'utilisateur porte son intérêt. Un exemple courant est un navigateur de cartes, tel que ceux que l'on peut trouver sur des sites de cartographie (`www.mappy.fr` ou `www.viamichelin.fr`). La plupart des logiciels de gestion de collections de photographies proposent ce type de technique, avec une vue de la collection (restreinte ou non) et une vue séparée pour l'affichage d'une photographie agrandie. Un exemple est le logiciel *Xnview* dont la figure 2.2 reproduit une copie d'écran.

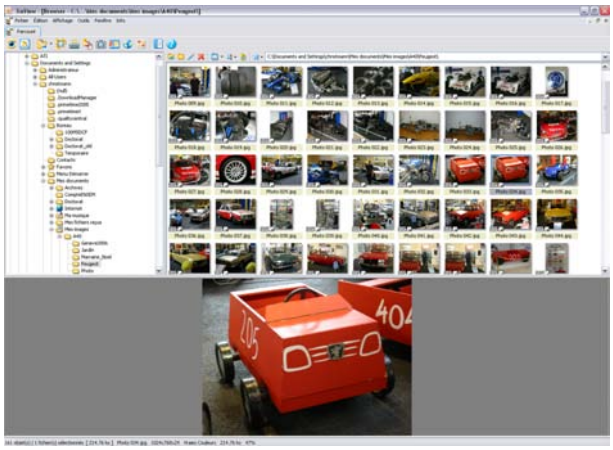


FIG. 2.2 – Logiciel *XnView* : exemple d'affichage *overview + details*

Dans le cadre de cette même approche, une adaptation des cartes arborescentes (*treemaps*) [Shneiderman, 1992 ; Bederson *et al.*, 2002] a également permis l'affichage d'une collection de photographies, par l'usage d'un algorithme de placement adapté aux données visuelles. *Photomesa* [Bederson, 2001] permet ainsi d'afficher toute la collection de photographies simultanément et d'agrandir une photographie particulière en préservant son contexte. L'affichage de la collection dans son intégralité offre la possibilité de se passer de moyens de navigation, mais l'inconvénient majeur est la faible lisibilité dans le cas d'un nombre important d'items.

La seconde technique, dite *focus + context*, affiche directement une zone à fort niveau de

détail au sein de la visualisation, avec des distortions progressives de l'affichage entre la zone à fort grossissement et le reste de la représentation, à faible résolution. Le maintien du contexte global est important, car il permet de conserver la compréhension du lien entre l'objet agrandi et le reste de la visualisation. Les vues en œil de poisson (*fisheye views*) [Furnas, 1986] et les arbres hyperboliques (*hyperbolic trees*) [Lamping *et al.*, 1995] en sont les exemples les plus connus. *Document Lens* [Robertson et Mackinlay, 1993] fait également partie de ces techniques. Utilisées principalement dans le cadre de structures arborescentes, ces visualisations permettent en outre l'affichage d'un nombre de nœuds élevés, jusqu'à 100 000 pour les vues en œil de poisson [Heer et Card, 2003]. Les applications restent très variées : arbres, calendriers, cartes géographiques... C'est au moyen de distorsions qu'il est possible d'afficher la quantité d'informations souhaitée. Le point fort est ici l'interactivité, puisqu'il est possible de modifier dynamiquement la position du focus et de le placer sur n'importe quelle partie de la représentation. Le principal inconvénient de ce type de représentation est le choix d'afficher la visualisation dans son ensemble, ce qui peut devenir peu exploitable dans le cas d'un nombre important d'objets à afficher. De plus, dans le cas précis des arbres hyperboliques, il faut constater que la lisibilité et la compréhension de la structure sont obtenues au prix d'une importante place « perdue ».

Ces deux techniques sont souvent opposées à des techniques plus conventionnelles de navigation dans de grands ensembles de documents visuels, appelées *zooming/panning* où, suivant la métaphore, il est possible de modifier la position de la caméra ou du document virtuel. La hauteur de « survol » permet alors de faire varier continûment le niveau de détail. De nouvelles techniques ont également été mises au point, notamment concernant la vue en perspective (*perspective viewing*) [Guiard *et al.*, 2006] : au cours de la navigation, il est possible de maintenir simultanément le contexte (vers l'horizon) et le détail (proche de la caméra). Cette technique est néanmoins plutôt adaptée à des documents textuels, car la déformation induite par la perspective rend à la fois le contexte peu exploitable, et les détails peu discernables. De nombreuses améliorations ont porté sur la navigation au sein d'un document, mais relativement peu s'intéressent à des représentations composées de multiples objets. Avec *OrthoZoom* [Appert et Fekete, 2006], il est par exemple possible de contrôler à la fois le déplacement et le zoom.

2.3.2 L'utilisation des données intrinsèques

Plutôt que de s'appuyer uniquement sur des techniques de visualisation, certains travaux proposent une adaptation plus fine au contenu des documents, en s'appuyant sur les données intrinsèques. Dans *TimeQuilt* [Huynh *et al.*, 2005], par exemple, les images sont classées temporellement. Cette approche se heurte à deux inconvénients majeurs : d'une part, au faible degré de fiabilité de la date, comme nous l'avons déjà mentionné ; d'autre part, au fait que toutes les photographies ne peuvent être affichées simultanément, une photographie devant être choisie pour représenter un *cluster* (groupement par date). Organiser les photographies temporellement semble légitime, mais ce type de représentation ne peut se suffire à lui-même.

Un autre point de vue intéressant est celui proposé dans *Photo Tourism* [Snavely *et al.*, 2006] : les photographies sont regroupées spatialement, suivant l'angle de vue et la distance à l'objet photographié. L'intérêt est qu'il suffit d'une seule photographie annotée pour que l'annotation soit diffusée aux autres photographies se rapportant au même objet. Cette représentation, bien

adaptée à des photographies de paysages, reste peu utile pour des portraits, des gros plans, ou des prises de vues artistiques. Ce travail est dans une certaine mesure une extension et un raffinement du *Personal Digital Historian* [Shen *et al.*, 2003 ; Moghaddam *et al.*, 2004] qui, en introduisant la notion de « *browser* intuitif », autorise la navigation à partir de 4 questions : « qui ? quand ? où ? quoi ? ». La géo-localisation précise nécessite malheureusement la présence d'un GPS lors de la prise de vue, ce qui est encore loin d'être généralisé.

Dans *PhotoTOC* [Platt *et al.*, 2002], la date sert également à des fins de *clustering*, avec l'idée sous-jacente que les photographies attachées à une même période le sont aussi à un même événement. Même si cette organisation permet une recherche plus efficace et plus satisfaisante, sa portée reste limitée aux photographies personnelles, prises par l'utilisateur lui-même.

L'utilisation de données intrinsèques comme aide à l'organisation spatiale des documents au sein des affichages et, de ce fait, à l'exploration visuelle, semble pouvoir jouer un rôle non négligeable. Mais les solutions apportées à l'heure actuelle restent par trop limitées à des cas bien précis, difficilement généralisables. Ces approches ne peuvent se suffire à elle-mêmes pour aider l'utilisateur dans sa tâche de recherche et d'exploration. Il semble donc plus efficace de fournir aux utilisateurs des représentations simples (en apparence), mais dont les propriétés visuo-spatiales pourront en faire de sérieuses alternatives aux solutions actuelles. Et cela parce que, jusqu'à maintenant, l'effort n'a guère porté dans ce sens.

2.3.3 Faible innovation et absence d'études ergonomiques

Il est acquis que la navigation et la recherche visuelle sont *de facto* incontournables pour les tâches de recherche de documents visuels. Jusqu'à aujourd'hui, la grande majorité des logiciels de gestion de collections de photographies restent semblables, et s'en sont tenus aux principes énoncés par Combs et Bederson [Combs et Bederson, 1999], à savoir la nécessité d'afficher le maximum de documents simultanément (en faible résolution), ainsi que la possibilité d'examiner un document particulier dans sa résolution native. Ainsi, nous sommes bien souvent en présence de tableaux à deux dimensions pour l'affichage de la collection, dans lesquels la navigation est rendue possible par des ascenseurs (*scroll-bars*), l'image agrandie d'un item apparaissant dans une zone spécifique. Pourtant, des études récentes [Simonin *et al.*, 2005] ont montré que les structures matricielles ne sont pas les meilleures pour la recherche visuelle, car elles procurent un faible guidage du regard, et impliquent donc des temps de recherche plus élevés.

Parmi les travaux de recherche ayant porté sur le simple affichage de documents visuels, la plupart se sont appuyés sur des techniques 2D existantes. Pour certaines solutions innovantes, des études ont été réalisées afin de vérifier l'efficacité et les performances d'utilisation de prototypes. Par exemple, l'étude de l'utilisation d'une implémentation de *Photomesa* sur PDA a été menée auprès de 15 étudiants en informatique, qui ont successivement utilisé ACDSee et *Pocket Photomesa* [Khella et Bederson, 2004]. A noter que le plaisir d'utilisation a été plus important avec ce dernier. Pourtant, il n'y a aucune différence statistiquement significative en termes de performances entre les deux systèmes. Moghaddam *et al.* ont également présenté une étude ergonomique de leur nouvel algorithme de placement, mais seuls 6 sujets y ont pris part [Moghaddam *et al.*, 2004]. La conception de nouveaux arrangements d'affichage proposés par

Kustanovitz et Shneiderman [Kustanowitz et Shneiderman, 2005] a également donné lieu à une étude ergonomique, mais avec seulement 4 participants. Même si les jugements furent positifs, les performances ne furent pas prises en compte, ce qui limite l'intérêt scientifique de cette étude.

Jusqu'à présent, la portée des différentes études a été relativement limitée : les méthodes pour recueillir les données sont bien souvent trop simplistes, et les protocoles expérimentaux sont décrits de façon imprécise. Ainsi, il est difficile d'évaluer les facteurs de nature à influencer les résultats. Il ne s'agit pour la plupart d'entre elles que d'études « préliminaires », dont les résultats sont donc à prendre avec précaution, du fait du faible nombre de sujets et des incertitudes quant aux protocoles expérimentaux. Il faut également noter que les études présentant des comparaisons de nouvelles représentations avec des solutions existantes [Khella et Bederson, 2004 ; Wiss et Carr, 1999] fournissent des résultats difficiles à interpréter, du fait du non-contrôle de nombreux facteurs, par exemple, les variations d'une condition à l'autre ou d'une étude à l'autre, de la disposition des items à l'écran ou des moyens d'interaction. Ainsi, l'impact de chaque facteur sur les résultats qualitatifs et quantitatifs ne peut être clairement isolé et étudié. La qualité ergonomique des visualisations interactives innovantes de grandes collections de documents visuels reste donc à prouver.

2.4 Les représentations 3D interactives

Jusqu'à maintenant, peu de travaux de recherche ont porté sur les représentations 3D interactives destinées à l'affichage de grandes masses de documents visuels. Depuis une dizaine d'années, les cartes graphiques des ordinateurs supportent l'affichage de représentations en 3D dimensions. La puissance de calcul des ordinateurs, même des plus courants, permet de gérer de manière réaliste les interactions avec de telles représentations. Il reste à étudier leur intérêt potentiel pour l'utilisateur.

2.4.1 Naissance des représentations 3D, et premières limites

De nouvelles représentations 3D ont été présentées dès le début des années 90 : ainsi les murs en perspective [Mackinlay *et al.*, 1991] permettent d'afficher des données temporelles, avec une plus grande quantité d'informations visibles simultanément. L'avantage supplémentaire, non négligeable, est que la distinction entre contexte et détail est visualisée de manière simple par la perspective, c'est-à-dire par des distorsions naturelles (voir la figure 2.3). Ce type de représentation cumule de fait plusieurs avantages, par rapport à une représentation « plate ». La métaphore est très simple à comprendre pour les utilisateurs, et il est facile de modifier le niveau de détail. De plus, les déformations sont relativement faibles, et le regard est naturellement guidé vers la partie centrale de la représentation. Il reste que ce type de visualisation n'est pas réellement adapté à la visualisation de collections de documents visuels.

De nombreux travaux ont également porté sur les données hiérarchiques, avec différentes représentations, comme les arbres coniques (*cone trees*, voir figure 2.4) ou une adaptation des cartes arborescentes (*treemaps*), en 3D [Bladh *et al.*, 2004]. L'apport de la troisième dimension est justifié par l'affichage simultané d'un plus grand nombre de nœuds, mais également par la

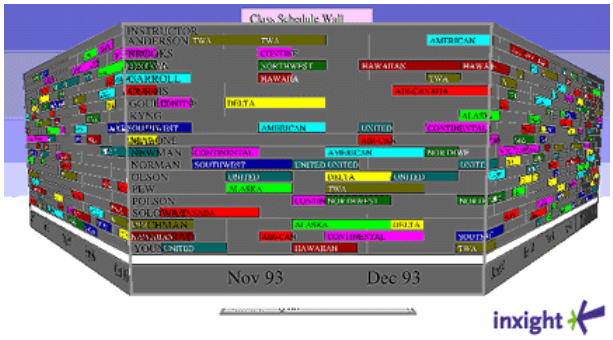


FIG. 2.3 – Murs en perspective

volonté de faciliter la compréhension des relations structurelles et hiérarchiques. D'ores et déjà, nous pouvons noter que l'apport de la troisième dimension n'est pas toujours positif : une évaluation des arbres coniques [Cockburn et McKenzie, 2000] montre que les utilisateurs sont plus lents pour retrouver des fichiers avec ce système (comparé à l'explorateur de fichiers *Windows*). Il convient néanmoins de remarquer que les commentaires des sujets indiquent que cette représentation permet de mieux appréhender la structure de l'espace d'informations. En outre, le

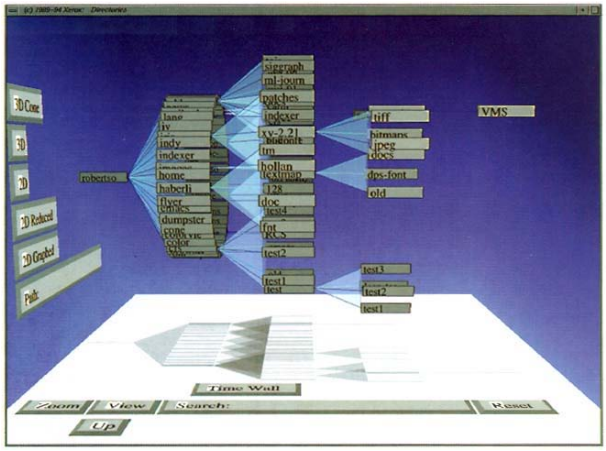


FIG. 2.4 – Arbre conique : l'arborescence est présentée horizontalement et comprend ici 4 niveaux (plus la racine) ; les boutons situés à gauche permettent de changer le mode de représentation

principal problème auquel se heurte ce type de représentation, et d'ailleurs la grande majorité des représentations 3D, a trait au phénomène d'occlusion : certains éléments sont masqués par d'autres, ce qui tend à diminuer le nombre d'items visibles simultanément, même si la perception de la structure demeure.

En plus d'être souvent destinées à l'affichage d'informations bien spécifiques, et de ce fait peu adaptées à l'affichage de documents visuels, la plupart de ces représentations sont souvent difficiles à appréhender par le grand public. Les performances de l'affichage sont bien trop souvent mises en avant, au détriment de la facilité d'accès, du confort ou encore du plaisir d'utilisation, même si ce dernier aspect est plus subjectif. L'augmentation de la performance implique trop souvent une complexité croissante, ce qui exclut de fait le grand public. *Source Viewer 3D*, une

représentation permettant d'analyser le code source [Marcus *et al.*, 2003] en est un exemple frappant (voir la figure 2.5).

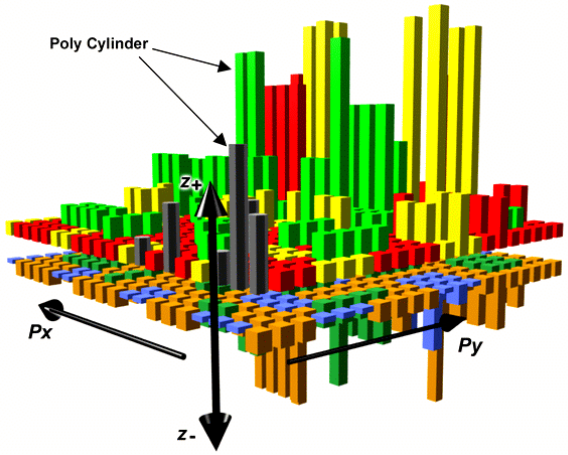


FIG. 2.5 – *Source Viewer 3D* : chaque *poly cylinder* représente une ligne de code, la couleur représente le type de structure de contrôle et la hauteur, le niveau d'emboîtement

Un autre exemple pourrait être *Information Cube* [Rekimoto et Green, 1993], une représentation permettant d'afficher un contenu hiérarchique dans un cube. Ce type de représentation est réellement difficile à comprendre par les utilisateurs [Wiss et Carr, 1999], ce qui empêche de manière évidente une large diffusion. Et les exemples similaires sont malheureusement nombreux.

L'emploi de la troisième dimension ne doit donc pas se faire à n'importe quel prix : il faut savoir en évaluer l'apport, et également les contreparties. Réaliser des maquettes de visualisation en trois dimensions est aujourd'hui chose relativement aisée, avec la large diffusion de bibliothèques 3D, et même d'ateliers de réalité virtuelle. Faire en sorte que ces représentations puissent être comprises et utilisables facilement se révèle être d'une toute autre difficulté. C'est le soin apporté à la conception, par l'utilisation de modèles simples et réalistes, ainsi qu'à l'évaluation ergonomique détaillée de ces nouvelles représentations, qui doit permettre une plus large diffusion.

2.4.2 Les métaphores, tentatives de se rapprocher du monde réel

Les problématiques ayant trait à la conception de représentations destinées à l'affichage de grandes quantités d'informations, ainsi qu'aux moyens d'interaction à imaginer pour pouvoir manipuler et explorer ces informations, sont aujourd'hui plus que jamais d'actualité. L'accès de plus en plus courant à l'informatique et aux grands espaces de connaissances impose de mettre en œuvre des techniques de visualisation et d'interaction simples, immédiatement compréhensibles par le grand public. Dans le milieu de l'informatique grand public, promouvoir de nouveaux concepts reste très difficile, car l'apport, que ce soit en termes de confort ou de performance, doit être immédiat, sous peine de susciter un phénomène de rejet. Les habitudes d'utilisation sont un obstacle majeur au développement de solutions novatrices. De nombreuses propositions de solutions ont été publiées au cours des vingt dernières années. C'est la sortie de l'*Apple*

Macintosh en 1984 qui fut le premier pas décisif dans l'avènement de l'informatique grand public. La métaphore du bureau permet de rendre plus évidentes les différentes fonctionnalités offertes par le système, par analogie avec le monde réel : l'utilisateur dispose, sur un bureau virtuel, une calculatrice, une corbeille, des dossiers ...

La troisième dimension, gage de plus grande fidélité au monde réel, renforce encore l'usage de métaphores. Le bureau (*desktop*) peut prendre la forme d'une ville, dans laquelle les documents sont représentés par des immeubles et les liens hiérarchiques par des routes [Rogers *et al.*, 1994]. Le *WebBook* [Card *et al.*, 1996] exploite la métaphore du livre, tandis que *WebForager* met en œuvre celle d'une bibliothèque permettant de stocker des documents virtuels. Ces travaux représentent une première approche convaincante, dans la mesure où ils visent à rendre l'interaction avec de multiples documents affichés simultanément, plus rapide et plus agréable. Cette volonté a été partagée ensuite par de nombreux travaux de recherche. Nous en allons en détailler deux en particulier, car ils apportent des nouveautés marquantes.

Data Mountain

Le cadre applicatif choisi par [Robertson *et al.*, 1998] est les « favoris » des navigateurs web : résumée de façon simpliste, l'idée est d'utiliser un plan incliné pour y disposer une centaine de miniatures de pages web (voir figure 2.6). Cette visualisation, en 3D, se propose de tirer parti



FIG. 2.6 – *Data Mountain* avec l'affichage de 100 favoris

de la mémoire spatiale des utilisateurs pour leur permettre de retrouver plus rapidement les favoris qu'ils auront eux-mêmes positionnés. La troisième dimension permet de se rapprocher du monde réel, avec le recours à la métaphore de la montagne. La représentation devient plus réaliste. Les règles simples, telles la relation taille-éloignement ou le masquage, sont respectées. La perspective est donc utilisée pour afficher davantage d'informations simultanément, tout en offrant une meilleure lisibilité de l'organisation des items. Ce regroupement facilite en outre le traitement visuel pré-attentif lors de la recherche de favoris [Treisman, 1985]. Un travail important a été effectué pour rendre compréhensibles et réalistes les déplacements des items au sein de la visualisation. La notion de profondeur a été renforcée par l'emploi d'indices visuels (ombres, animations) et sonores. Une étude ergonomique a été menée pour évaluer de manière qualitative

et quantitative cette interface. Comparée au gestionnaire de favoris d'*Internet Explorer*, *Data Mountain* est une alternative crédible, les temps demandés pour classer ou retrouver des favoris, ainsi que le nombre d'erreurs y étant plus faibles. De plus, les participants, dans le questionnaire final, ont marqué une préférence significative pour cette représentation.

Task Gallery

Bien avant l'interface *Aero* sous *Microsoft Vista* ou *Beryl/XgL* sous *Linux*, *Task Gallery* [Robertson *et al.*, 2000] proposait déjà une amélioration notable du système usuel de gestion de fenêtres en utilisant la troisième dimension. Les différentes fenêtres apparaissent plaquées sur les murs, sols et plafonds d'une galerie d'art virtuelle, sorte de long couloir. L'avantage principal est non seulement de pouvoir visualiser plusieurs fenêtres en même temps, mais également de pouvoir les organiser. Les études menées suggèrent que *Task Gallery* aide l'utilisateur à gérer ses fenêtres, tout en étant « amusant ». L'intérêt majeur se trouve dans la métaphore 3D, qui se propose de tirer parti de la mémoire spatiale, de par les relations spatiales induites par la disposition des fenêtres, et l'influence positive des images des fenêtres, comme cela a été démontré par des études antérieures [Czerwinski *et al.*, 1999]. La justification du choix de la métaphore est importante : les utilisateurs, principalement les non-joueurs, tendent à se sentir perdus au sein d'environnements 3D dans lesquels ils doivent naviguer ; c'est pourquoi l'usage d'un couloir droit, associé à l'analogie avec la galerie d'art, est tout à fait adapté. Cela impose de fait une navigation égocentrique, et par là même la création d'outils de navigation simples et intuitifs. Il faut enfin remarquer que cette interface, bien que destinée à remplacer un gestionnaire de fenêtres, s'appuie sur des captures des différentes fenêtres actives : il serait donc envisageable de l'utiliser pour la gestion de collections de photographies et de tirer profit de cet environnement 3D et des outils de navigation associés. Bien entendu, *Task Gallery* se prête mal à la gestion de grandes collections de photographies, mais ce prototype ouvre d'intéressantes perspectives.

2.4.3 Les enjeux de la troisième dimension

Comme nous avons pu le voir jusqu'à maintenant, le recours à la troisième dimension offre à la fois des avantages et des inconvénients. Essayons-nous à un récapitulatif de ceux-ci. Le premier intérêt, le plus fondamental, est de pouvoir afficher plus d'informations simultanément, en utilisant la profondeur et donc les nouvelles possibilités de disposition inhérentes à cette dimension. La troisième dimension permet en outre de se rapprocher du monde réel par la création d'environnements plus réalistes, et facilite donc la compréhension et la manipulation de ces représentations par analogie avec des situations réelles. La simplicité d'utilisation et le confort peuvent s'en trouver facilités, si tant est que l'on ait apporté suffisamment de soin à la conception. Le niveau de performance, aussi bien en termes de visualisation que d'utilisation, peut également augmenter, même s'il reste plus difficile à mesurer.

Un axe de recherche important vise à évaluer comment les représentations faisant usage de la troisième dimension peuvent assister de manière plus efficace les capacités cognitives et, plus particulièrement, la mémoire spatiale mise en jeu lors de la recherche d'une image familière en naviguant dans une collection elle aussi familière. L'étude présentée dans [Tavanti et Lind,

2001] porte sur la représentation de structures arborescentes : une des conclusions, basée sur les différences significatives de performance entre les deux interfaces, est que la représentation 3D peut contribuer à améliorer de manière significative l'efficacité de la mémoire spatiale. Le protocole peu rigoureux de cette expérimentation interdit d'accorder beaucoup de crédit à cette conclusion. Ce constat est partagé par Cockburn [Cockburn, 2004] qui a répliqué l'étude de Tavanti, mais en intégrant les facteurs non contrôlés par le protocole expérimental de cette étude : orientation verticale/horizontale de la fenêtre, utilisation de lettres d'habitude utilisées plutôt horizontalement, taille des icônes et de l'espace entre elles. Les résultats de cette expérimentation montrent le caractère limité des effets de la troisième dimension : on n'observe aucune différence significative quant à l'efficacité de la mémoire spatiale. Ces résultats sont en concordance avec ceux de Cockburn et McKenzie [Cockburn et McKenzie, 2001 ; Cockburn et McKenzie, 2002]. Il faut pourtant tout de même relativiser ces conclusions, dans le sens où nous sommes en présence de tâches très artificielles (retrouver des lettres ou des drapeaux dans une arborescence). Et, bien entendu, la recherche des avantages de la troisième dimension ne doit pas se limiter à la meilleure utilisation des capacités cognitives : d'autres aspects, et non des moindres, rentrent en ligne de compte, comme un usage plus agréable et confortable, une meilleure ergonomie au niveau de l'interaction.

D'autres études ont eu pour but d'évaluer l'apport de la troisième dimension sur différents types de tâches [Wiss et Carr, 1999] en comparant plusieurs représentations 3D d'informations hiérarchiques. Les principales conclusions de ces études semblent contradictoires : pour les logiciels étudiés (*CamTree*, *Information Cube* et *Information Landscape*), plus la représentation se veut réaliste, plus la complexité d'usage et de compréhension pour les utilisateurs devient importante. Ce problème, crucial, impose donc de particulièrement soigner le couple représentation/navigation, de façon à ce que l'utilisateur ne se sente pas perdu pas dans le monde virtuel.

Une comparaison de DataMountain à une interface 2D équivalente [Cockburn et McKenzie, 2001], pour le classement et la recherche de favoris (sous forme de miniatures), aboutit à la conclusion que les performances sont meilleures en 2D, mais ce résultat n'est pas statistiquement significatif. *A contrario*, on note une forte préférence pour la vue 3D, même si celle-ci est contrainte et que l'utilisateur ne peut se déplacer virtuellement que suivant deux axes. Il semble que ce soit surtout l'impression de profondeur qui suscite la sensation d'une troisième dimension réaliste, et qu'il soit inutile de recourir à des méthodes de navigation offrant les 6 degrés de liberté pour parvenir à donner cette sensation. Ce constat peut expliquer la faible différence observée entre les deux interfaces considérées. Une seconde étude [Cockburn et McKenzie, 2002] centrée sur la comparaison entre des interfaces physiques et virtuelles en 2D, 2D 1/2 et 3D, confirme ces résultats : les auteurs sont sceptiques quant à l'apport du 3D à la recherche d'informations dans des représentations en perspective. Si les questionnaires indiquent que les utilisateurs jugent les interfaces 3D moins efficaces, il faut remarquer que ces résultats s'appuient sur des représentations statiques, avec un type de tâche bien spécifique, relativement éloigné de la recherche dans des collections non structurées de documents visuels.

Compte tenu des différentes observations et remarques que nous avons pu établir jusqu'à maintenant, basées sur l'étude des travaux de recherche antérieurs et sur la prise en compte, non seulement des besoin des utilisateurs, mais également des outils et techniques qui sont à notre

disposition aujourd'hui, nous allons, dans la partie suivante, présenter les objectifs et motivations de notre étude, ainsi que nos hypothèses de travail.

2.5 Objectifs, Motivations et Hypothèses

Le but de notre étude est d'évaluer de manière détaillée l'efficacité, la simplicité et le confort d'utilisation de deux visualisations 3D interactives de grandes collections non structurées de documents visuels, destinées à faciliter les activités de recherche et d'exploration visuelles. Cet objectif principal est motivé, notamment, par la revue critique des recherches antérieures sur la visualisation de masses de données et leur application à la visualisation de grands ensembles de documents multimédias ou visuels, étude dont nous allons d'abord résumer les conclusions majeures.

Malgré les efforts effectués dans le domaine de la recherche par mots-clés, l'intérêt de ce type de technique se heurte à la faible motivation du grand public pour annoter leurs documents visuels. Cette tâche fastidieuse, bien qu'allégée par l'existence de méthodes d'indexation automatique basée sur le contenu, est un frein majeur à l'usage des moteurs de recherche pour la recherche de documents visuels. Les difficultés majeures tiennent d'une part au volume des informations à traiter et, d'autre part, à la pauvreté des relations entre les items. Nous sommes donc, de fait, en présence de collections peu, voire même non structurées. Les méta-données intrinsèques, par la généralité et/ou la non fiabilité de leur contenu, ne peuvent jouer le rôle de « liant » entre les différentes entités d'information. Même en considérant le cas idéal d'un contenu informationnel correctement annoté, deux obstacles se dresseraient encore. D'une part, les moteurs de recherche ne peuvent fournir l'unique résultat attendu, du fait du caractère souvent ambigu des mots-clés choisis, ou encore de leur champ trop vaste, qui ne peuvent caractériser correctement les documents visuels correspondants. D'autre part, il est clairement établi que dans certaines situations, la recherche par *browsing* est préférable, principalement lorsque l'utilisateur n'a pas totalement cerné ce qu'il cherche (ce qui correspond peu ou prou à des données non familières). La visualisation de masses d'informations visuelles, la navigation et la recherche visuelle dans de tels ensembles représente donc, aujourd'hui encore, un domaine scientifique comportant de nombreux problèmes ouverts.

Il est facile de constater que de nombreuses solutions commerciales existent sur le marché pour faciliter la gestion de collections de documents visuels. Une part écrasante de ces logiciels utilise le même type d'interface utilisateur, à savoir un affichage matriciel (les documents sont disposés sous forme d'un tableau, en lignes et en colonnes) et une zone supplémentaire pour l'affichage détaillé (ou *zoom*) d'un document sélectionné. Pour la navigation, le recours à des ascenseurs (*scroll-bars*) horizontaux et/ou verticaux est généralisé. Pourtant, des travaux ont montré que l'exploration visuelle est loin d'être optimale avec ce type de structure. De plus, l'interaction avec des ascenseurs, par les nombreux mouvements antagonistes qu'elle demande, accroît la fatigue de l'utilisateur, et lui demande une précision de pointage importante.

Parmi les nombreuses propositions issues des travaux de recherche récents, peu sont adaptées à l'affichage et à la navigation dans des collections de documents visuels. De plus, les études ergonomiques menées pèchent trop souvent par leur faible qualité : elles sont souvent décrites

de façon imprécise, et les facteurs éventuels, souvent mal identifiés, ce qui interdit l'obtention de résultats fiables. Le manque d'études approfondies des performances et du degré de satisfaction des utilisateurs est un frein important aux avancées dans le domaine. L'accent est par trop souvent mis sur les performances de la visualisation (afficher un maximum d'informations, par exemple), au détriment de la simplicité d'utilisation et de la qualité de l'interaction. Enfin, les représentations destinées au grand public sont loin d'être les plus nombreuses, et nous sommes bien souvent en présence de représentations absconses pour l'utilisateur *lambda*.

Depuis quelques années, la puissance des ordinateurs a atteint un niveau laissant aux concepteurs la totale liberté de réaliser les interfaces qu'ils désirent. Il n'est pas uniquement question de puissance de calcul brute, mais également de puissance graphique, permettant l'affichage de nouvelles représentations en trois dimensions avec, comme ligne d'horizon, la capacité de se rapprocher du monde réel, et donc de produire des représentations plus simples à comprendre, mais également dotées de moyens d'interaction mieux adaptés. Cette troisième dimension bouleverse la donne, dans le sens où il est possible d'afficher plus d'informations simultanément, mais également de tirer parti au mieux des capacités cognitives des utilisateurs, en particulier de leur mémoire spatiale. Pourtant, peu de représentations 3D sont aujourd'hui destinées à l'affichage de collections non structurées de documents visuels. Le travail sur les données non hiérarchiques reste toujours le parent pauvre de la recherche dans ce domaine, malgré son intérêt indéniable.

En outre, les visualisations 3D offrent des particularités intéressantes, du fait même de leur géométrie. Les vues en perspective sont naturellement un affichage de type *focus + context* dans lequel la profondeur permet une distorsion continue entre les objets situés à l'avant-plan et ceux situés à l'arrière-plan. Les aspects négatifs inhérents aux affichages en perspective, comme les phénomènes d'occlusion, se doivent d'être limités au maximum, pour maintenir le profit de la troisième dimension. L'usage de ce type de technique permet en outre d'aider à l'exploration visuelle par un meilleur guidage du regard entre zone d'intérêt et contexte.

Dans le cas de documents visuels, trois types de tâches sont particulièrement courants [Shneiderman, 1996] :

- le filtrage de la collection à la recherche de photographies vérifiant certaines propriétés particulières (appartenance à un thème particulier, ou documents provenant d'événements différents, mais partageant une caractéristique commune) ;
- la recherche de documents familiers, dont la position est connue ;
- la recherche de documents familiers, dont la position est inconnue.

Il est donc plus que primordial de disposer d'une représentation dont la structure serait simple à comprendre, qui serait à la fois confortable, performante et simple à prendre en main. Garantir à l'utilisateur une meilleure mémorisation spatiale permettrait de réduire les temps de recherche, et donc de minimiser l'effort à accomplir pour retrouver un document visuel.

Les études empiriques ou expérimentales présentées dans les chapitres suivants visent à évaluer la représentation 3D que nous proposons. Ces dernières ambitionnent d'offrir une vue d'ensemble cohérente et simple de la collection, tout en garantissant la possibilité de disposer d'un niveau de détail plus important pour l'examen d'un document visuel particulier. Nous attendons également une mémorisation plus précise et sans effort de la position de documents visuels dans la représentation, utile pour les recherches ultérieures. Nos hypothèses de recherche sont les

suivantes :

- hypothèse A : L'usage de la troisième dimension, associée à une représentation adaptée, permet d'accroître le confort d'utilisation, en particulier celui de la navigation, en garantissant un meilleur guidage du regard lors la recherche d'un item particulier de la collection, de par la distorsion naturelle induite par la perspective. Ce meilleur guidage influe de manière positive sur la longueur des parcours visuels et sur la rapidité de défilement, avec pour conséquence une plus grande précision et un plus grand confort.
- hypothèse B : Si les vues canalisent le regard vers les zones de la visualisation où les miniatures sont agrandies et/ou non déformées (hypothèse A), la vue externe où les images agrandies sont non déformées permet un meilleur guidage/contrôle des mouvements oculaires que la vue interne ; elle conduit à des parcours visuels plus courts et plus systématiques.
- hypothèse C : Plus que l'action de manipuler un objet virtuel représentant la collection, c'est de se déplacer dans un espace virtuel en trois dimensions qui est susceptible d'améliorer la précision de la mémorisation de la position des documents visuels dans la visualisation de la collection. Le déplacement dans une scène virtuelle, par son côté plus intuitif et réaliste, peut offrir une expérience plus forte et plus riche. Les différences entre ces deux types de représentation sont à même d'influencer non seulement les performances des utilisateurs, mais également leurs jugements subjectifs, les environnements immersifs étant connus pour apporter une réponse émotionnelle plus forte, ainsi qu'un plus grand sentiment d'engagement [Robertson *et al.*, 1998]. Il faut également ajouter que les utilisateurs d'ordinateurs sont familiers de l'interaction avec des objets 2D, tandis que le déplacement dans un espace 3D virtuel reste peu habituel pour la grande majorité d'entre eux.
- hypothèse D : Les représentations 3D sont susceptibles de recueillir majoritairement les préférences des utilisateurs par rapport aux interfaces 2D classiques. De ce fait, elles peuvent remplacer de manière immédiate, et profitablement, les représentations matricielles usuellement utilisées aujourd'hui.

Dans le chapitre suivant, nous présentons la représentation 3D retenue, ainsi que ses deux implémentations : la première étant un objet virtuel en 3D que les utilisateurs peuvent manipuler, tandis que la seconde est un espace virtuel dans lequel les utilisateurs peuvent se déplacer. Nous expliciterons leurs caractéristiques communes, les éléments qui font leur spécificité et, éventuellement, leur supériorité par rapport aux représentations 2D et 3D connues.

Les études présentées dans les parties suivantes permettront dans un premier lieu de valider, grâce à une étude préliminaire, l'intérêt de la représentation et de ses deux implémentations et leur validité par rapport aux hypothèses de travail. Cette étude aura également pour rôle d'entériner le protocole expérimental, et d'avoir un premier retour sur l'interface ainsi que sur les moyens d'interaction. La seconde étude, bâtie sur l'étude préliminaire, permettra de vérifier les hypothèses B et C et donc de répondre aux questions suivantes : est-ce qu'une représentation immersive est plus, ou moins, efficace et performante qu'une représentation non immersive ? Quels sont les jugements subjectifs provoqués par chaque type de vue ? La troisième s'attachera particulièrement à la vérification des hypothèses A et D, et ainsi évaluer sous quels aspects et dans quelle mesure les représentations 3D sont meilleures que les représentations 2D actuelles.

Chapitre 3

Choix d'une représentation 3D interactive

Dans ce chapitre, nous présentons la représentation 3D retenue, ainsi que les deux implémentations interactives que nous avons conçues, et qui servent de base aux trois études ergonomiques présentées dans ce mémoire. Dans un premier temps nous rappelons les contraintes qui nous ont guidés dans le choix d'une représentation appropriée. Ensuite, nous passons en revue les différentes solutions envisageables et détaillons la représentation choisie, ainsi que les deux visualisations qui en sont issues. Dans un deuxième temps, nous décrivons les fonctionnalités que nos visualisations implémentent, et nous esquissons un rapide panorama des différences entre les trois études. Nous terminons ce chapitre par une comparaison de nos propositions avec les solutions existantes dans le domaine de l'affichage de grands ensembles multimédias, et une esquisse des applications possibles de notre représentation.

3.1 Une représentation, 2 vues différentes

3.1.1 Rappel des contraintes

Il nous appartient de produire une représentation simple et performante, autant par la qualité de l'interaction proposée que par la disposition et le nombre d'items visibles simultanément. Jusqu'à maintenant, nous avons pu remarquer que les logiciels actuels de gestion de collections de documents visuels sont, dans une grande mesure, semblables. L'interaction fait toujours appel à des ascenceurs, et la disposition reste trop souvent une structure matricielle. Les autres types d'affichages, comme les *timelines* (ou affichages linéaires temporels) par exemple, restent cantonnées à des besoins bien particuliers, peu fréquents. La conception d'une représentation pertinente implique également le choix d'une disposition cohérente des items : une des conclusions de [Combs et Bederson, 1999] est que l'accroissement du nombre d'items doit aller de pair avec l'utilisation de techniques classiques, telles que les structures matricielles et les interfaces dites zoomables. Bien que cela paraisse contradictoire, il faut conserver une disposition « classique » tout en proposant une représentation garantissant un meilleur guidage du regard et un

meilleur confort visuel, que ceux offerts par les matrices 2D. La représentation doit également faciliter la mémorisation de la position des photographies familières.

L'usage de la troisième dimension doit nous permettre, naturellement, d'afficher davantage d'informations simultanément. Mais avant tout de garantir une meilleure compréhension et un meilleur confort d'utilisation, en fournissant une représentation dont la disposition des items à l'écran soit la plus naturelle et la plus rationnelle possible. Comme cela a été présenté dans la partie précédente, réaliser des interfaces en trois dimensions doit permettre de mieux tirer parti des capacités cognitives des utilisateurs, en particulier de leur mémoire spatiale ; cela peut se révèler très utile pour retrouver, par exemple, certains items de la collection plus rapidement. Gardons de plus à l'esprit qu'une double difficulté se dresse toujours devant nous : d'une part, inciter l'utilisateur à essayer une nouvelle représentation, et d'autre part faire en sorte que celui-ci se l'approprie, passé l'effet de nouveauté. Il est donc incontournable d'être en mesure de proposer des caractéristiques supérieures aux gestionnaires classiques de collections de documents visuels.

L'étude de Sebrechts *et al.* [Sebrechts *et al.*, 1999] semble indiquer que l'apport de la troisième dimension n'est pas immédiat : les temps de réponse avec l'interface 3D proposée par les auteurs ne sont équivalents à ceux obtenus avec l'interface 2D qu'après 6 sessions d'expérimentation. La complexité de leur représentation 3D explique ces résultats, et milite pour le choix d'une représentation simple à appréhender. C'est sans aucun doute la règle à respecter pour minimiser le temps d'adaptation, et ainsi atténuer le possible phénomène de rejet. De plus, d'après [Chen, 2000] les capacités cognitives (et parmi celles-ci, l'habilité spatiale) sont un des facteurs principaux à l'origine des différences interindividuelles. La réalisation d'une nouvelle représentation doit se faire avec l'objectif de ne pas pénaliser les utilisateurs dotés de faibles capacités cognitives : nous nous situons clairement dans une stratégie compensatoire [Messick, 1976], dans le sens où nous souhaitons amoindrir l'impact des différences interindividuelles.

L'usage d'une nouvelle dimension apporte une possibilité supplémentaire, celle de pouvoir réaliser des environnements immersifs, apportant plus de réalisme aux tâches informatiques. Au delà de cette considération, la troisième dimension peut également rendre la métaphore de la manipulation plus tangible. Nous avons donc deux possibilités : étendre le réalisme des interfaces basées sur la manipulation directe, ou leur donner un caractère immersif, en permettant à l'utilisateur de se « déplacer » dans un monde virtuel dans lequel seraient disposés les objets de la collection. L'objet de nos travaux est de comprendre et de démontrer laquelle de ces deux métaphores est la plus adaptée à l'exploration et à la recherche au sein de grandes collections de documents visuels, et ainsi, d'être à même de fournir aux concepteurs d'interfaces graphiques les conseils et les recommandations qui leur permettront de réaliser les interfaces de demain.

La problématique est de concevoir une représentation qui permette l'implémentation réaliste de ces deux métaphores tout en tenant compte des remarques précédentes et de la contrainte supplémentaire, à savoir qu'elles soient semblables en tout autre point de façon à rendre possible l'évaluation, par comparaison, de leur apport respectif à l'amélioration de l'interaction. C'est l'objet des parties suivantes.

3.1.2 Perspective et fidélité

Choisir un modèle 3D simple est pour nous primordial, et nous avons de plus une contrainte forte supplémentaire, que nous n'avons pas encore introduite, et qui est incontournable : la fidélité au contenu. Le contexte de l'étude étant les documents visuels, le rendu de ces documents à l'écran est un point crucial : au delà d'une certaine déformation, le risque de dénaturer le contenu d'un document est grand, avec pour conséquence immédiate une identification visuelle impossible. Il convient d'observer que certains contenus sont plus robustes que d'autres à ce type de problème : même légèrement dénaturé, un objet réel ou un paysage sera plus facile à reconnaître qu'un visage. Mais comme nous ne faisons pas d'hypothèse *a priori* sur le contenu des documents visuels, il faut amoindrir au maximum toute déformation pour assurer une certaine généricité à notre représentation.

Le plan en perspective

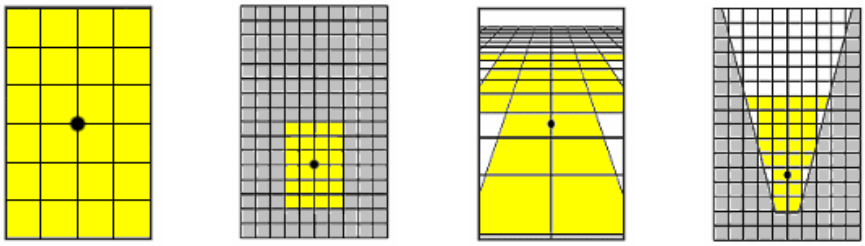
Cette contrainte exclut de fait les représentations sous forme de plans en perspective, la déformation étant trop importante, autant pour les objets « proches » de l'utilisateur que pour les objets situés à l'arrière-plan. De plus, l'équilibre entre le nombre d'items « exploitables » et l'intensité des déformations est quasi-impossible à atteindre. La figure 3.1 montre 16 photographies plaquées sur un plan en perspective : à l'avant-plan, les photographies sont très déformées, et rendraient difficiles l'identification d'un visage ; à l'arrière plan, discerner le thème de la photographie devient très difficile. Même si l'angle de vue permet de rendre visibles plus d'items qu'une



Fig. 3.1 – Exemple de plan en perspective avec 16 photographies

vue 2D classique, l'effet de la perspective annule en réalité cet avantage. Comme le montre la figure 3.2, le gain est pratiquemment nul : les figures 3.2.b et 3.2.d montrent l'intégralité du document, quand la figure 3.2.a (respectivement 3.2.c) montre la vue courante dans une représentation 2D (respectivement en perspective). La coloration jaune montre la partie réellement exploitable : bien que la surface visible soit plus importante dans l'affichage en perspective, la surface exploitable est similaire à la surface visible en 2D.

Cela illustre qu'il est possible de trouver une disposition 2D (par nature exempte de toute distorsion) pour laquelle la taille moyenne des items est plus importante que dans la vue en perspective. Il est bien entendu possible d'augmenter l'angle de la caméra par rapport au plan pour réduire les problèmes liés à la perspective, mais cette manipulation ferait perdre le gain de la troisième dimension. Il convient donc de constater que ce type de représentation, proposé



(a) (b) (c) (d)

FIG. 3.2 – Comparaison 2D (a, b) - plan en perspective (c, d) : la vue globale du document est représentée par les schémas b et d, les schémas a et c représentant la vue courante ; la partie réellement « exploitable » est en jaune

par [Guiard *et al.*, 2006] pour parcourir rapidement un fichier textuel, s'il peut être adapté à certaines situations, se révèle inadapté à l'affichage de documents visuels. De plus, le plan incliné ne convient pas réellement aux deux métaphores que nous souhaitons implémenter, car il favorise davantage la locomotion (déplacement dans l'environnement constitué des éléments de la collection) que la manipulation.

Les formes géométriques classiques

Vernier et Nigay [Vernier et Nigay, 1997] présentent différentes représentations, basées sur des formes géométriques simples, pour l'affichage de résultats de recherches par mots-clés. Une des représentations, la sphère, semble être une candidate sérieuse, car elle peut convenir à la fois pour induire un sentiment d'immersion (vue interne) et pour renforcer la métaphore de la manipulation directe (vue externe). Pourtant, comme le plan en perspective, ce type de représentation présente plusieurs inconvénients rédhibitoires. D'une part, le problème de la distorsion des photographies reste entier, dès que l'on ne considère plus l'image faisant face à l'utilisateur. D'autre part, la sphère étant uniformément continue, le repérage est très difficile, quel que soit le point de vue, intérieur ou extérieur. Même si cette considération est secondaire à ce stade de la réflexion, il faut également raisonner en termes de puissance de calcul : avec le nombre élevé de documents que nous souhaitons afficher, la mise à jour de l'affichage au cours des interactions va être coûteuse. En effet, à chaque mouvement de la sphère autour de son centre, il faut recalculer la forme et la taille des images visibles. A moins de trouver un algorithme rendant cette opération peu coûteuse et de disposer d'un matériel puissant, l'affichage sera très lent. Le public que nous visons étant des plus larges, il faut garder à l'esprit certaines contraintes techniques concernant le matériel courant. Même si l'utilisation d'une sphère tronquée permettrait de réduire ce dernier inconvénient, le problème des déformations et celui de l'absence de repères demeureraient.

Les vues présentant à la fois des déformations cartésiennes en x et polaires inverses (respectivement polaires) paraissent séduisantes au regard des problèmes soulevés par la sphère. Ces déformations correspondent aux vues intérieure et extérieure d'un cylindre vertical. Au final, ce choix s'impose, car cette représentation cumule les avantages que nous attendions, tout en minimisant les inconvénients des autres représentations évoquées précédemment. D'une part, ces

deux vues permettent d'implémenter les deux métaphores que nous souhaitons comparer : la vue interne, par sa concavité, favorise le sentiment d'immersion dans un monde virtuel représentant les objets de la collection ; la vue externe, par sa convexité, suggère la notion de manipulation d'un objet virtuel représentant la collection. D'autre part, le phénomène de distorsion est relativement limité et ne concerne réellement qu'une partie de l'affichage, c'est-à-dire les zones proches des bords de l'écran. Enfin, la continuité visuelle à assurer lorsque le cylindre tourne autour de son axe est limitée à une dimension, ce qui rend plus simple le repérage au sein de la collection. Cette unique dimension rend la navigation dans le cylindre facile à appréhender, quand les rotations de la sphère, par rapport à son centre, induisent des mouvements complexes.

3.1.3 Description des deux vues et implémentation

Nous avons donc choisi deux vues d'un cylindre générique pour représenter la collection. Les miniatures des documents visuels sont plaquées sur les faces latérales intérieure (vue interne, VI) ou extérieure (vue externe, VE) d'un cylindre vertical. Ces deux vues sont représentées dans la figure 3.3. L'adoption d'un cylindre vertical plutôt qu'horizontal est justifiée par le fait que ce



(a) Vue intérieure (b) Vue extérieure

FIG. 3.3 – Vues intérieure et extérieure du cylindre vertical

dernier semble moins à même de susciter un sentiment d'immersion. Le choix délibéré de considérer deux vues différentes d'un même objet s'explique par le désir de garantir une comparaison non biaisée des performances et des préférences subjectives des sujets concernant l'utilisation de chaque vue. Ce choix de conception réduit donc significativement le nombre de facteurs pouvant influencer les comportements, les performances et les préférences des utilisateurs. Les déformations propres à chaque vue, ainsi que la nature de la métaphore seront donc à mettre en relation avec les performances observées, particulièrement en termes d'efficacité et de confort visuel.

L'adoption de ces représentations, et donc le plaquage des miniatures sur les parois du cylindre, permet une utilisation limitée de la troisième dimension (la profondeur), et évite l'apparition de phénomènes d'occlusion tendant à diminuer significativement le nombre d'items visibles simultanément. La limitation des distorsions étant un enjeu primordial, nous avons choisi de

fixer un diamètre largement supérieur à la largeur de l'affichage, de façon à réduire les effets de perspective, sans pour autant les annuler, ce qui amoindrirait la force des métaphores. Outre les distorsions, l'effet de la perspective se traduit naturellement par la taille des items : dans la vue interne, les miniatures sont de plus faible taille dans la région centrale de l'écran, tandis que cette réduction se manifeste dans les zones latérales pour la vue externe. De manière plus précise, le choix du diamètre a été fait empiriquement, avec le souci de conserver un juste équilibre entre la réduction de la distorsion et la préservation du caractère *focus + context* de notre représentation. Dans la pratique, un sixième de la représentation est visible à l'origine.

Les utilisateurs sont habitués aux représentations classiques sous forme matricielle, avec des interacteurs prenant la forme d'ascenseurs pour la navigation. Ces représentations sont bornées, donc disposent d'un « début » et d'une « fin ». Contrairement à ces représentations, la continuité du cylindre nous place face à un problème analogue à celui de la sphère. Lors de la rotation du cylindre autour de son axe, l'utilisateur peut rapidement perdre tout point de repère. Pour remédier à cette difficulté, une solution simple a été adoptée et implémentée : l'utilisation d'une « fente noire », qui est en réalité une « colonne » du cylindre, vierge de photographies. Ainsi l'espace devient visuellement discontinu et l'utilisateur a la trace d'un parcours complet de la collection.

Le déplacement vertical (ou horizontal) utilisé pour naviguer dans les structures matricielles prend ici la forme d'une rotation, possible dans les deux directions. Comparées aux tableaux 2D classiques, les représentations proposées permettent de faciliter et d'optimiser les examens répétés de l'ensemble de la collection : les utilisateurs ne sont pas obligés de changer le sens de défilement (*scrolling*) ou de revenir au début de la représentation après chaque exploration totale de la collection. D'autre part, toute information est atteignable en un demi tour au plus, à condition de choisir judicieusement le sens de parcours. Cela devrait, suivant la position de l'item recherché et la force de la mémorisation de cette position, amoindrir les temps de recherche.

L'implémentation de ces deux représentations de documents visuels a été identique pour l'étude préliminaire et les deux études suivantes. Nous avons choisi d'utiliser un cylindre composé de 80 colonnes (moins une, laissée vierge) et de 11 lignes, soit au total 869 documents visuels. D'abord développées sous oRis [Harrouet, 2000] pour l'étude préliminaire, les représentations furent ensuite réalisées sous ARéVi [Harrouet *et al.*, 2006], cet atelier offrant des possibilités plus importantes d'assistance au développement. Nous ne mentionnons dans ce chapitre que les points communs aux représentations utilisées dans les trois études. Les différences seront expliquées en détail dans les chapitres décrivant ces études.

3.1.4 Discussion préliminaire

Ces deux représentations en perspective d'une même collection de documents visuels ont des propriétés géométriques différentes. Et ces différences sont susceptibles d'influencer les stratégies d'exploration visuelle, ainsi que les préférences subjectives. En effet, du fait de la concavité de la vue interne, la taille des images décroît depuis les bords latéraux de l'affichage jusqu'au centre, alors que la convexité de la vue externe fait que les images près des bords latéraux de l'écran sont de taille plus faible que celles au centre. Il convient également de noter que les

déformations ne sont pas équivalentes d'une vue à l'autre. Avec le même nombre d'images affichées simultanément, et le même angle visuel de la caméra virtuelle, la courbure diffère légèrement, avec pour conséquence des déformations non équivalentes (voir la figure 3.4). En outre, la convexité

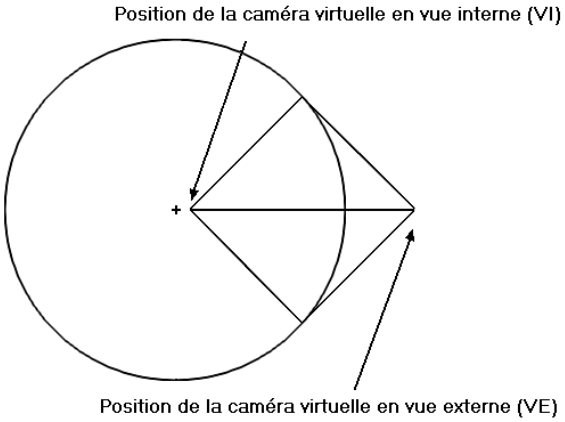


FIG. 3.4 – Non équivalence des vues interne et externe

de la vue externe est de nature à inciter les utilisateurs à envisager l'interaction comme la manipulation d'un objet virtuel, tandis que la vue interne est propre à induire une sensation d'immersion dans la représentation. Par conséquent, il nous a paru intéressant d'évaluer et de comparer l'efficacité, le confort et l'attractivité de ces deux vues qui, par leur simplicité et le caractère intuitif des métaphores d'interaction qu'elles suggèrent, sont susceptibles d'être bien accueillies par le grand public.

Le choix de ces représentations trouve également une autre motivation [Robertson *et al.*, 2000], en complément des points précédemment explicités : l'utilisateur, pour ne pas se sentir « perdu » face à une nouvelle représentation, doit pouvoir se rapprocher de concepts familiers. Dans ce but, *Task Gallery* utilise la métaphore d'une galerie d'art, censée faciliter le repérage spatial dans la visualisation en assimilant la navigation à des déplacements dans le monde réel. Les représentations que nous proposons permettent également une analogie avec le monde réel : le cylindre extérieur, par sa forme, ses propriétés et la métaphore d'interaction induite, rappelle les présentoirs de cartes postales, que l'on peut faire tourner devant soi ; le cylindre intérieur, par sa forme, suggère une sorte de carrousel, dans lequel une personne pourrait se déplacer.

D'autres avantages inhérents à ces représentations font également d'elles un choix logique et cohérent. D'une part ce type de représentation favorise une navigation et/ou une interaction plus simple qu'un environnement 3D classique dans lequel l'utilisateur dispose de tous les degrés de liberté. Ici la forme circulaire implique une rotation pour le parcours en largeur de la collection, et donc « canalise » une partie des déplacements. En 1995, beaucoup d'applications 3D ne fournissaient en réalité qu'une navigation en 2D [Plaisant *et al.*, 1995] : ce qui était la conséquence d'une limitation technique peut aujourd'hui être un choix délibéré. Les déplacements proposés sont moins riches que ceux qu'il est possible de réaliser dans un espace 3D, mais ce compromis permet de ne pas perturber les novices, tout en gardant un niveau de réalisme élevé. D'autre part, la géométrie choisie permet de conserver une présentation de structure matricielle, et ainsi

de respecter les principes de Shneiderman [Shneiderman, 2003] en organisant les groupes d'items en structures alignées. En outre, comme nous l'avons déjà signalé, les phénomènes d'occlusion courants dans d'autres types de représentations sont absents de celles que nous proposons. Enfin, un des derniers avantages, et non des moindres, concerne la généricité de notre interface : alors qu'il est difficile de faire à la fois un outil générique (du point de vue recherche) et un outil spécifique (du point de vue de l'utilisateur final), comme le souligne Catherine Plaisant [Plaisant, 2004], notre représentation peut être utilisée avec tous les types de documents visuels, peut s'intégrer non seulement dans un gestionnaire de photographies mais également dans la partie visualisation d'un explorateur de fichiers. Ajoutons enfin que la géométrie choisie, relativement simple, permettra d'isoler plus facilement les facteurs pouvant influer sur les résultats qualitatifs et quantitatifs que nous obtiendrons.

3.2 Des métaphores d'interaction adaptées

L'exposition du choix de notre représentation effectuée, nous allons maintenant expliciter les différentes fonctionnalités implémentées. Leur nombre a été volontairement limité, les utilisateurs n'utilisant que les fonctionnalités de base des logiciels pour leur usage courant [Rodden et Wood, 2003]. Notre approche étant avant tout basée sur la navigation et l'exploration visuelle, il faut fournir aux utilisateurs les moyens de parcourir la collection, d'examiner en détail un item particulier et, bien entendu, la possibilité de se repérer facilement dans la collection. Le détail de ces différentes fonctionnalités fait l'objet des sections suivantes, après une première discussion sur le choix du périphérique d'interaction.

3.2.1 Le périphérique, un choix important

La principale contrainte est de produire des moyens d'interaction fonctionnellement équivalents pour les deux vues (étude préliminaire et étude 1) ainsi que pour la représentation 2D (étude 2). Le même périphérique, quel qu'il soit (clavier, souris ou joystick par exemple), sera donc partagé pour toutes les représentations, pour ne pas créer de différence à ce niveau. Le joystick a été rapidement abandonné, n'étant pas assez communément utilisé : certains utilisateurs non familiers pourraient être déroutés par son fonctionnement. Cette dernière remarque exclut également tout périphérique peu courant ou « évolué », comme les souris à 6 degrés de liberté ou encore les tablettes graphiques. L'ajout de facteurs dont il serait difficile d'évaluer l'impact sur les résultats est bien évidemment à proscrire. Nous avons également fait le choix de n'utiliser qu'un seul périphérique, afin de simplifier au maximum la tâche de l'utilisateur, et lui permettre de focaliser son attention sur la tâche à réaliser plutôt que sur l'utilisation conjointe de deux outils distincts.

Parmi les périphériques « classiques », il reste donc à choisir entre la souris et le clavier, sachant que chacun présente des avantages et des inconvénients pour chaque vue du cylindre. Le choix de l'un de ces deux périphériques est motivé par le fait qu'ils sont assez courants pour que leur maîtrise soit acquise rapidement, quel que soit l'utilisateur. Le clavier semble être, *a priori*, plus adapté à la vue interne, puisqu'il existe une relation évidente entre les flèches direc-

tionnelles et les déplacements correspondants dans l'environnement 3D. Néanmoins, les flèches directionnelles du clavier ne permettent qu'un déplacement suivant deux dimensions, l'ajout d'une troisième dimension imposant d'utiliser des touches supplémentaires. A l'opposé, la souris est davantage en adéquation avec l'interaction de la vue externe qui implémente la manipulation directe, couramment rencontrée dans les interfaces classiques.

Nous écartons le clavier pour plusieurs raisons : d'une part, l'utilisation de touches supplémentaires laisse envisager un problème de mémorisation, à moins de réaliser un clavier spécial comportant des pictogrammes associant une touche à une action spécifique ; d'autre part, avant d'être vraiment familier avec les différentes commandes, l'utilisateur serait obligé de détourner son regard de l'écran, pour appuyer sur la touche correcte. Les expérimentations présentées ci-après, par leur durée limitée, ne permettent pas à l'utilisateur de se familiariser avec l'utilisation des touches au point de ne plus avoir à regarder le clavier. Ces raisons nous ont poussé à choisir la souris comme seul et unique périphérique disponible dans les deux métaphores.

Nous allons maintenant présenter les différentes fonctions d'interaction et de navigation mises à la disposition des utilisateurs. Etant donné que l'implémentation de ces fonctions a évolué d'une expérimentation à l'autre, à la faveur des commentaires des sujets, ces fonctions et les interacteurs permettant de les activer seront décrits dans les chapitres suivants, consacrés à chaque expérimentation.

3.2.2 Le déplacement de/dans la représentation

Un environnement 3D, qu'il soit de nature immersive ou non, permet l'usage des 6 degrés de liberté, contrairement aux représentations 2D qui sont contraintes à des déplacements dans le plan. Même si l'on peut considérer que la modulation du facteur de zoom (par exemple dans le cas de la visualisation d'un document textuel) peut s'apparenter à la notion de profondeur. Pourtant, donner à l'utilisateur le pouvoir d'un déplacement sans contrainte impose un facteur de risque important. Nous souhaitons nous adresser à la plus grande majorité des utilisateurs, dont le degré de familiarité avec les environnements 3D nous est inconnu. Une interaction trop riche serait donc très difficile à appréhender, et rendrait la navigation dans la collection laborieuse et hasardeuse. D'autant que l'usage d'une souris impose une simplification des mouvements possibles. Il nous semble donc judicieux de « canaliser » les mouvements des utilisateurs afin de garantir une certaine simplicité d'usage, tout en conservant les avantages de l'existence de la troisième dimension. Ainsi, nous avons délibérément ignoré tout déplacement peu naturel, c'est-à-dire n'étant pas en adéquation avec les métaphores d'interaction.

Etant donnée la forme singulière de la représentation, le premier déplacement envisagé est la rotation du cylindre autour de son axe. Ce défilement latéral, naturel, est analogue au défilement horizontal avec les ascenseurs, tel que l'on peut le trouver dans les interfaces courantes ; voir, par exemple, la partie droite de l'explorateur de fichiers (sous *Windows*), ou le parcours d'un document *pdf* ou *word*.

Il est également possible de moduler la distance entre la caméra virtuelle et la représentation. Ce type de déplacement présente plusieurs intérêts : d'une part, dans le cadre des expérimentations, nous sommes obligés de fixer le nombre d'images pour que « toute chose soit égale par

ailleurs ». Ce qui implique une disposition identique, ainsi qu'un nombre rigoureusement identique de documents visibles initialement, au moins au début de l'expérimentation. Pourtant, les utilisateurs ne sont pas égaux sur le plan des capacités visuelles, et ils peuvent avoir des préférences différentes concernant le nombre d'items affichés simultanément. Les deux vues 3D offrant un *focus + context* naturel, nous souhaitons conserver cette caractéristique, tout en répondant aux deux contraintes énoncées ci-avant. C'est pourquoi ce déplacement est utile, en permettant de modifier simultanément et dynamiquement la taille des images et leur nombre à l'écran, via un mouvement facilement compréhensible par l'utilisateur. Cela rapproche notre interface d'une ZUI (*Zoomable User Interface*) au sens décrit par Combs et Bederson [Combs et Bederson, 1999].

Le déplacement vertical, potentiellement utile lorsque la distance entre la représentation et la caméra virtuelle implique de tronquer une partie de l'affichage verticalement, peut également s'avérer utile dans le contexte des études que nous avons réalisées. Car la solution de modifier dynamiquement la disposition (c'est-à-dire diminuer le nombre de lignes et augmenter la nombre de colonnes) n'est pas acceptable, ayant pour effet de modifier la perception de l'environnement 3D de l'utilisateur, et de lui faire perdre les repères acquis lors de la navigation, aussi bien sur la longueur de parcours que sur la mémorisation des items survolés, et de leurs positions relatives.

D'autres mouvements de libre rotation ont été éliminés : celui décrivant le tangage, c'est-à-dire un mouvement de basculement avant-arrière, ainsi que celui décrivant le roulis, c'est-à-dire un mouvement de basculement latéral. Ces mouvements ne sont pas naturels pour l'observateur et risqueraient d'introduire une certaine forme de confusion. Le périphérique retenu ne permet d'ailleurs pas la gestion simple de ce type de déplacement.

3.2.3 Le dispositif de zoom

L'utilisateur doit pouvoir modifier le niveau de détail d'un item de la collection, pour réaliser par exemple un examen plus approfondi. La modulation de la distance entre la caméra virtuelle et le cylindre rend possible cette modulation du niveau de détail. L'utilisation d'un dispositif de zoom « classique » permet un accès rapide à une visualisation de grande définition d'un document visuel particulier. Ces deux solutions, qui ne sont pas totalement équivalentes, sont à mettre en relation avec la métaphore implémentée : la seconde solution est plus en accord avec la manipulation, tandis que la première étant plus fidèle à la notion de locomotion.

3.2.4 Le dispositif de repérage

Dans une visualisation classique, dans laquelle les items sont disposés sous forme matricielle, et dont la taille dépasse celle de l'écran, la navigation se fait au moyen d'un (ou deux) ascenseur(s). En plus de servir d'interacteur, l'ascenseur (ou *scrollbar*) fournit deux informations supplémentaires : la position de celui-ci par rapport à ses bornes indique la position courante, tandis que sa taille indique le ratio entre taille totale de la visualisation et partie affichée. De cette manière, l'utilisateur peut à chaque instant connaître la position de la caméra virtuelle par rapport au document.

Dans notre représentation, nous avons choisi de laisser une « colonne » vierge d'images, comme

indicateur d'un parcours complet de la collection. Cet indicateur fournit une information ponctuelle sur la position de l'utilisateur, mais est notoirement insuffisant pour fournir une information précise à chaque instant. Il faut proposer un dispositif supplémentaire permettant un repérage précis en temps réel. La principale difficulté réside dans le fait de trouver une adéquation entre le dispositif et la représentation cylindrique, donc continue. Un ascenseur serait ici totalement inadapté, puisqu'il serait incapable de traduire cette continuité, et trahirait d'une certaine façon l'apport de la troisième dimension. L'adoption d'une mini-représentation homothétique à la représentation réelle est de ce fait la plus appropriée, pour donner de façon claire et non équivoque la position de l'utilisateur dans le cylindre réel. Une contrainte supplémentaire est que la mise à jour de la position doit tenir compte de la métaphore : la vue intérieure du cylindre doit favoriser l'impression de déplacement dans la collection, tandis que la vue extérieure doit se rapprocher de la métaphore de la manipulation et donner à l'utilisateur l'impression qu'il déplace l'objet virtuel représentant la collection. Ainsi, il faut que, dans le premier cas, la caméra virtuelle soit en mouvement, la mini-représentation restant fixe, alors que, dans le second cas, le point de vue ne doit pas changer, c'est la représentation qui se déplace.

3.2.5 Bref résumé

Nous souhaitons proposer à l'utilisateur les fonctionnalités les plus simples possible, le but recherché étant de permettre une navigation rapide et intuitive, nécessitant un apprentissage quasi-inexistant, afin de satisfaire un large public, allant des novices aux utilisateurs expérimentés. C'est pourquoi nous avons limité la nature des mouvements possibles : le défilement vertical et/ou horizontal des interfaces classiques se retrouve ici dans le mouvement de rotation du cylindre autour d'un axe vertical, que nous pourrions qualifier de défilement latéral/horizontal. La modulation du niveau de zoom global (rapport entre le nombre d'images visibles et la taille de celles-ci), proposé dans tous les logiciels classiques, se retrouve de manière naturelle avec le changement de la distance entre la caméra virtuelle et le cylindre. Dispositif d'agrandissement et repérage spatial sont également adaptés à notre représentation 3D. Il n'a pas paru opportun d'ajouter d'autres fonctionnalités, pour que la plate-forme de visualisation garde un caractère générique. En fonction des attentes des utilisateurs, les concepteurs pourront ajouter des fonctionnalités *ad-hoc* à leur gré, mais cela dépasse de loin le cadre de notre étude.

3.3 Discussion

La représentation cylindrique se rapproche du mur en perspective [Mackinlay *et al.*, 1991], du moins dans la vue externe, et offre une visualisation de type *focus + context* naturelle. Au final, la seule déformation polaire d'une représentation matricielle classique permet d'avoir une représentation adaptée aux données structurées de manière linéaire suivant une ou deux dimensions. *A contrario*, elle est inadaptée à des données structurées de manière hiérarchique. Cela étant, nous avons démontré que l'usage de données non structurées, ou structurées linéairement, est finalement assez courant lors de la gestion de collections de documents visuels, que ceux-ci soient homogènes (une série de photographies appartenant à un même thème) ou non (les

résultats d'une requête, pouvant comprendre des documents visuels de nature différente). La nature des déformations appliquées aux images offre des avantages non négligeables : lorsque les utilisateurs tournent le cylindre, leur attention visuelle est naturellement canalisée sur quelques colonnes, en fonction des déformations (dues à la perspective) propres à chaque vue. Il a été démontré, du moins pour des affichages 2D de petites collections non structurées de photographies (30 par affichage), que certaines structures (elliptiques, radiales) permettent d'augmenter le confort et l'efficacité de la recherche de photographies visuellement familières ; elles guident, ou tout du moins canalisent, l'exploration visuelle, ce qui induit des parcours visuels plus courts et plus efficaces car plus systématiques [Simonin *et al.*, 2005]. Les représentations en perspective permettent en outre une visualisation multi-échelle régulière et continue, de par les effets de réduction et d'augmentation de taille durant la rotation : lorsque l'utilisateur poursuit un document visuel pendant la rotation, la précision de l'affichage évolue comme sous l'effet d'un zoom (avant/arrière) continu. La figure 3.5 permet de voir de manière schématique l'évolution de la taille d'une image, d'un bord à l'autre de l'écran, dans la vue externe. L'intérêt des visualisa-



FIG. 3.5 – Evolution de la taille d'un document

tions multi-échelles pour l'exploration et la recherche dans de grandes collections de documents a été démontré depuis longtemps. On peut par exemple se référer au *Visual Information Seeking Mantra* [Shneiderman, 1996]. Précisons que notre représentation, par sa géométrie, rend impossible la visualisation de la collection dans son intégralité. Mais dans le cas d'une collection non structurée, afficher 1000 documents visuels sur un écran d'une définition de 1280 par 1024 pixels impliquerait une taille d'environ 30 pixels de côté pour chaque document, ce qui est trop faible pour en discerner distinctement l'objet principal (figure 3.6).



FIG. 3.6 – Taille réelle sur un écran de 19 pouces (1280*1024) d'une image de 30 pixels (un veau couché sur de la paille)

La représentation que nous proposons est plus performante que les visualisations 3D positionnant les miniatures dans différents plans de profondeur différente, au sens où nous pouvons afficher davantage de documents visuels simultanément. Plaquer les miniatures sur la surface latérale d'un cylindre permet d'éliminer le phénomène d'occlusion que l'on peut observer par exemple dans *Data Mountain*, dans *Simple Landscape* ou *PhotoGoRound* de *TriVista Corporation* [Combs et Bederson, 1999]. Cette dernière représentation a une structure très proche du cylindre en vue externe, mais le choix de garder les documents visuels toujours face à l'utilisateur implique une forte occlusion qui annihile l'intérêt de ce type de représentation.

Comparativement au mur au perspective, la vue externe du cylindre se trouve être visuellement plus confortable, du fait de la surface continûment incurvée, alors que les angles résultant du contour en U du mur en perspective risquent d'interrompre la continuité du parcours visuel. Les

vues comparables à la vue interne du cylindre se retrouvent moins fréquemment dans les travaux scientifiques : cela peut s'expliquer parce que, contrairement aux vues classiques de type *focus + context*, les zones agrandies de la représentation sont rejetées sur les bords, et non situées au centre de l'écran. Une autre visualisation immersive similaire à la vue interne a été proposée récemment, dans un contexte différent, celui des bibliothèques numériques. Almeida *et al.* [Almeida *et al.*, 2006] rapportent la conception et l'implémentation d'un environnement 3D pour naviguer et lire plusieurs documents numériques simultanément. Les utilisateurs sont immergés dans une salle de lecture ; en arrière-plan, un mur semi-circulaire d'étagères comprend des collections de livres, et au premier plan sont présentés verticalement les livres en cours de consultation. L'utilisateur peut prendre des livres dans la bibliothèque et les consulter (tourner les pages). Mais il n'est pas possible de faire « tourner » le mur, et les étagères peuvent seulement être zoomées et dézoomées. Cet environnement a d'abord été implémenté sur un écran standard [Cubaud *et al.*, 2002] mais les auteurs travaillent actuellement à une version améliorée sur affichage hémisphérique *VisionStation* qui n'utilise plus la métaphore du trépied de la version initiale. Cette approche diffère du courant principal de la recherche sur les visualisations destinées aux bibliothèques virtuelles, qui se centrent sur la conception de métaphores de livres en 3D [Card *et al.*, 2004 ; Chu *et al.*, 2004].

3.4 Applications

Une mesure fréquente, pour l'évaluation d'une nouvelle technologie, est son adoption par d'autres, et son utilisation dans des produits commerciaux [Plaisant, 2004]. Bien entendu, cette adoption ne garantit en rien la qualité ergonomique de ces technologies, encore moins un niveau d'efficacité ou de confort d'utilisation. Toutefois, la large diffusion d'une idée, ou d'un concept, peut être une aide certaine dans la phase de maturation de la conception d'un logiciel, et permettre ainsi de nombreuses améliorations et éventuellement une orientation vers d'autres choix, différents ou voisins.

Dans l'état actuel, la représentation que nous présentons, associée à ses deux implémentations, est restée au stade de simple prototype de recherche. Mais les applications sont potentiellement grandes, si tant est que nous puissions attester de la supériorité des vues 3D sur les classiques interfaces 2D qui sont aujourd'hui les plus courantes. La première application reste naturellement l'évolution à court terme des navigateurs de fichiers intégrés directement dans les systèmes d'exploitation. Cela est techniquement possible, la majeure partie des ordinateurs récents étant capables de gérer des « bureaux » 3D. Et l'usage n'est pas limité aux stations personnelles, les bornes interactives peuvent également bénéficier de ce concept, car notre but est de produire une représentation qui soit assez générique pour correspondre à toutes sortes de documents visuels, avec une interaction qui puisse convenir à un large public par sa simplicité et son intuitivité. Il reste également à évaluer si une vue est meilleure que l'autre, où si les deux doivent être proposées conjointement aux utilisateurs.

Remarquons enfin, même si cela apparaît assez gratuit, qu'une des implémentations de notre représentation rencontre un succès d'estime chez les publicitaires, ce qui témoigne de son caractère esthétiquement attrayant : ainsi, la vue intérieure du cylindre à été utilisée dans deux publicités

en 2007, l'une d'*Apple* (voir figure 3.7) et l'autre de *Club Internet* (voir figure 3.8).



FIG. 3.7 – Publicité pour le boîtier *AppleTV*, Apple, 2007



FIG. 3.8 – Publicité pour une offre de *VOD*, Club Internet, 2007

Chapitre 4

Méthodologie expérimentale

Dans ce chapitre, nous présentons les éléments communs au protocole expérimental des trois expérimentations que nous avons menées pour comparer et évaluer les deux vues proposées, et donc pour valider nos hypothèses de travail. Après la présentation des tâches de recherche visuelle, nous détaillerons la réalisation du matériel expérimental, le choix des participants et enfin le déroulement des passations. Le recueil des données, les éléments différant d'une expérimentation à l'autre et la présentation des résultats et de leur analyse seront développés dans les chapitres suivants, propres à chaque expérimentation.

4.1 Tâches de recherche visuelle

4.1.1 Généralités

Nous avons insisté dans le second chapitre sur la motivation du choix de la visualisation interactive de documents visuels comme objet d'étude : ce sont les types de documents les plus couramment manipulés par le grand-public, ainsi que par de nombreux professionnels et experts. L'explosion des loisirs numériques et, en particulier, la démocratisation des appareils de photographie numérique, ainsi que la simplicité d'échange apportée par les grands réseaux de partage que sont par exemple *Yahoo Flickr* ou *Google Picasa Web Album* contribuent à cet état de fait. La recherche par mots-clés, qui se fonde sur une indexation (automatique ou non) n'est pas la solution la plus adaptée, et surtout impose de nombreuses contraintes. Le recours à la navigation et à l'exploration de grandes collections, structurées ou non, reste incontournable, pour autant que nous soyons capables de proposer des interfaces simples, intuitives, confortables et performantes.

Afin d'évaluer l'apport de la représentation que nous proposons, il est nécessaire de concevoir des tâches réalistes et simples, dont le scénario tend à se confondre avec les actions fréquemment réalisées par les utilisateurs lors de l'interaction avec de grandes collections de documents visuels. Parmi les différents types de tâches réalisables avec un logiciel de gestion de collections de photographies, la taxonomie proposée dans [Plaisant *et al.*, 1995] en considère deux, entre autres, qui sont particulièrement fréquentes :

- l'exploration : l'espace de recherche est inconnu de l'utilisateur, qui l'explore pour la première fois. Ainsi, celui-ci ne dispose pas d'une image mentale de l'espace, et n'a pas de points de repère *a priori* ;
- la navigation : l'environnement est connu de l'utilisateur de manière plus ou moins précise, et ce dernier peut ainsi se repérer d'après les différents objets de cet environnement, pour atteindre l'un d'entre eux situé en un point particulier de l'espace de recherche.

Ces deux types de tâches, présentés ici de manière générique, se rapprochent de deux des trois scénarios envisageables dans le cas de la manipulation de grandes quantités d'informations :

- l'information recherchée est connue, c'est-à-dire familière visuellement et on connaît sa position, à savoir le chemin pour y accéder. C'est le cas, par exemple, lors de la recherche d'un document connu dans une arborescence globalement familière ou, dans le cas de données non structurées, lorsque le souvenir de plusieurs points de repère entre le point de départ et la position du document permet de retrouver ce dernier ;
- l'information recherchée est connue, mais sa position est ignorée. Pour garder le même type d'exemple, cela se produit lors de la recherche d'un document connu, mais dont la position est inconne (ou oubliée) ;
- l'information recherchée est non familière visuellement, et sa position naturellement inconnue. Nous pouvons citer l'exemple la recherche d'une image vérifiant plusieurs critères dans une collection de documents visuels : je recherche la photographie d'un coucher de soleil sur une mer calme, parmi des photographies de vacances qu'un tiers m'a données.

Ainsi, le premier scénario implique plutôt une tâche de navigation, car l'unique but dans ce contexte est de retrouver un document particulier, sans prendre la peine d'observer tous les autres documents, mais seulement quelques documents clés. Le dernier scénario entraîne quant à lui une exploration minutieuse de la collection. Le deuxième scénario est à la frontière entre ces deux types de tâches, au sens où il mêle navigation et exploration.

La navigation et l'exploration de grandes collections de documents visuels sont des tâches fréquentes et réalistes, qui possèdent en outre un intérêt supplémentaire, celui de ne pas faire intervenir les mêmes caractéristiques fonctionnelles. L'exploration demande une interface simple, avec la possibilité de zoomer sur des documents particuliers, tout en maintenant visible le contexte, de façon à créer facilement des points de repère. La navigation suppose une niveau de détail moindre, le recours au zoom étant moins fréquent, et la vitesse de défilement souvent plus élevée. Les représentations que nous proposons ont pour objectif de pouvoir concilier à la fois navigation et exploration.

Nous avons retenu les scénarios les plus contrastés (le premier et le dernier), car ce sont ceux qui privilégient le plus nettement l'une des deux formes d'interaction possibles avec des masses de données, l'exploration pour le premier, la navigation pour le second. Ce qui permettra de mettre en exergue l'influence du type de tâche sur la représentation, si elle existe.

Concernant les documents visuels, nous faisons le choix d'utiliser des photographies, car ce sont, parmi les documents visuels, ceux les plus couramment utilisés, ceux que l'on trouve en plus grand nombre, aussi bien localement sur nos disques durs, qu'à distance via Internet. De plus, il y a une adéquation entre les scénarios choisis et les photographies, car ils correspondent parfaitement à la manipulation de ce type de documents. Nous sommes d'accord avec [Plaisant,

2004] sur l'importance du choix d'utiliser de grands ensembles de données réalistes et de proposer sur eux des tâches réalistes. Il est certain qu'utiliser des objets abstraits, comme par exemple des rectangles de couleur [Guiard *et al.*, 2006] permet de contrôler plus facilement tous les facteurs susceptibles d'influencer les comportements et les performances des participants. En particulier, cela permet d'éliminer l'influence de la variabilité inter-individuelle quant au références culturelles associées au contenu des photographies. Mais l'expérimentation avec des objets abstraits suppose la réalisation de tâches artificielles, dont les résultats seront difficilement interprétables, notamment en termes de recommandations de conception, et qui donneront lieu à des jugements subjectifs dénués d'intérêt pratique. Notre choix impose, certes, une conception rigoureuse du protocole expérimental, difficile à réaliser, pour supprimer, ou tout au moins atténuer au maximum, l'influence de facteurs non contrôlables. La contrepartie évidente sera que les résultats des études expérimentales présentées dans ce mémoire pourront être facilement applicables à la conception de logiciels de gestion de collections de photographies, et de manière plus large, de collections de documents visuels.

Dans le reste de ce document, nous nommerons **T1** et **T2** les deux types de scénarios de tâches retenus :

- T1 : recherche d'une photographie correspondant à des critères thématiques et visuels spécifiques, c'est-à-dire recherche d'une photographie non familière, dont la position au sein de la représentation est inconnue ;
- T2 : recherche d'une photographie familière, dans une collection familière de photographies, c'est-à-dire l'action de retrouver une photographie visuellement familière, dont la position dans la représentation a été mémorisée lors d'interactions précédentes avec la représentation.

Les tâches T2 sont appropriées pour la comparaison de la possible influence de la représentation visuelle sur la mémoire spatiale, et donc pour évaluer les hypothèses B et C (voir chapitre 2). Les deux tâches (T1 et T2) sont à même de tester la validité des autres hypothèses de recherche (A et D). Néanmoins, l'hypothèse A n'a été évaluée que dans la seconde expérimentation, pour T2, à l'aide de mesures oculométriques (utilisation d'un *eye-tracker*).

Les tâches T1 et T2 sont plus réalistes que celles présentées, par exemple, dans les études de Tavanti et Lind [Tavanti et Lind, 2001], puis de Cockburn [Cockburn, 2004], où les auteurs utilisent des ensembles composés respectivement de lettres de l'alphabet suédois et de drapeaux de pays. Les tâches T1 et T2 sont d'autre part plus simples que celles étudiées dans [Robertson *et al.*, 1998], où les auteurs utilisent des miniatures de pages web en tant que *bookmarks*. Par conséquent, nos études seront plus à même de fournir des résultats d'une portée scientifique plus significative que ceux décrits dans ces études. La simplicité et le réalisme des tâches de recherche conduiront à des analyses, des interprétations et des conclusions plus fiables et directement applicables à la conception des visualisations interactives.

4.1.2 Tâche T1 : conception

Description de la tâche

De manière pratique, la tâche T1 consiste en la recherche de plusieurs photographies spécifiques, d'après une description verbale. Afin qu'un participant puisse être capable de rechercher une photographie qu'il n'a jamais vue, il faut qu'il soit en mesure de se créer sa propre image mentale de celle-ci. Le recours à une description verbale se révèle donc être l'unique solution. Cela permet en outre de se rapprocher de situations réalistes, dans lesquelles un utilisateur peut être amené à chercher un document visuel dans une collection, d'après une description donnée par un tiers.

L'élaboration des descriptions est une étape critique, un point sensible du protocole expérimental, des études antérieures ayant montré la difficulté de cette entreprise [Suh *et al.*, 2003] : certains sujets avaient en effet des difficultés à trouver une cible, car ils avaient des présupposés erronés sur celle-ci. Pour expliquer ce phénomène, les auteurs parlent d'*attentional control settings* [Folk *et al.*, 1992]. Dans sa taxonomie des différents types de données, Shneiderman [Shneiderman, 1996] insiste d'ailleurs sur la complexité de la rédaction d'une description textuelle, en rappelant qu'une photographie est souvent citée comme valant mille mots. Nous avons opté pour des descriptions très détaillées, afin de limiter autant que faire se peut les problèmes rencontrés par Suh *et al.* Les descriptions verbales sont proposées aux participants sous forme textuelle, plutôt qu'orale, pour que l'utilisateur ait le temps de mémoriser la description à son rythme et de se créer une représentation mentale fidèle. Nous avons en outre fait le choix de ne pas afficher les descriptions à l'écran car les expérimentations sont relativement longues, et cela ne ferait qu'accroître la fatigue visuelle. Il reste plus agréable et confortable de travailler « sur papier » pour les informations verbales. Cela permet de plus aux sujets de consulter librement et facilement les descriptions au cours de la recherche.

Description d'une tâche de type T1

Les sujets doivent rechercher dans chaque collection de photographies, une photographie particulière - que nous nommerons **cible** dans la suite de cette présentation - décrite verbalement. Les descriptions correspondent de manière non ambiguë à une seule et unique photographie de la collection. Les descriptions sont toutes constituées de 4 à 6 phrases. Elles sont de plus structurées suivant un même schéma. Celui-ci ayant évolué entre l'étude préliminaire et la seconde étude, il sera donc détaillé dans les chapitres correspondants. Naturellement, les collections et les cibles varient d'une recherche à l'autre.

Après la lecture de la description de la cible à chercher (le temps est illimité), les participants commencent la recherche, limitée en temps. Cette limite a été imposée, d'une part pour conserver un temps total d'expérimentation raisonnable, et d'autre part pour éviter tout découragement entraînant une frustration du sujet. Dans ce premier type de tâche, il a été demandé aux sujets de se concentrer davantage sur la précision que sur la rapidité. Lorsque le sujet pense avoir trouvé la cible, il la sélectionne et clique sur un bouton prédéfini pour indiquer au système la fin de la recherche. A ce moment, ou bien lorsque le temps imparti est dépassé, donc quel que soit le

résultat de la recherche, une courte animation se lance afin de permettre au sujet de visualiser la cible qu'il fallait trouver, ainsi que sa position sur la représentation. Ainsi, tous les participants ont l'opportunité de se familiariser avec les caractéristiques visuelles de la cible, ainsi qu'avec sa position dans la représentation. Cette animation se déroule de la manière suivante : la vue initiale de la collection est affichée, puis la représentation défile automatiquement jusqu'à ce que la cible apparaisse sur la colonne centrale de l'écran, puis la cible est agrandie. L'animation emprunte le chemin le plus court pour atteindre la cible, c'est-à-dire le sens de rotation qui permet d'atteindre celle-ci le plus rapidement.

Le sujet est alors invité à lire la description de la cible suivante, et ainsi de suite, pour toutes les cibles d'une vue, puis pour chaque vue (3D Interne et 3D Externe dans l'étude préliminaire et la première étude, et 3D Interne, 3D Externe et 2D dans la seconde étude). Le sujet se trouve alors dans les conditions requises pour réaliser la tâche 2.

4.1.3 Tâche T2 : conception

Toutes les recherches de cibles de la tâche T2 sont réalisées après celles de la tâche T1, et consistent à retrouver les mêmes cibles dans les même collections que celles utilisées dans la tâche T1. Cela assure que les cibles dans la tâche T2 sont visuellement familières à tous les participants, même dans le cas de recherches infructueuses lors de T1, et qu'il en est de même pour leurs positions dans les collections. Pour indiquer au sujet la cible à rechercher, celle-ci est affichée isolément au centre de l'écran durant 3 secondes. Cette durée a été fixée empiriquement, d'après des études antérieures menées au sein de l'équipe. Ensuite, la collection correspondante s'affiche, et la recherche visuelle peut alors commencer. L'enchaînement est le même que dans les recherches de la tâche T1, excepté l'animation finale, qui est inutile dans le contexte de la tâche T2. L'ordre de présentation des cibles au cours de T2 est différent de celui adopté pour T1 afin d'éviter que les effets de la mémoire temporelle (mémoire épisodique) n'interfèrent avec ceux de la mémoire spatiale.

4.2 Matériel visuel expérimental

4.2.1 Choix du matériel visuel

Les photographies utilisées, en couleur, ont été collectées sur Internet dans des bases de données libres de droits, et classées suivant des thèmes très variés, comme le sport, les fleurs, les voitures, les animaux, etc. Chaque thème comprend une vingtaine de photographies. Une distinction est faite entre les photographies représentant des paysages (paysages urbains et naturels, scènes d'intérieur), et les photographies représentant des objets. Cette distinction est basée sur des résultats expérimentaux qui suivent le modèle de perception visuelle *Coarse-to-Fine* [Parker *et al.*, 1992 ; Schyns et Oliva, 1994]. En particulier, ce modèle montre l'influence de la fréquence spatiale dans l'organisation temporelle de l'exploration de la scène et du processus de reconnaissance. La distribution des fréquences spatiales hautes et basses dans les photographies de paysages et les photographies d'objets est différente, car les photographies de paysages incluent

plus de de détails que les photographies d'objets communs ou les portraits. Il a de plus été démontré que retrouver des photographies familières dans de petites collections (30 items) sur un affichage 2D statique est significativement plus long pour des photographies de paysages [Simonin *et al.*, 2005]. C'est pourquoi nous classons les photographies d'animaux, ainsi que les portraits, dans la catégorie des « objets » , comme les auteurs de [Simonin *et al.*, 2005].

Pour l'étude préliminaire, ainsi que pour la première étude, nous disposions de 3 ensembles de 40 thèmes, composés de 22 photographies chacun, plus un ensemble de 900 photographies utilisées pour les tâches de recherche d'entraînement qui permettaient au sujet de prendre en main le logiciel proposé. Soit un total de 2680 photographies, plus 900. Dans la seconde étude, la base de données a été élargie, et nous disposions de 4 ensembles de 44 thèmes, composés de 20 photographies chacun, plus un ensemble de 880 photographies pour les tâches de recherche d'entraînement. Soit un total de 3520 photographies (plus 880), ce qui nous a permis de diversifier davantage les collections. La figure 4.1 présente une série de photographies sur le thème des feux d'artifice. Pour les 3 études, les photographies ont toutes été redimensionnées pour correspondre



FIG. 4.1 – Photographies sur le thème des feux d'artifice

à un ratio (largeur/hauteur) de 4/3. La définition choisie a varié suivant les études (de 128x128 pixels à 512x384 pixels), en fonction des contraintes techniques, le but étant d'obtenir la plus grande qualité d'affichage possible tout en gardant des temps de chargement acceptables et une fluidité suffisante. La définition des vignettes affichées est cependant toujours la même, soit 72 pixels par pouce, de même que la profondeur des couleurs (24 bits).

4.2.2 Choix des cibles

Pour les 3 études, nous avons choisi « manuellement » les différentes photographies servant de cibles. La méthodologie a été la suivante : nous avons parcouru les différentes séries de photo-

graphies, à la recherche de photographies paraissant simples à décrire, et présentant une saillance visuelle prononcée, gage d'un examen rapide de la collection et ce faisant d'une recherche plus efficace. La saillance visuelle pouvant s'exprimer sous la forme de plusieurs critères, tels que la couleur, le cadrage, le nombre d'objets, la forme et la direction [Itti et Koch, 2000]. La seconde contrainte est l'équi-distribution des cibles entre la catégorie des objets et celle des paysages.

Pour l'étude préliminaire, une première sélection de cibles potentielles a été effectuée, en essayant d'obtenir le même nombre de cibles appartenant à un même thème, afin qu'un thème ne soit pas sous- (ou sur-) représenté. Après avoir rédigé la description verbale des photographies sélectionnées, nous avons exclu celles dont les descriptions manquaient de clarté, et qui ne permettaient pas la constitution d'une image mentale fidèle. Nous avons ensuite soumis les descriptions à plusieurs personnes, pour les valider, c'est-à-dire pour gommer au maximum les références culturelles, et supprimer tout élément ambigu, en particulier au niveau des formes et des couleurs. Pour les deux études suivantes, certaines cibles de l'étude préliminaire ont été supprimées suite aux difficultés rencontrées par les sujets pour les identifier. D'autres cibles ont été choisies, et le même processus a été réitéré pour élaborer les descriptions des cibles.

4.2.3 Constitution des collections

Chaque ensemble de photographies utilisé pour la recherche d'une cible, sera appelé **collection** dans la suite du document. Dans les 3 études présentées, une collection pour la représentation 3D (vue interne ou vue externe) comprend 869 photographies, et dans la seconde étude, une collection pour la représentation 2D comprend 858 photographies. Chaque collection (une par cible à chercher) comprend bien entendu la cible à chercher, ainsi que toutes les autres photographies appartenant au même thème. Les autres photographies sont choisies aléatoirement dans la base de données, de manière à équilibrer le nombre d'objets et le nombre de paysages. Une cible apparaît uniquement dans la collection où elle sert de cible. En raison du nombre total de photographies qui, bien qu'important, s'avère insuffisant pour obtenir des collections entièrement différentes pour les tâches de type T1, nous avons choisi de renouveler partiellement les photographies d'une collection à l'autre en arrangeant différemment les photographies conservées.

4.3 Choix des participants

8 personnes ont participé à l'étude préliminaire, 20 à la première étude et à la seconde étude. Pour chaque étude, nous avons équilibré le nombre de participants entre hommes et femmes. Tous les participants retenus étaient bien entendu volontaires, afin que les expérimentations se déroulent dans des conditions normales, en termes d'intérêt et de motivation. Tous les participants sont des utilisateurs expérimentés d'ordinateurs, avec un niveau d'études au minimum BAC+2, majoritairement dans le domaine de l'Informatique. La grande majorité des participants faisaient partie du LORIA et étaient à ce titre doctorants, post-doctorants ou stagiaires. Nous avons choisi de restreindre la plage d'âge, afin de limiter l'hétérogénéité des résultats en termes de performances et de préférences, certaines études mettant en lumière un impact de l'âge sur

les performances motrices des utilisateurs dans les tâches visuelles [Dollinger et Hoyer, 1996]. Egan [Egan, 1988] a également montré comment les différences individuelles peuvent affecter les performances. Borgman [Borgman, 1989] s'est intéressé plus spécifiquement aux tâches de recherche d'information pour lesquelles il a mis en évidence un effet de l'âge sur les performances.

Tous les participants ont effectué un test de vision, réalisé avec le matériel BIOPTOR (TM. Reg. US. Pat. Off ; by Stereo Optical Co., Inc). Celui-ci s'appuie sur les tests suivants :

- **vision de loin**
  - fusion verticale : ce test mesure le positionnement vertical des yeux lorsqu'il n'y a pas de *stimuli* similaires permettant de les maintenir en fusion ;
  - fusion latérale : ce test mesure le positionnement latéral des yeux lorsqu'il n'y a pas de *stimuli* similaires permettant de les maintenir en fusion ;
  - fusion centrale : ce test mesure la capacité d'une petite zone centrale de l'œil à obtenir la fusion de 2 *stimuli* disjoints, sans l'assistance de la vision périphérique (contrairement aux deux tests précédents).
  - ce test mesure la capacité du sujet à reconnaître des lettres (lettres de Snellen) de différentes tailles pour l'œil gauche, l'œil droit et les deux yeux ;
  - deux tests mesurent la capacité du sujet à juger les distances relatives en profondeur, lorsque tous les indices visuels, exceptée la disparité binoculaire, sont éliminés ;
  - discrimination des couleurs : ce test, construit grâce à 4 reproductions précises des assiettes Pseudo-Isochromatiques d'Ishahari, permet de mesurer l'aptitude du sujet à discriminer les couleurs.
- **vision de près**
  - fusion latérale ;
  - fusion centrale ;
  - acuité visuelle.

Tous ces tests portent sur la vision binoculaire, excepté les tests d'acuité visuelle, qui portent également sur la vision monoculaire. Le test BIOPTOR étant effectué dès le début de l'expérimentation, il permettait d'exclure les candidats daltoniens ou ceux ayant de grandes difficultés visuelles. Cela n'a, du reste, jamais été le cas, tous les candidats ayant une vision normale.

4.4 Déroulement des expérimentations

Pour les 3 expérimentations, le déroulement a été similaire, avec l'enchaînement suivant :

1. test visuel : comme cela a été expliqué dans le paragraphe précédent, ce test a pour premier objectif d'exclure tout participant n'ayant pas une vision « normale ». Le second objectif est de nous fournir des données plus fines quant aux aptitudes visuelles de chacun, pour une mise en relation éventuelle avec les préférences et les stratégies d'exploration visuelle.
2. questionnaire initial : cette consultation est la porte d'entrée de notre expérimentation, elle comprend quelques phrases expliquant son but général et nos motivations. Il permet de recueillir des informations générales sur le profil des participants, tels leur niveau de culture générale (diplôme en cours de préparation) et leur niveau de connaissances en Informatique : systèmes d'exploitation et périphériques utilisés, et type d'activité couramment pratiquée :

Internet, logiciels grand public, logiciels graphiques, programmation. Ce questionnaire est reproduit en annexe.

3. présentation de l'expérimentation, puis de la nature du « travail » à réaliser : différentes représentations utilisées (2 dans l'étude préliminaire et la première étude, 3 dans la seconde étude) et tâches de recherche proposées. Les consignes particulières sont également exposées : en particulier, l'attention du sujet est attirée sur le fait que, bien que le temps soit limité, il faut viser la précision de recherche plutôt que la rapidité dans T1.
4. début des tâches de recherche : le sujet commence par une des vues, dont le fonctionnement lui est expliqué dans le détail. Puis la structure des descriptions est présentée, et le sujet peut s'entraîner sur deux recherches de cibles sans enjeu, pour qu'il puisse bien comprendre le travail à réaliser. Les différentes collections correspondant aux tâches de type T1 pour cette vue sont présentées et le sujet effectue les recherches de cibles dans ces collections à partir de leurs descriptions textuelles. L'autre (ou les deux autres) vue(s) sont ensuite expérimentées successivement pour T1 ; à chaque nouvelle vue, des explications et un court entraînement précèdent la réalisation des tâches de type T1. A l'issue des recherches portant sur la tâche 1, le but de la tâche 2 est rappelé, et le sujet effectue toutes les tâches de recherche de cibles de type T2, les vues lui étant présentées dans le même ordre que pour T1. Dans la seconde étude, les participants sont équipés d'un *eye-tracker* à optique de tête pendant la réalisation des tâches T2. En dehors des phases d'explication et d'entraînement, pour lesquelles l'expérimentateur reste présent avec le sujet, ce dernier est laissé seul dans une salle spécialement aménagée, sans fenêtre, afin de pas provoquer de quelconques distractions. En cas de problème, ou pour prévenir l'expérimentateur de la fin d'une des étapes de l'expérimentation, le sujet dispose d'un téléphone avec une ligne directe le liant à l'expérimentateur.
5. le sujet est invité à remplir un questionnaire final d'évaluation globale de l'expérimentation incluant, en particulier, celle des représentations qui lui ont été proposées. Dès le début de l'expérimentation, le sujet est placé en position d'évaluateur, et non pas de simple acteur de façon à ce qu'il se comporte naturellement, sans se sentir observé, évalué. Lui donner ce rôle est également de nature à accroître son intérêt et sa motivation pour la réalisation des tâches expérimentales. Ce questionnaire est reproduit en annexe.
6. enfin, le sujet participe à un entretien semi-directif, dans lequel il est amené à communiquer ses impressions sur l'expérimentation, les différentes représentations visuelles des collections et les améliorations éventuelles qu'il serait souhaitable d'apporter aux visualisations interactives testées. Pour éviter de rendre mal à l'aise le sujet dans le cas d'un enregistrement, l'expérimentateur prend des notes durant le débriefings et les retranscrit une fois le sujet parti. Le schéma de ces entretiens est donné en annexe.

Dans ce chapitre, nous avons présenté les aspects communs aux trois études concernant, notamment, les tâches de recherche, les représentations et le matériel visuel utilisés. Les spécificités de chaque expérimentation seront présentées dans les chapitres suivants ; les caractéristiques communes décrites dans ce chapitre seront simplement résumées pour mémoire.

Chapitre 5

Etude préliminaire

Nous présentons dans ce chapitre une étude à caractère exploratoire, visant à déterminer la faisabilité et l'intérêt éventuel des représentations 3D pour l'interaction avec de grandes collections de photographies. En particulier, le but est d'évaluer la pertinence du choix de notre représentation 3D, à savoir un cylindre en vue externe et interne, et celle des métaphores d'interaction proposées : manipulation directe (vue externe) *versus* locomotion-immersion (vue interne). Cette étude nous a été utile pour l'élaboration d'une étude comparative approfondie de ces deux vues. L'évaluation, faite sur 8 sujets, porte sur les deux types de tâches présentées dans le chapitre précédent, à savoir la recherche de plusieurs cibles d'après une description textuelle de leur contenu (T1), puis la recherche de ces mêmes cibles présentées isolément au centre de l'écran dans les mêmes collections (T2). La première tâche favorise l'exploration, la seconde la navigation.

Les deux interfaces ont été positivement évaluées par les sujets, mais les préférences comme les performances se partagent entre les deux visualisations interactives. Malgré cela, il faut souligner que 7 des 8 sujets donnent leur préférence à la vue interne, pour ce qui est du confort visuel. La métaphore de la locomotion, si elle semble prometteuse, souffre de quelques défauts, qui la rendent moins performante par rapport à la métaphore de la manipulation, en particulier au niveau du zoom d'une photographie.

Après la description de la méthodologie et du protocole expérimental, nous présentons la méthodologie d'analyse des données recueillies, ainsi que l'examen des résultats. Le choix des représentations, ainsi que la méthodologie expérimentale, communs aux trois études, ont été présentés dans les deux chapitres précédents. Ainsi, nous ne ferons que de brefs rappels, en apportant précisions et explications en cas de différence.

5.1 Méthodologie

Dans cette première étude, la troisième dimension doit permettre de dépasser la métaphore de la manipulation, la plus usuelle dans les logiciels 2D, en autorisant le recours à la métaphore de la locomotion, c'est-à-dire le déplacement de l'utilisateur dans l'environnement virtuel, plutôt que le déplacement de l'environnement virtuel par rapport à lui. Ce type de métaphore, favorisant l'immersion, même s'il est des plus courants dans les jeux vidéos, est pratiquement absent des

logiciels « utilitaires » destinés au grand public. L'analogie avec la carte géographique et la carte routière permet de bien comprendre les aspects sous-jacents à ces deux métaphores. Le suivi d'un itinéraire sur une carte routière nécessite de bouger cette dernière, tandis que la même opération sur une grande carte murale nécessite le déplacement de l'observateur. Dans le premier cas, c'est une interaction avec l'espace de représentation, tandis que dans le second cas, c'est un déplacement dans l'espace de représentation. Il faut donc que la (les) représentation(s) de la collection permette l'implémentation des ces deux métaphores de navigation, dans des environnements 3D, sans privilégier l'une par rapport à l'autre. C'est cette contrainte qui a guidé le choix du cylindre vertical[8], associé à :

- une vue interne, implémentant la métaphore de la locomotion (VI) ;
- une vue externe, implémentant la métaphore de la manipulation (VE).

Afin d'étudier l'impact de ces deux métaphores sur la navigation et l'exploration dans de grandes collections non structurées de photographies, nous avons conçu une étude expérimentale portant sur la recherche visuelle de photographies suivant deux tâches courantes dans le cas de la consultation de collections :

- tâche 1 : recherche de photographies possédant des caractéristiques spécifiques décrites textuellement. La position des photographies dans la visualisation et leur contenu visuel sont inconnus de l'utilisateur (T1) ;
- tâche 2 : recherche des mêmes photographies, après présentation visuelle de celles-ci durant 3 secondes. Les photographies, ainsi que leur position dans la visualisation, sont donc connues (T2).

L'expérimentation repose sur la recherche de 6 cibles par vue (VI et VE), pour chaque tâche (T1 et T2), auxquelles s'ajoutent 2 cibles d'entraînement pour chaque vue dans la tâche 1. Les cibles cherchées dans la tâche 2 sont identiques à celles de la tâche 1, mais elles sont présentées dans un ordre différent.

5.1.1 Hypothèses testées

Cette étude a un caractère exploratoire. Elle doit nous permettre d'étudier l'intérêt des utilisateurs potentiels pour la représentation 3D que nous proposons, et de raffiner, d'après les commentaires des participants, les caractéristiques propres à chaque vue, principalement au niveau des moyens d'interaction. Nous n'avons de ce fait pas formulé d'hypothèse de recherche au sens strict du terme, mais plutôt des hypothèses de bon sens :

- hypothèse A : la vue interne (VI), par sa géométrie induisant une plus faible distorsion globale des items, permet un meilleur confort visuel et propose une meilleure vision globale ;
- hypothèse B : les performances sont influencées par le type de tâche et par la métaphore d'interaction. Autrement dit, les performances sont meilleures dans la tâche 2, et les performances intra-sujet ne sont pas égales pour les deux vues ;
- hypothèse C : les préférences sont influencées par les performances.

[8] cf. paragraphe 3.1.3

5.1.2 Conception de la représentation

Le cylindre offre, comme la sphère, un point de vue interne et externe, ce qui permet de concilier les deux modes d'interaction. Le choix du cylindre vertical plutôt qu'horizontal s'explique par le fait que cette orientation est plus favorable à la métaphore de la locomotion en vue interne, sans que cela n'affecte la métaphore de la manipulation en vue externe. Dans la locomotion, l'utilisateur se situe « à l'intérieur » du cylindre, ce qui est de nature à susciter une impression d'immersion. La disposition des images sur la surface externe du cylindre incite, quant à elle, à assimiler la manipulation de la visualisation à celle d'un objet virtuel, le cylindre se trouvant devant l'utilisateur.

Comme pour la sphère, mais dans une moindre mesure, le choix du cylindre, en tant que représentation continue, implique un repérage difficile. L'utilisateur, lors de la rotation du cylindre autour de son axe, peut facilement perdre tout point de repère, sans indication de la réalisation d'un tour complet. Pour pallier ce défaut, une solution simple a été adoptée et implémentée : l'utilisation d'une « fente noire », qui est en réalité une colonne du cylindre vierge de photographies. Ainsi l'espace est visuellement discontinu, ce qui offre à l'utilisateur un point de repère global.

Le cylindre comprend 869 photographies réparties en 11 « lignes » et 79 « colonnes ». Cette répartition a été établie de manière empirique, pour répondre à deux contraintes. D'une part, pour disposer d'assez de colonnes de façon à ce que la courbure soit perceptible de manière évidente, sans ambiguïté. D'autre part, pour privilégier les actions de rotation par rapport aux translations (verticales). Ce qui se traduit par la contrainte d'afficher toutes les lignes lors de la vue initiale, sans que la taille des photographies ne soit trop faible ; le nombre de photographies affichées simultanément sous forme de vignettes doit donc être compris entre 100 et 200 selon les dimensions de l'écran. L'exigence d'obtenir des temps de réaction de l'affichage réalistes explique également cette restriction. Les rotations ont été privilégiées car ces actions favorisent la distinction intuitive entre les deux métaphores, les translations étant plus ambiguës.

5.1.3 Interaction avec les représentations

Faciliter l'adhésion à chacune des deux métaphores impose l'emploi, pour chacune des vues du cylindre, de moyens d'interaction spécifiques. L'étude expérimentale comparative impose quant à elle que ces moyens d'interaction soit comparables, c'est-à-dire qu'ils proposent les mêmes fonctionnalités avec une mise en œuvre semblable, de façon à ne pas biaiser les résultats. Dans la vue interne, l'utilisateur doit avoir la sensation de se déplacer dans la représentation tandis que, dans la vue externe, il doit avoir le sentiment de manipuler la représentation. Pour cela, nous avons développé deux interactions distinctes, mais reposant sur les mêmes fonctionnalités de base offertes par tout gestionnaire de collection de photographies :

- déplacement : l'ensemble de la collection n'étant pas visible simultanément, du fait de la géométrie de la représentation, un déplacement latéral (rotation) doit être proposé ;
- zoom : afin de permettre un examen détaillé, il doit être possible d'agrandir une photographie particulière ;
- point de vue courant : l'utilisateur doit pouvoir connaître à tout instant la position de

l'affichage courant dans la collection.

Métaphore de la locomotion - vue interne du cylindre

Pour assurer la cohérence entre les moyens d'interaction avec la vue interne et la métaphore de la locomotion, il nous faut examiner plus en détail la notion de zoom. Reprenons l'exemple de la carte murale : lorsque l'on veut suivre un itinéraire, il faut naturellement se déplacer le long de la carte ; lorsque l'on veut regarder une zone en détail, il faut s'approcher de la carte, en adaptant sa hauteur de vue. Dans notre vue 3D, cela suppose donc que l'utilisateur puisse se déplacer dans toutes les directions :

- horizontalement, parallèlement à la paroi intérieure du cylindre, vers la gauche ou la droite (ce qui revient à une rotation selon l'axe du cylindre) ;
- perpendiculairement au cylindre, afin de s'éloigner ou se rapprocher du cylindre, et ainsi moduler le niveau de détail ;
- verticalement, vers le haut/bas, à distance constante du cylindre. Le rapprochement se faisant perpendiculairement au plan de l'écran, il est obligatoire de centrer l'image à l'écran l'image que l'on veut agrandir, avant de s'en approcher.

Un interacteur composé de 6 flèches est proposé à l'utilisateur. Ces flèches sont assimilables à des boutons, sur lesquels l'utilisateur doit cliquer pour se déplacer dans la direction correspondante. L'utilisation des flèches en tant que boutons est rendue obligatoire par le choix de la souris : nous avons donc transposé les flèches directionnelles du clavier sur une un objet graphique avec lequel on peut interagir par simple clic. Cet interacteur a été positionné en bas à droite de l'écran, afin d'être facilement repérable, et de ne pas entraver l'immersion dans l'environnement 3D. La figure 5.1 présente cet interacteur.

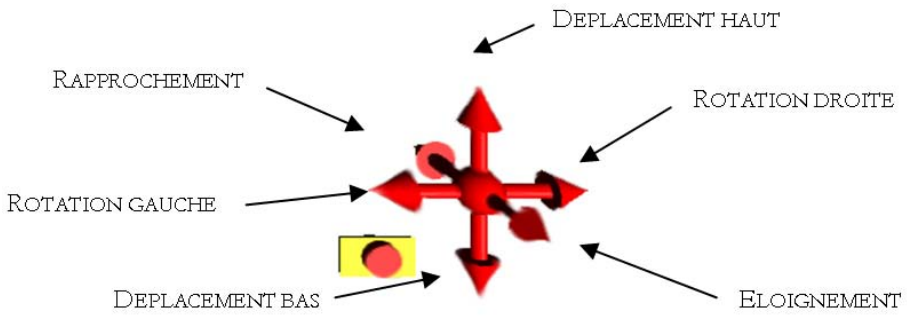


FIG. 5.1 – Interacteur pour la vue interne

La sphère au centre des flèches n'a pas d'autre fonction que celle de favoriser la distinction entre ces flèches. Pour arrêter un mouvement en cours, quel qu'il soit, il est nécessaire de cliquer sur un bouton supplémentaire. Il aurait été possible de cliquer à nouveau sur la flèche ayant permis de lancer le déplacement, mais cela aurait nécessité un effort de mémorisation. Ce bouton est placé à proximité des flèches, légèrement sur la gauche. De plus, nous imposons que toute transition d'un type de déplacement à un autre passe par le bouton stop, afin d'interdire tout déplacement composé. L'apparence choisie pour ce bouton est celle d'un bouton d'arrêt d'urgence pour lever toute ambiguïté sur sa fonction.

Métaphore de la manipulation - vue externe du cylindre

Afin de donner à l'utilisateur l'impression qu'il manipule réellement le cylindre, le choix s'est porté sur une interaction « directe ». Pour lancer la rotation, l'utilisateur doit cliquer sur la paroi du cylindre, plus précisément sur une partie sans image (la bordure entourant chaque image) et déplacer la souris dans le sens où il désire que s'effectue la rotation. Le *drag* est utilisé ici pour simuler l'action de rotation du cylindre. Cette solution est une variante de la manipulation directe plus intuitive et mieux adaptée au 3D que les classiques ascenseurs.

Pour garantir le respect de la métaphore de la manipulation, le zoom doit ici correspondre à la sélection d'une photographie particulière. Dans les faits, il suffit à l'utilisateur de cliquer sur une photographie pour que celle-ci soit reproduite agrandie au centre de l'écran, avec un facteur d'agrandissement identique à celui que l'on obtient dans la la vue interne lorsqu'on se rapproche au maximum de la face intérieure du cylindre. Pour faire disparaître la reproduction agrandie, un simple clic sur cette dernière suffit.

Notons que les déplacements verticaux et ceux perpendiculaires au cylindre (éloignement-rapprochement) n'ont pas été implémentés, car ils ne sont plus utiles ici. En effet, dans le cas de la métaphore de la locomotion, le déplacement de type rapprochement-éloignement est utilisé pour agrandir une photographie, tandis que le déplacement vertical est utilisé uniquement à des fins de centrage vertical. Le zoom implémenté dans la métaphore de la manipulation évite le recours à ces déplacements.

Discussion

Ces deux interfaces ont des fonctionnalités équivalentes mais la syntaxe de l'interface de la vue interne est plus complexe, pour certaines actions, que l'interface de la vue externe, en particulier en ce qui concerne le déplacement de type rapprochement-éloignement comparé au zoom. Mais cette différence est une conséquence du fait que nous avons choisi de privilégier la cohérence de l'implémentation de chaque métaphore pour tirer le meilleur parti des possibilités offertes par chacune d'elles. Nous tenons compte de cette différence dans l'interprétation des résultats obtenus lors de l'expérimentation.

Nature des déplacements

Pour des raisons techniques (i.e., les performances d'affichage graphique), il était impossible lors de la réalisation de la première étude d'obtenir des déplacements continus des visualisations : le manque de fluidité de l'affichage rendait le défilement latéral haché, ce qui risquait d'entraîner une certaine fatigue visuelle. Les ressources des équipements courants ne pouvaient alors assurer la fluidité du déplacement à l'écran de 869 images avec une résolution suffisante et une prise en compte réaliste des changements de point de vue (perspective) induits par les rotations. Nous avons donc décidé que le défilement latéral, sur lequel s'appuie principalement la recherche d'une cible, serait discontinu, avec un défilement colonne par colonne et une pause d'une durée de 1,6 secondes entre deux pas de défilement consécutifs. Ce type de déplacement favorise un parcours séquentiel de la collection, et vise ainsi à améliorer la rapidité de la recherche. Le fait d'imposer

une vitesse de défilement constante ne semble d'ailleurs pas avoir un caractère contraignant : dans l'expérience basée sur le logiciel *PhotoGoRound* (dont les visualisations ont une géométrie proche de celle des vues que nous proposons), les sujets ont jugée inutile la possibilité qui leur était offerte de moduler la vitesse de rotation [Combs et Bederson, 1999]. Enfin, en ce qui concerne les déplacements verticaux et les déplacements de type éloignement-rapprochement, nous avons maintenu la continuité, car ces mouvements étant plus plus faciles (matériellement) à réaliser de manière continue (d'un point de vue perceptif), sont beaucoup moins fatigants visuellement que les défilements latéraux.

5.1.4 Représentation du point de vue courant de l'utilisateur

La « fente » dans le cylindre peut s'avérer insuffisante pour fournir à l'utilisateur sa position précise à l'intérieur du cylindre (métaphore de la locomotion) ou la position de la colonne au centre de l'écran sur la surface du cylindre (métaphore de la manipulation). Il est donc nécessaire d'ajouter à l'interface un objet qui permette un repérage spatial plus précis. Nous appellerons cet objet la « carte » dans la suite de ce chapitre.

Cette carte a été réalisée en trois dimensions, sous la forme d'un cylindre homothétique à la représentation 3D. Ce cylindre qui possède la particularité d'être transparent afin de minimiser le masquage de la représentation évolue en cohérence avec la représentation 3D de la collection courante. Il est placé en bas à gauche de l'écran. La colonne de photographies au centre de l'écran est représentée sur cette carte par une bande verticale rouge. La position de cette bande sur le cylindre évolue en fonction des interactions de l'utilisateur. Toutefois, dans le souci de respecter les métaphores, quelques différences sont à souligner :

- dans la locomotion, l'utilisateur se déplaçant dans l'environnement 3D, la carte reste immobile et c'est la bande rouge (qui correspond à l'angle de prise de vue) qui se déplace ;
- dans la manipulation, l'utilisateur étant immobile par rapport au cylindre, la bande rouge reste fixe et c'est la carte qui se déplace.

La figure 5.2 présente les deux points de vue du dispositif de repérage.

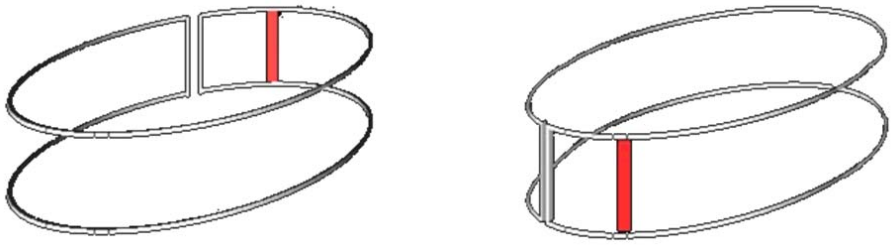


FIG. 5.2 – Représentation du point de vue courant pour la locomotion (à gauche) et la manipulation (à droite)

5.1.5 Implémentation

L'implémentation s'est faite sous *oRis* [Harrouet, 2000], un environnement de programmation créé par Fabrice Harrouet, enseignant-chercheur au Laboratoire d'Informatique Industrielle de l'Ecole Nationale d'Ingénieurs de Brest[9]. C'est un langage de programmation multi-agents, à la syntaxe proche du langage C++. *oRis* permet de créer et de gérer des objets 3D dans une scène et propose un environnement d'exécution multi-tâches, ce qui nous a permis de réaliser l'application complète avec ce langage. Le choix d'*oRis* est motivé par le fait que ce langage était utilisé par un autre membre de l'équipe *MErLIn* pour la création de scènes 3D animées. L'auteur de cet environnement logiciel ayant constitué une documentation en ligne des classes et des méthodes, l'implémentation n'a pas posé de réels problèmes, si ce n'est au niveau de l'application des textures. En effet, chaque photographie affichée à l'écran se décompose en réalité en un objet 3D (un cube) et une texture. Et cette texture doit être au format *bitmap* en couleurs indexées, présentant un format de 256 par 280 pixels au maximum. Cette résolution particulière a demandé l'adaptation du format des photographies en conséquence. Toute l'interface est composée d'objets en 3D, que ce soit les flèches de commande, les différents boutons, le système de repérage, et même le texte donnant les instructions à suivre en cours d'expérimentation.

Les figures 5.3 et 5.4 présentent les deux vues, interne et externe, de notre représentation.



FIG. 5.3 – Vue intérieure du cylindre vertical

[9] www.enib.fr/~harrouet



FIG. 5.4 – Vue extérieure du cylindre vertical

5.2 Protocole expérimental

5.2.1 Généralités

Pour tester les hypothèses de travail A, B et C[10], nous avons élaboré deux types de tâches de recherche de photographies : la première étant la recherche de cibles d'après une description textuelle (T1)[11], la seconde étant la recherche de ces mêmes cibles, après une présentation visuelle de 3 secondes (T2)[12]. Les deux tâches proposées sont donc des tâches de sélection de cibles visuelles, dans un temps limité à 270 secondes (ce temps permet de réaliser deux tours complet de cylindre).

Chaque participant avait donc à chercher des cibles particulières dans différentes collections, pour chaque vue (VI et VE) et chaque tâche (T1 et T2). Pour les tâches de type T1, les cibles, leurs positions dans les collections et les collections elles-mêmes diffèrent d'une tâche de recherche à l'autre. Les tâches de type T2 utilisent le même matériel visuel mais dans un ordre de présentation différent afin d'éviter tout phénomène d'apprentissage et rendre les tâches de recherche visuelle aussi réalistes que possible.

Un total de 16 photographies ont été choisies pour tenir le rôle de cibles (autant d'objets que de paysages). Elles ont été réparties de la manière suivante : dans T1, les participants devaient rechercher 6 cibles + 2 cibles d'entraînement par vue. Ces mêmes cibles ont naturellement été

[10] cf. paragraphe 5.1.1
[11] cf. paragraphe 4.1.2
[12] cf. paragraphe 4.1.3

utilisées dans T2. Les 2 cibles d'entraînement avaient pour rôle de familiariser les participants avec le fonctionnement de l'interface et la nature de la tâche T1.

Rappelons que la validité et la pertinence de ces descriptions ont été testées au cours d'une pré-expérimentation : cela a permis de procéder à quelques ajustements, principalement focalisés sur la clarté et la facilité de compréhension de la description (e.g., utilisation d'un vocabulaire simple).

5.2.2 Choix des cibles et descriptions textuelles

Les cibles ont été choisies pour leur simplicité à être décrites, ainsi que pour leur saillance visuelle[13]. La figure 5.5 présente les 12 cibles retenues, 6 par métaphore. Les 4 cibles d'entraînement ne sont pas représentées, car elles ne sont pas prises en compte dans les analyses qualitatives et quantitatives. L'ordre de présentation des cibles pour chaque vue et chaque tâche est présenté dans l'annexe A.



(a) Vue intérieure



(b) Vue extérieure

FIG. 5.5 – Photographies servant de cibles

Le rôle des descriptions est de dépeindre parfaitement les cibles, rendant toute confusion de la cible avec d'autres photographies impossible et permettant à chaque participant de se constituer une image mentale de la cible la plus fidèle possible. Néanmoins, ces descriptions ne devaient pas dépasser quelques lignes, pour éviter d'accroître de façon pénalisante la charge de travail cognitive des participants. Pour concilier ces deux impératifs, chaque description est élaborée en deux parties. Dans la première partie, assez courte, le sujet découvre le thème général de la photographie, ainsi que les informations permettant de la détecter rapidement (couleur, cadrage, direction). La seconde partie de la description fournit des informations complémentaires, qui permettent au sujet de conforter ou d'invalider son choix, en apportant des éléments discriminants supplémentaires. Pour être facilement différenciées, ces deux parties sont imprimées de couleurs différentes (5.6).

A titre d'illustration, une des descriptions utilisées est reproduite ici (figure 5.7) avec un codage par couleur permettant de visualiser les différents types d'informations introduits systématiquement dans chaque description.

[13] cf. paragraphe 4.2.2



Un large chemin de planches brun clair s'enfonce dans un sous-bois. Le feuillage peu fourni laisse voir les troncs sombres et élancés des arbres à contre-jour, et entrevoir le ciel blanchâtre.

De chaque côté du chemin, une bordure en bois.
Sur ses bords et sur le sol, quelques feuilles mortes rousses.

(a) (b)

FIG. 5.6 – Cible (a) et description correspondante (b)



Un large chemin de planches brun clair s'enfonce dans un sous-bois. Le feuillage peu fourni laisse voir les troncs sombres et élancés des arbres à contre-jour, et entrevoir le ciel blanchâtre.

De chaque côté du chemin, une bordure en bois.
Sur ses bords et sur le sol, quelques feuilles mortes rousses.

FIG. 5.7 – Nature des différentes informations présentes dans une description

Les descriptions de l'ensemble des cibles choisies pour les trois études sont présentées dans l'annexe A.

5.2.3 Constitution des collections

Le nombre de photographies de la base de données, bien qu'important, interdit de réaliser 12 collections totalement distinctes. Cette limitation impose un renouvellement partiel des photographies d'une collection à l'autre, la disposition des images conservées étant modifiée. Dans la pratique, une série de 1200 photographies a été utilisée pour chaque vue (VI et VE) ; les cibles ont été choisies de manière équitable entre paysages et objets ; de même, les collections comprennent un nombre identique d'objets et de paysages. La stratégie pour construire une collection est la suivante :

1. toutes les photographies appartenant au thème de la cible sont incluses dans la collection, soit 30 photographies ;
2. 22 photographies sont prélevées au hasard pour chaque thème restant (39), soit au total 888 photographies ;
3. 19 photographies sont supprimées de manière aléatoire parmi les 888 (hormis les images du thème de la cible) ;
4. les 869 photographies résultantes sont alors disposées aléatoirement à la surface du cylindre.

De cette manière, les 1200 photographies d'une série interviennent dans la constitution des 6 collections d'une métaphore. Les 12 cibles sont exclues des collections dans lesquelles elles ne

sont pas l'objet de la recherche, afin de contrôler strictement la durée de leur exposition à la vue des sujets. Les cibles sont réparties spatialement de manière aléatoire sur le cylindre.

5.2.4 Conditions initiales

La cible n'est jamais visible lors de l'affichage initial, pour forcer les sujets à interagir avec la visualisation. La position initiale du cylindre est identique pour les deux vues : les bords horizontaux de la bande constituée par la paroi du cylindre sont équidistants des bords haut et bas de l'écran. La paroi du cylindre est à une distance suffisante pour que les sujets des photographies puissent être perçus facilement, mais insuffisante pour que l'on puisse identifier la cible avec certitude : cette démarche délibérée doit inciter les sujets à zoomer ou à s'approcher de la paroi pour une identification précise. Nous avons donc fixé, dans chaque vue, la distance entre la représentation et la caméra virtuelle de façon à obtenir l'affichage simultané du même nombre de photographies pour les deux vues ; ainsi, les images situées sur les bords dans la vue interne sont de taille semblable et donc aussi facilement « identifiables » que celles situées au centre dans la vue externe. La « fente », c'est-à-dire la colonne sans photographie, est toujours affichée initialement à gauche de l'écran, laissant entrevoir le début d'une colonne de photographies à sa gauche.

5.2.5 Choix des participants

Les participants sont au nombre de 8 et sont tous volontaires. La sélection s'est effectuée en tenant compte des contraintes suivantes :

- avoir un niveau courant de familiarité avec l'informatique et la maîtrise des périphériques habituels (clavier et souris) ;
- appartenir à une tranche d'âge homogène, pour obtenir des temps de réaction comparables ;
- avoir des profils variés, en termes de formation et de centres d'intérêt, pour représenter la diversité du grand public.

Les participants ont été choisis parmi les stagiaires du LORIA. Les données les concernant sont synthétisées dans le tableau 5.1.

Numéro	Age	Sexe	Formation
1	27	M	M1 mathématiques-informatique
2	34	M	M1 Audiovisuel - DUT Informatique
3	19	F	Elève ingénieur INSA 1ère année (préparation intégrée)
4	22	M	Elève ingénieur ENST 2ème année
5	24	M	M2 Robotique
6	23	F	M1 mathématiques-informatique
7	20	F	L3 mathématiques-informatique
8	27	F	M2 bio-informatique

TAB. 5.1 – Identification des participants

La moyenne d'âge des participants est de 24 ans et demi ; tous sont familiers avec les environnements informatiques et utilisent fréquemment des ordinateurs dans le cadre de leurs études.

5.2.6 Déroulement de l'expérimentation

Deux types de tâches sont à réaliser (T1 et T2), et ce pour chaque vue interactive (VI et VE), ce qui fait donc une suite de quatre blocs de tâches à réaliser. Naturellement, les tâches de type T1 doivent s'effectuer avant celles de type T2. L'ordre de présentation des deux vues interactives n'a pas été équilibré : en raison du faible nombre de sujets, l'influence éventuelle de ce facteur est contrôlée en fixant sa valeur. De cette manière, l'influence de l'ordre d'exécution des tâches n'est pas inexistante, mais identique pour tous les sujets. Arbitrairement, nous avons décidé de commencer par la métaphore de la locomotion ; il nous a semblé cohérent de garder ce même ordre dans les deux tâches, afin de ne pas favoriser l'une ou l'autre métaphore. L'ordre est donc le suivant :

T1	Locomotion	2 cibles d'entraînement puis 6 cibles
	Manipulation	2 cibles d'entraînement puis 6 cibles
T2	Locomotion	6 cibles identiques à celles utilisées dans T1 pour la Locomotion
	Manipulation	6 cibles identiques à celles utilisées dans T1 pour la Manipulation

Le participant doit tout d'abord remplir un questionnaire initial (voir annexe A) permettant d'identifier son profil ; ensuite le but de l'expérimentation lui est présenté, avec une explication des deux modes d'interaction et des deux types de tâches qu'il lui faut réaliser. Afin de motiver les participants, de réduire l'effet de la situation expérimentale et de donner du réalisme à leur activité, ceux-ci ont été placés en situation d'évaluateurs de l'interface, et non pas en simples acteurs, lors de la consigne globale donnée oralement au début de l'expérimentation.

Le sujet prend ensuite place dans la salle d'expérimentation. Il commence par les tâches de type T1 et l'utilisation de la vue interne (VI, la métaphore de la locomotion) : les moyens d'interaction avec la visualisation lui sont présentés, tandis qu'il découvre l'interface (une collection spéciale a été prévue à cet effet). Des consignes générales lui sont données oralement, avec une insistance sur la recommandation de privilégier la précision plutôt que la rapidité, bien que les recherches soient bornées temporellement. Le participant peut alors librement s'approprier les moyens d'interaction, et une fois qu'il estime les maîtriser suffisamment, il peut entreprendre la recherche des deux cibles d'entraînement, puis des six cibles, en dehors de la présence de l'expérimentateur (l'expérimentateur peut toujours être joint par téléphone, via une ligne directe). Ces recherches achevées, il passe à la seconde métaphore, à savoir la manipulation, et reçoit à nouveau les explications nécessaires sur les différents moyens d'interaction. Le sujet aborde ensuite les tâches de type T2 sans recevoir aucune information particulière de la part de l'expérimentateur. Cependant, s'il en fait la demande, le déroulement d'une recherche visuelle de type T2 lui est brièvement rappelé. Lorsque le sujet a réalisé les 24 tâches de recherche, il répond à un questionnaire final (voir annexe A) dans lequel il peut noter ses observations sur les différentes questions posées. L'expérimentation se conclut par un entretien semi-directif (ou débriefing), au cours duquel le sujet est amené à communiquer ses impressions sur l'expérimentation, et plus précisément sur les deux métaphores proposées.

5.3 Méthodes d'analyse

Etant donné le caractère exploratoire de l'étude, l'analyse a porté sur les comportements des sujets à partir de l'ensemble des informations recueillies. A l'issue de l'expérimentation, 3 sources distinctes d'informations sont disponibles, pour chaque sujet :

- les données et mesures relatives à son activité et à ses performances ;
- l'entretien de débriefing auquel il a participé ;
- les questionnaires initial et final.

5.3.1 Données et mesures

Chaque action des sujets a été datée et enregistrée. De manière exhaustive, les enregistrements comprennent : les déplacements des vues (durée, direction et amplitude), les images agrandies (pour la vue externe) et, pour chaque cible, la durée de la recherche, son résultat (image sélectionnée ou temps de recherche dépassé).

Tâche 1

Outre les mesures que nous venons de présenter, les points suivants ont été étudiés :

- l'examen du premier affichage : l'analyse des temps d'examen de l'affichage initial de chaque collection permet de savoir si le sujet effectue ou non un balayage visuel des photographies affichées à l'écran avant de mettre en œuvre un défilement latéral ;
- le sens de défilement latéral : ici, nous voulons savoir si le sujet utilise un sens de rotation constant d'une image à l'autre ;
- les arrêts du défilement : nous étudions plus particulièrement leur durée, leur motif et la distance de parcours séparant deux arrêts successifs du défilement ;
- les autres déplacements : les autres déplacements sont les déplacements verticaux et les déplacements de type éloignement/rapprochement. Nous considérons leur amplitude et tentons de comprendre leur motivation.

Concernant la métaphore de la manipulation (VE), les déplacements autres que la rotation ne sont pas étudiés, puisque non implémentés. En revanche, nous disposons de la connaissance des images agrandies, puisque cette métaphore met en œuvre un zoom par sélection à la souris, ce qui n'est pas le cas pour la métaphore de la locomotion (VI).

Tâche 2

En plus des mesures communes aux deux vues, nous observons également la capacité du participant à choisir le sens de rotation optimum ainsi qu'à identifier la cible lorsqu'elle apparaît à l'écran pour la première fois, afin dévaluer la qualité du repérage spatial dans les deux métaphores.

5.4 Etude quantitative

5.4.1 Temps de recherche

Dans cette partie sont présentées les performances des participants, évaluées en termes de temps de recherche cumulés, pour VE et VI, et T1 et T2. Le graphique 5.8 synthétise ces résultats. Les temps de recherche sont globalement plus courts pour T2 que pour T1, ce qui est cohérent

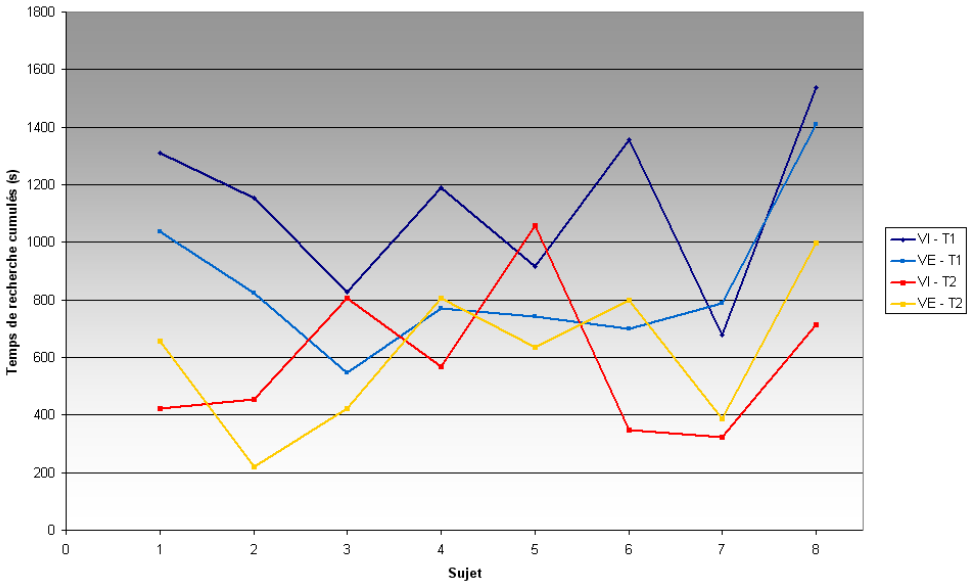


FIG. 5.8 – Temps de recherche cumulés pour chaque sujet

avec la nature des deux tâches : recherche d'une photographie inconnue (T1) versus recherche d'une photographie familière (T2). La différence entre les deux vues est assez marquée dans T1, et clairement en faveur de la vue externe : la moyenne des temps de recherche cumulés est de 1121,3 secondes pour VI, contre 852,9 secondes pour VE. Sur les 8 sujets, 6 sont nettement plus rapides avec VE (`S1` à `S6`) et les temps (avec VE *vs* VI) sont proches pour les deux autres (`S7` et `S8`). Les différences inter-individuelles sont plus marqués dans T2 : 4 sujets sont nettement plus rapides avec VI (`S1`, `S4`, `S6` et `S8`) et 3 sujets sont nettement plus rapides avec VE (`S2`, `S3` et `S5`). Pour le participant `7`, les temps sont très proches. La moyenne des temps de recherche cumulés est de 586,5 secondes pour VI, contre 615,2 secondes pour VE.

5.4.2 Nombre de succès

Les participants devaient rechercher 6 cibles par vue et par tâche. Nous examinons donc ici le nombre de cibles trouvées, pour ces 4 conditions. Les données sont présentées dans le tableau 5.2.

Dans T1, on peut remarquer que le nombre d'erreurs cumulées pour tous les sujets est similaire dans VI et VE (respectivement 16 et 20). Les erreurs sont bien plus faibles dans T2, ce qui

Sujet		*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Tâche 1	*VI*	3	5	6	4	3	3	6	2
	VE	4	6	6	6	1	1	1	3
Tâche 2	*VI*	6	5	6	6	5	6	6	3
	VE	6	6	6	5	5	4	6	5

TAB. 5.2 – Nombre de succès par vue (VE, VI) et par tâche (T1, T2)

s'explique par la facilité de la tâche. Le nombre de photographies étant relativement restreint, il est difficile de conclure en faveur de l'une ou l'autre vue.

5.4.3 Qualité du repérage spatial - T2

La tâche 2 nous permet d'évaluer la qualité du repérage spatial, d'après plusieurs critères :
- l'examen initial de la collection : aucune des cibles n'étant positionnée dans l'affichage initial, un examen avant la première rotation indique une faiblesse au niveau de la mémorisation ;
- le sens de défilement latéral : pour retrouver chaque cible, le participant doit normalement choisir le sens de rotation permettant de minimiser la longueur de parcours ;
- nombre de passages devant la cible : le participant doit retrouver la cible en faisant au plus un tour ;
- précision : naturellement, le participant doit retrouver chaque cible, dans chaque vue.

Néanmoins, nous ne considérons pas ces deux derniers critères, en raison de la simplicité de la tâche 2 : les résultats (nombre de succès) sont très proches dans VE et VI, mis à part pour le sujet `8`, qui semble avoir rencontré de grandes difficultés dans la vue interne (seulement 3 succès, contre 5 dans VE).

Examen initial

L'examen initial consiste en un balayage visuel de l'ensemble des images affichées initialement à l'écran, avant de lancer la rotation. Le critère retenu pour évaluer la présence d'un examen initial est la durée entre l'affichage initial de la visualisation et le début de la première rotation. Lorsque le temps « d'inaction » dépasse 4 secondes, nous considérons qu'il y a examen initial. Nous estimons à 4 secondes le temps nécessaire au sujet pour prendre ses repères et lancer la rotation ou un autre déplacement, autrement dit le temps qu'il lui faut pour activer la commande de son choix. Le tableau 5.3 synthétise ces résultats.

L'examen initial n'ayant pas lieu de se produire dans T2 si la mémorisation a été correcte, il est surprenant de constater que dans la vue externe, cet examen a été réalisé par tous les sujets pour au minimum 2 des 6 cibles, tandis que 5 sujets ont systématiquement omis cette étape dans la vue interne. Cette différence importante semble indiquer une mauvaise mémorisation de la position des images lors de l'interaction avec la vue externe, même si cela n'est pas confirmé par les entretiens de débriefing.

Sujet	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
VI	0	0	0	6	3	0	0	5
VE	6	2	4	6	6	3	4	5

TAB. 5.3 – Examens initiaux cumulés pour chaque sujet dans VE et VI (T2)

Sens de défilement latéral

Les résultats concernant le choix du sens de rotation sont reportés dans le tableau 5.4. Les 6 cellules grisées correspondent à des problèmes d'enregistrement lors de l'expérimentation, les cellules dont le texte est imprimé en rouge correspondent au choix du sens de rotation non optimum, tandis que le texte imprimé en bleu correspond à un sens de rotation constant pour toutes les cibles d'une vue. Quant au texte imprimé en noir, il correspond au choix du sens de rotation optimum.

Sujet	*VI*						*VE*					
	1	*2*	*3*	*4*	*5*	*6*	*1*	*2*	*3*	*4*	*5*	*6*
1												
2	G	G	D	D	D	D	G	D	D	D	D	D
3	G	G	G	D	D	G	G	D	G	D	G	D
4	D	D	G	G	G	D	G	G	G	G	G	G
5	D	D	G	D	G	D	G	G	G	G	G	
6	G	D	G	D	G	D	D	G	G	D	G	D
7	D	G	D	D	G	D	G	G	G	D	G	D
8	D	D	D	D	D		G	G	G	G	G	G

TAB. 5.4 – Sens de rotation pour chaque sujet dans VE et VI (T2) : un mauvais choix de sens de rotation est représenté en rouge, et un sens de rotation constant est représenté en bleu

Sur les sujets `2` à `8`, nous pouvons remarquer que 4 sujets ont conservé un sens constant de rotation dans VE, contre 1 seul dans VI. En outre, le sens de rotation choisi, même s'il est optimal, n'est pas la preuve d'un choix délibéré. Si l'on exclut les 3 sujets ayant adopté un sens de rotation constant pour VE, les résultats des 4 autres sujets sont proches d'une vue à l'autre, pour chacun d'eux. On ne peut donc conclure qu'une des vues favorise plus que l'autre la mémorisation de la position de la cible dans la visualisation.

Discussion

Les différences, lorsqu'elles existent pour un sujet, restent peu marquées : il en est ainsi pour le nombre de passages devant les cibles, ainsi que pour le nombre de succès. Toutefois, la présence d'un examen initial systématique dans la vue externe traduit une mauvaise mémorisation de la position des cibles dans les visualisations. Ce résultat, associé au fait que 4 sujets ont choisi un

sens de rotation constant avec VE, contre 1 seul avec VI, suggère une plus grande difficulté de repérage spatial dans la vue externe. Les entretiens de débriefing contribuent à renforcer cette interprétation ; en effet, plusieurs sujets ont indiqué que la représentation du point de vue courant (la carte) était plus difficile à utiliser avec la vue externe qu'avec la vue interne.

5.5 Etude qualitative

L'étude qualitative repose principalement sur l'analyse des entretiens de débriefings et des questionnaires finaux. Elle nous a permis d'évaluer la compréhension des métaphores d'interaction par les sujets, et de mettre en lumière les préférences de ces derniers. De plus, elle a conforté la pertinence du choix de la troisième dimension. La connaissance des améliorations souhaitées par les participants s'est avérée utile pour mettre en œuvre l'étude expérimentale décrite dans le chapitre suivant, qui porte sur un nombre plus important de participants.

5.5.1 Pertinence des cibles choisies et de leurs descriptions textuelles (T1)

Il s'agit ici d'évaluer la première tâche proposée aux participants (T1), en se focalisant principalement sur la pertinence des descriptions et des cibles choisies. Le tableau 5.5 caractérise, cible par cible, le nombre d'erreurs cumulées (pour les 8 sujets), ainsi que la nature de celles-ci :

- une erreur provoquée par la confusion avec une autre photographie de la collection ;
- la non-sélection d'une cible présumée, par manque de temps.

	VI						*VE*					
Cible	*1*	*2*	*3*	*4*	*5*	*6*	*1*	*2*	*3*	*4*	*5*	*6*
Nombre d'échecs cumulés	1	3	3	2	3	2	3	1	0	2	2	0
dont confusions avec autres photographies	1	0	1	0	0	0	0	0		1	0	
dont temps dépassé	0	3	2	2	3	2	3	1		1	2	

Tab. 5.5 – Nombre d'erreurs cumulées (8 sujets) et nature de celles-ci, par vue (VI et VE) pour T1

Pour la vue interne, 3 cibles (casiers, ilôt, enfants) n'ont pas été trouvées par 3 des 8 sujets, ce qui indique la présence d'une plus grande difficulté pour ces cibles. Par ailleurs, aucune cible n'a été trouvée par tous les sujets. Dans la vue externe, 2 cibles (éclair, chemin) ont soulevé des difficultés assez importantes : en moyenne, plus de 12 photographies, ont été agrandies par les sujets pendant la recherche de la cible 1. Comme la grande majorité des échecs sont imputables à un manque de temps, ils sont par nature ambigus. On constate néanmoins pour VE, que les cibles ayant suscité le plus d'échecs ont celles qui ont occasionné le plus grand nombre d'images agrandies, ce qui suggère l'influence importante des caractéristiques visuelles des cibles sur la difficulté de leur identification. Le succès de la recherche, cela semble évident, est nettement influencé par la « difficulté » de la cible. Il reste à comprendre les critères visuels susceptibles d'expliquer cette difficulté.

Homogénéité de la difficulté des cibles

Deux conclusions, triviales *a posteriori*, peuvent être formulées :
- les photographies présentant une forte saillance visuelle, qu'elle soit due aux différences de couleurs présentes dans la photographie ou à sa couleur générale par rapport à celle des autres photographies, sont plus facilement identifiées par les sujets ;
- les photographies relevant d'un thème aisément identifiable visuellement sont plus facilement trouvées par les sujets.

Ces conclusions sont confirmées par les participants lors des entretiens de débriefing : ils estiment que certaines cibles n'étaient pas clairement identifiables comparées aux autres photographies de la collection en raison de leur manque de saillance visuelle. Les débriefings indiquent en outre que les cibles de type « paysage » sont plus difficiles à trouver que les objets. Ainsi, les sujets précisent qu'il est plus simple de construire une représentation mentale d'un objet relevant d'un thème familier : « on sait à quoi ressemble un bébé phoque » ou « le paon, c'était facile ». Notons qu'aucune cible n'a rencontré un échec total, étant toujours trouvée par au moins 5 sujets sur 8, quelles que soient la tâche et la métaphore d'interaction. Il convient enfin de remarquer que les photographies ayant posé problème dans la tâche 1 ont également suscité des erreurs dans la tâche 2, mais dans une moindre mesure. Il reste finalement que les images « simples » (i.e., comportant peu de détails et représentant des objets/paysages familiers) sont plus facilement mémorisées [Attneave, 1955].

Composition des descriptions

D'après les participants, tous les détails donnés dans les descriptions sont utiles. L'analyse des photographies agrandies (VE) le confirme : elles sont différentes d'un sujet à l'autre, avec des caractéristiques visuelles non semblables. Chaque participant « pioche » dans les descriptions des éléments qui lui sont personnels, pour construire sa représentation mentale de la cible à chercher. Les facteurs d'ordre culturel exercent sur ce filtrage une influence importante.

Stratégies de filtrage

Nous nous intéressons ici aux stratégies de filtrage, et plus particulièrement à la relation entre celles-ci et les thèmes des photographies. On remarque que pour une cible dont le thème est facilement identifiable, les informations sur la couleur passent au second plan : cette particularité s'observe dans le cas de l'éclair (VE - cible 1), cible pour laquelle les sujets ont agrandi des ciels nocturnes zébrés d'éclairs, mais pas forcément « bleutés ». Il en est de même pour la cible avec l'homme barbu (VE - cible 6) : les sujets ont agrandi des photographies présentant une personne portant un chapeau, ignorant la « tenture rayée jaune et blanche » et la « barbe ». De manière générale, la description influence la stratégie de filtrage visuel : si le thème leur semble simple à identifier, les sujets ont recours à un filtrage par thème et éventuellement par couleur. Dans le cas contraire, ils font en majorité un filtrage par couleur, et éventuellement par direction et par forme, en cas de couleurs peu saillantes.

Les descriptions, par leur précision, ont fourni à chaque sujet des informations suffisantes pour

lui permettre de se créer une image mentale pertinente de chaque cible. Elles étaient suffisamment évocatrices et suggestives pour que la tâche 1 soit exécutée sans problème, excepté pour une ou deux cibles qui ne présentaient pas de saillance visuelle et/ou thématique. Il semble donc que la tâche 1 réponde à nos attentes : les descriptions textuelles, telles qu'elles ont été construites, permettent bien de retrouver sans ambiguïté les cibles en favorisant la création de représentations mentales pertinentes, ce qui assure le réalisme des tâches de type T1.

5.5.2 Locomotion *versus* Manipulation

Zoom *versus* rapprochement

Pour la vue externe, un simple clic sur une photographie permet de l'agrandir, tandis que la métaphore retenue pour la vue interne impose à l'utilisateur de s'approcher pour réaliser la même action. C'est le respect des métaphores d'interaction qui induit cette différence entre les deux vues. Les fréquences d'utilisation de chaque dispositif sont reportées dans le tableau 5.6.

Sujet	**T1**		**T2**	
	VI	**VE**	**VI**	**VE**
1	3	82	0	16
2	6	30	0	6
3	3	43	0	10
4	1	28	0	11
5	0	30	0	6
6	4	18	0	6
7	5	17	0	6
8	5	26	0	13

TAB. 5.6 – Nombre cumulé d'utilisations des dispositifs de zoom, par vue et par tâche

L'observation de la fréquence d'utilisation des couples rapprochement-éloignement lors de l'interaction avec la vue interne montre qu'ils ont été utilisés avec parcimonie : pendant T1, ces déplacements ont été utilisés en moyenne 0,6 fois par cible, davantage à des fins de confirmation avant la validation. Dans T2, ils n'ont jamais utilisés. *A contrario*, les sujets ont, dans l'ensemble, fréquemment zoomé dans la vue externe (en moyenne 5,7 fois par cible dans T1 et 1,5 fois dans T2). En particulier, le recours au zoom est fréquent pour les cibles leur semblant difficiles : la cible 1 a été agrandie plus de 12 fois en moyenne.

La fréquence d'utilisation du dispositif de zoom est clairement en faveur de la vue externe. Deux hypothèses peuvent expliquer cette différence :

- la vue interne permettrait une meilleure vision des photographies ;
- le déplacement de type éloignement-rapprochement serait trop coûteux en temps ou trop complexe à mettre en œuvre.

Le contenu des débriefings désigne la seconde hypothèse : les 6 premiers sujets ont souligné l'intérêt du zoom dans le cas de la métaphore de la manipulation, car il permet d'agrandir une

image rapidement sans modifier l'environnement ni commander l'arrêt du défilement. De fait, ces sujets auraient apprécié de retrouver le même dispositif dans la métaphore de la locomotion, pour accélérer les recherches. En effet, ils estiment que les déplacements autres que la rotation, en particulier ceux de type éloignement-rapprochement, sont trop coûteux en temps et en nombre de « clics ». Seul le sujet `6` a plébiscité le zoom de la vue interne et aurait souhaité son adaptation dans la métaphore de la manipulation. L'observation des temps de recherche confirme également cette seconde hypothèse : si tous les participants sont plus rapides avec VE dans T1, 3 d'entre eux sont plus rapides avec VI dans T2, là où les dispositifs de zooms sont très peu utilisés.

Moyens d'interaction

Plusieurs sujets considèrent la locomotion plus adaptée à des débutants, en raison même des moyens d'interaction proposés : l'usage de l'interacteur est plus intuitif, car l'utilisateur associe facilement chaque flèche au déplacement qu'elle provoque dans l'espace 3D.

Concernant la vue interne, certains sujets souhaiteraient pouvoir utiliser un autre périphérique, à la place ou en complément de la souris, pour rendre les déplacements plus aisés. Si les participants 4 et 6 auraient apprécié un joystick, les participants 2 et 5 jugent que ce type de périphérique serait peu pratique à utiliser. L'utilisation conjointe de la souris et du clavier a également été suggérée ; ce mode d'interaction est emprunté au monde des jeux vidéos, source que citent volontiers les sujets, et fonctionne de la manière suivante : dans les jeux en 3D à la première personne (lorsque le joueur « emprunte » la vision du monde virtuel au personnage qu'il dirige), l'utilisateur guide les déplacements de son personnage avec les flèches du clavier, et oriente le regard de celui-ci (que l'on pourrait ramener à la caméra virtuelle) avec la souris. L'action conjuguée de la souris et du clavier permet des déplacements précis, l'orientation du regard influençant l'orientation du déplacement. Les sujets souhaiteraient donc avoir plus de liberté de mouvement et se déplacer directement dans la direction souhaitée, plutôt que devoir combiner rotations, déplacements verticaux et éloignements-rapprochements. Il convient néanmoins de noter que ce mode d'interaction n'est pas forcément évident et intuitif pour un utilisateur n'ayant jamais pratiqué ce type de jeux.

Concernant la vue externe, certains sujets ont trouvé difficile de devoir cliquer entre les images pour effectuer une rotation, à cause de la précision demandée par cette action (`S3`, `S4` et `S8`). Les autres sujets ont jugé cet aspect pratique, car il évite de quitter la représentation des yeux.

Dans l'état actuel, aucun mode d'interaction ne semble exempt de défauts, dans la mesure où les remarques et améliorations formulées ne convergent pas et parfois même sont contradictoires. La locomotion semble prometteuse à condition d'y apporter les modifications nécessaires pour rendre la navigation plus rapide et l'interaction plus souple. Une explication partielle à la mauvaise appréciation du dispositif de zoom associé à cette métaphore est lié au défilement colonne par colonne (choisi comme compromis technique) et à la nécessité de centrer une photographie à l'écran avant de s'en approcher. Ce qui implique, dans la pratique, d'attendre que la photographie à agrandir parvienne au centre de l'écran, ce qui peut paraître long et de ce fait frustrant. Deux solutions sont envisageables :

– transposer directement le dispositif de zoom implémenté dans la métaphore de la manipu-

lation. Mais ce choix aurait pour conséquence d'amoindrir la force de la métaphore ;
- automatiser le déplacement vers le centre de l'écran et le rendre fluide donc continu. Cette solution aurait l'avantage de préserver la métaphore de la locomotion, mais demanderait une programmation différente et imposerait de fréquentes animations.

Bien sûr, ces propositions sont à tester et évaluer car elles ne préjugent en rien de l'accueil effectif des utilisateurs potentiels.

Cylindre intérieur *versus* cylindre extérieur

Contrairement aux modes d'interaction avec la représentation, les préférences concernant les deux vues du cylindre sont plus tranchées. Sept des 8 sujets donnent leur préférence à la vue interne pour le meilleur confort de vision et la vision globale de meilleure qualité qu'elle apporte : ces participants motivent leur choix d'une part par les déformations moins importantes induites par cette vue et d'autre part par le sentiment de voir davantage d'images simultanément. Ils remarquent que seule la colonne centrale est réellement visible dans le cas de la vue externe, les images latérales présentant d'importantes déformations, alors qu'un plus grand nombre d'images semblent être de taille similaire dans la vue interne. La géométrie des vues interne et externe illustre cette remarque : dans la vue interne, les images latérales, c'est-à-dire les plus déformées, sont également les plus grandes, car situées plus près du sujet. La taille compense d'une certaine façon la déformation. Cette caractéristique est inversée dans le cas de la vue externe, où les images latérales déformées sont les plus petites, ce qui rend leur observation difficile.

Au delà de cette différence, la géométrie des vues semble avoir une influence non négligeable sur les stratégies de parcours visuel, d'après les informations recueillies dans les débriefings : lors de la rotation du cylindre, la plupart des sujets affirment observer la colonne entrante dans la vue interne, alors que leur regard se porte plus naturellement sur la colonne centrale (ou la troisième colonne après la colonne entrante) dans la vue externe.

Ajoutons enfin qu'un sujet a fait la comparaison avec un écran de cinéma pour justifier sa préférence du cylindre intérieur, en précisant que la concavité accroît le sentiment d'immersion.

Et la 3D ?

Mis à part un seul sujet, tous les autres trouvent intéressant le choix de la 3D alliée à la représentation cylindrique ; selon eux, elle permet, par rapport aux visualisations 2D classiques, une recherche visuelle plus efficace dans un ensemble plus grand d'images affichées simultanément. Un avantage supplémentaire réside dans la continuité de l'espace de recherche, contrairement à une interface 2D classique.

Jugements globaux

Certains sujets placent l'efficacité au-dessus du confort visuel ou du confort d'interaction, ce qui peut résulter en partie de la contrainte de trouver les cibles en temps borné. Ces sujets préfèrent la métaphore de la manipulation, en raison de la facilité d'utilisation du zoom, en

particulier de sa rapidité et de son faible coût en nombre d'actions à effectuer.

De manière détaillée :

- les sujets `1`, `2` et `7` préfèrent globalement la manipulation, du moins en l'état actuel des choses, pour sa simplicité et son efficacité. Ils préfèrent néanmoins la locomotion pour le confort visuel. Si la métaphore de la locomotion pouvait être enrichie d'un dispositif de zoom identique à celui de la manipulation, ils donneraient clairement leur préférence à la locomotion. Ces trois sujets sont tous familiers des systèmes informatiques, mais leur profils sont assez différents, que ce soit en termes d'âge (27, 34 et 20 ans), de sexe (deux hommes et une femme) et d'utilisation des logiciels graphiques, puisque les deux premiers sujets ne s'en servent que très peu contrairement au troisième ;
- les sujets `3` et `4` privilégient le confort d'utilisation et préfèrent de ce fait la locomotion, car la commande y est plus intuitive et la visualisation plus agréable. Comme les 3 sujets précédents, ils auraient toutefois apprécié un dispositif de zoom plus rapide et intuitif. Et dans ce cas, ils auraient donné leur préférence sans réserve à cette vue. Ces deux sujets sont relativement jeunes (19 et 22 ans), et de sexe différent.
- le sujet `5` préfère nettement la métaphore de la manipulation, car il a moins l'impression d'être passif, et trouve la vision plus agréable.
- le sujet `6` se rapproche des sujets `1`, `2` et `7` pour ce qui est de sa préférence pour la manipulation, mais il aurait aimé avoir la possibilité de faire des déplacements de type éloignement-rapprochement, comme dans la métaphore de la locomotion.
- le sujet `8` préfère la métaphore de la locomotion sans restriction.

Cela confirme les faiblesses de l'implémentation de la locomotion, au niveau de l'agrandissement des photographies, puisque cette fonctionnalité s'avère plus utile dans la tâche 1, qui est par nature une tâche d'exploration. Mais la locomotion, en tant que métaphore nouvelle, ne bénéficie pas encore des moyens logiciels dont nous disposons actuellement pour la manipulation qui s'inspire fortement des fonctionnalités utilisée dans les interfaces 2D classiques.

5.6 Conclusions générales

Cette étude préliminaire exploratoire visait à comparer deux vues d'un cylindre vertical, associée à deux métaphores d'interaction, la manipulation et la locomotion. Au delà des hypothèses de travail, il s'agissait non seulement d'évaluer l'intérêt d'utilisateurs potentiels pour cette nouvelle représentation 3D, mais également de valider le protocole expérimental, en particulier le recours aux descriptions textuelles pour la tâche 1.

Dans un premier temps, nous faisons une synthèse des résultats quantitatifs et qualitatifs obtenus, avec un retour sur les hypothèses de travail. Nous discutons ensuite la pertinence du choix de respecter strictement les métaphores d'interaction choisies lors de l'implémentation de l'interaction. Cette discussion nous permettra d'esquisser les améliorations à mettre en œuvre, en vue de la réalisation de l'expérimentation présentée dans le chapitre suivant.

5.6.1 Synthèse des résultats

En termes de rapidité, si la recherche dans VE est plus rapide pour 6 sujets sur 8 dans la tâche 1 (2 participants ont des temps voisins dans les 2 vues), dans la tâche 2 elle s'avère nettement plus rapide pour 4 sujets avec VI quand 3 autres sujets obtiennent des meilleurs temps dans VE. Du reste, pour un même sujet, les temps de recherche sont très variables d'une vue à l'autre. Logiquement, les temps de recherche des cibles sont plus courts dans la tâche 2 (T2), comparés à ceux de la tâche 1 (T1) en raison de la différence de difficulté de ces tâches. Concernant les succès des recherches visuelles, aucune vue n'est significativement supérieure à l'autre, même si la vue externe semble légèrement meilleure. Concernant la tâche 1, cela peut s'expliquer par la non-homogénéité des cibles. Concernant la seconde tâche, le nombre de succès est très proche d'une vue à l'autre pour un même sujet et toujours très proche du taux de réussite maximum pour l'ensemble des sujets, ce qui s'explique par la simplicité de ce type de tâche.

Le repérage spatial semble meilleur dans la vue interne : dans la tâche 2, le recours systématique des sujets à un examen initial dans VE suppose une mauvaise mémorisation de la position des cibles (celles-ci n'étaient jamais présentes sur la vue initiale) ; le choix fréquent d'un sens de rotation constant avec VE (pour T2) suppose également une mauvaise mémorisation de la position de la cible dans la visualisation, et donc du sens optimal de rotation à choisir pour y accéder.

Comme le supposait l'hypothèse B, la rapidité est influencée par la nature de la tâche et le mode d'interaction. En effet, la tâche 1 impose un parcours séquentiel relativement lent de la collection, parcours au cours duquel le sujet confronte les photographies affichées à la représentation mentale qu'il s'est construite de la cible. Durant cette recherche, le sujet est donc amené à agrandir plus de photographies pour comparer sa représentation mentale de la cible avec les candidates possibles. Dans la locomotion, le sujet est handicapé par le coût temporel du déplacement de type éloignement-rapprochement, contrairement au zoom de la manipulation qui peut se faire en un seul clic. Pour éviter d'avoir à s'approcher, les sujets arrêtaient le défilement pour observer une image plus en détail, ce qui a pour effet d'allonger notablement le temps de recherche. La tâche 2 se prête davantage au parcours « libre » de la collection, parcours dans lequel le sujet adopte une vue plus globale, à la recherche de la photographie qui l'intéresse (*i.e.*, la cible) dont il connaît à la fois l'apparence et la position. Le fait qu'il ne soit pas obligé d'agrandir une photographie dans cette tâche gomme donc les problèmes d'utilisation liés aux moyens d'interactions offerts par la locomotion. Ces remarques suggèrent également qu'il est nécessaire de modifier la conception du zoom dans la locomotion, pour atteindre le même niveau de performances que le zoom de la manipulation.

5.6.2 Préférences des sujets

Sur les 8 sujets, 7 ont accueilli favorablement la représentation 3D sous forme de cylindre. Si les préférences globales sont partagées entre les vues interne et externe du cylindre, elles semblent néanmoins pencher vers la vue interne : 3 sujets préfèrent cette vue, et 3 autres lui donneraient leur préférence, à condition de disposer d'un mécanisme de zoom efficace. Ce résultat, ramené au nombre restreint de sujets, ne nous permet pas d'avoir une opinion tranchée sur la

« meilleure » métaphore. Nous pouvons remarquer que l'influence de la géométrie des vues est grande sur les préférences, contrairement à son influence sur les performances. Cela contredit l'hypothèse C, mais se justifie par le fait que certains participants placent le confort visuel au-dessus du niveau de performance. Le confort visuel supérieur de la vue interne est d'ailleurs souligné par 7 des 8 sujets, car les déformations naturelles que celle-ci impose sont plus en accord avec les stratégies de parcours visuel : lors de la rotation du cylindre, l'attention se porte sur les colonnes « entrantes », pour lesquelles les images sont les plus grandes dans la vue interne. Ajoutons que la non-équivalence des déformations entre les 2 vues est également de nature à influencer le confort visuel, la variation de la taille des photographies dans la vue interne étant moins marquée que dans la vue externe, où les photographies sur les bords latéraux de l'écran présentent d'importantes déformations et sont de taille très petite. Ces résultats sont en accord avec l'hypothèse A, formulée au début de ce chapitre.

La compréhension des métaphores diffère également. D'une part, les sujets ont préféré l'utilisation du point de vue courant (la carte) de la vue interne, car il leur était plus facile d'associer « leurs » mouvements avec ceux du curseur (la barre rouge sur la carte). Dans la vue externe, le fait que la bande transparente tourne semble avoir été déroutant. Nous pouvons interpréter cela comme une difficulté à assimiler et à intérioriser la métaphore de la manipulation. L'ordre de présentation des vues a pu également exercer une influence en faveur de la vue interne que tous les sujets ont utilisée en premier.

5.6.3 Respect des métaphores et améliorations

Comme nous avons pu le voir concernant les performances, la fonctionnalité permettant l'examen détaillé d'une photographie (zoom) a joué un rôle non négligeable dans les résultats obtenus. Cela est également vrai pour les jugements subjectifs, qui accordent une importance prépondérante à cette fonctionnalité pour la tâche 1. On peut reconnaître que l'implémentation du zoom dans la métaphore de la locomotion était perfectible. C'est sans doute le prix à payer pour avoir pris le parti de respecter au maximum les caractéristiques de la métaphore. Cela étant, la lenteur et la complexité de mise en œuvre sont davantage des conséquences indirectes du défilement discontinu, choisi pour répondre aux contraintes techniques d'alors. Aucun participant ne s'est réellement plaint de la nature discontinue du mouvement dans le cas d'une simple rotation. Mais les participants, pour pouvoir agrandir une photographie, devaient attendre qu'elle arrive au centre avant d'effectuer le mouvement de rapprochement. Et c'est dans ce cas que la lenteur du mouvement a été perçue négativement.

Comme cela a été suggéré par certains participants, il serait possible d'automatiser la phase de zoom dans la locomotion, mais cela nécessiterait un clic de sélection d'une image, ce qui est davantage conforme à la métaphore de la manipulation. Tous les sujets, sauf un, ont apprécié le zoom implémenté dans la vue externe, et 3 (`S1`, `S2` et `S7`) en font une condition nécessaire pour donner leur préférence à la métaphore de la locomotion. L'adoption, dans la vue interne, du zoom de la métaphore de la manipulation est donc à envisager. La fidélité à la métaphore de la locomotion s'en trouve diminuée, il est vrai, mais la compréhension de cette métaphore ne devrait pas s'en trouver affectée, car elle s'exerce davantage sur les déplacements de rotation. De ce fait, il faut conserver des interacteurs spécifiques à chaque vue, ainsi que les modes de

déplacement associés, dans un référentiel allocentré (celui du cylindre) pour la vue interne, et dans un référentiel égocentré pour la vue externe. Même si cette différence de référentiel n'est pas forcément bien assimilée par l'ensemble des sujets, elle est importante pour la cohérence de chaque métaphore. Un avantage supplémentaire de cette solution est de doter les deux vues de moyens d'interaction strictement identiques, ce qui facilite la comparaison entre les deux vues en termes d'efficacité et de confort d'utilisation. Enfin, déplacer rapidement l'ensemble de la visualisation, d'abord pour placer au centre de l'affichage la photographie à agrandir, puis pour la rapprocher de l'utilisateur, risque de s'avérer fatigant visuellement, surtout si ces déplacements sont effectués rapidement et de manière continue de façon à obtenir des temps d'affichage identiques des images agrandies pour les deux vues.

Mis à part l'agrandissement des photographies, le second point à améliorer est la rotation. Comme nous venons de le voir, le défilement discontinu n'est pas critiquable en tant que tel. Mais c'est le fait que sa vitesse (ou plutôt sa fréquence d'activation) soit fixe qui a posé des problèmes. Souvent jugée trop faible dans T2, et trop rapide dans T1, la vitesse ne pouvait pas être adaptée aux capacités et aux souhaits de chacun. Il serait intéressant de pouvoir moduler la vitesse au cours d'une recherche, car le flux d'informations « intéressantes » n'est pas constant durant la rotation. Le passage à l'atelier de réalité virtuelle *AReVi* a permis de gagner suffisamment en performances pour rendre possible l'implémentation des différentes améliorations proposées ici lors de la mise en œuvre de l'étude expérimentale décrite dans le chapitre suivant.

5.6.4 Conclusion

L'étude préliminaire a tenu le rôle que nous attendions d'elle, c'est-à-dire permettre de valider les différents choix de conception concernant les vues, les moyens d'interaction, les tâches et donc, globalement, aussi bien le protocole expérimental que la réalisation des métaphores. Au vu des informations recueillies, l'intérêt de la représentation cylindrique pour l'affichage des collections de données non structurées de photographies est réel. Même si la métaphore de la locomotion semble prometteuse, sous réserve de procéder à quelques améliorations, il faut maintenant comparer les deux vues de manière plus approfondie, en faisant intervenir un plus grand nombre de sujets, afin d'être en mesure d'établir laquelle des deux vues est la plus à même de remplacer efficacement les représentations matricielles classiques.

Chapitre 6

Première étude

La conception de cette étude repose, comme l'étude préliminaire, sur une approche expérimentale. Les deux vues (intérieure et extérieure) du cylindre 3D vertical, sur lesquelles sont plaquées 869 photographies, constituent le matériel visuel. Ce chapitre décrit l'étude empirique menée sur 20 participants en vue d'évaluer l'efficacité, la qualité du repérage spatial et le confort de chaque vue, pour deux types d'activités visuelles : la recherche, à partir d'une description verbale, d'une photographie inconnue, dont la position dans la visualisation est également inconnue, et la recherche d'une photographie familière dont la position dans la collection est connue.

L'étude préliminaire a démontré l'intérêt de huit utilisateurs potentiels pour notre représentation. Les résultats indiquent, entre autres, un meilleur repérage spatial dans la vue interne. En tenant compte des remarques des participants, concernant principalement les améliorations à apporter à ces deux vues interactives, les moyens d'interaction avec elles ont été partiellement repensés : la gestion de l'agrandissement d'une photographie et les interacteurs contrôlant le déplacement des vues dans l'espace virtuel ont été revus. Le but poursuivi est de rendre l'interaction plus simple et intuitive, tout en facilitant la comparaison entre les deux vues. Ces modifications seront explicitées et motivées au début de ce chapitre.

En résumé, d'après les questionnaires finaux et les entretiens de débriefing, 10 participants à cette première étude ont préféré la vue interne, 4 la vue externe, et 6 n'ont pas exprimé de préférence marquée. 12 participants ont réalisé de meilleures performances avec l'une ou l'autre vue, et les mesures suggèrent une légère supériorité de la vue interne sur la vue externe. Deux stratégies d'exploration visuelle ont été observées : lors du défilement latéral, les participants portaient leur attention soit sur les colonnes « entrantes » (colonnes latérales), soit sur les colonnes centrales. Les distorsions géométriques, comme la nature du défilement, ont exercé une influence sur les stratégies d'exploration visuelle. L'influence du caractère immersif de la vue interne s'est, en revanche, avérée marginale. L'analyse des comportements indique que les visualisations 3D ont le potentiel de favoriser la mémoire spatiale. Ces résultats suggèrent que les vues en perspective peuvent faciliter la recherche visuelle au sein de collections non structurées de documents visuels.

Le plan de ce chapitre est similaire au précédent. Nous présentons la méthodologie, avec les hypothèses que nous souhaitons vérifier, et les modifications apportées aux représentations. Dans un second temps, nous exposons le protocole expérimental et les méthodes d'analyse des données

recueillies. Nous terminons par le commentaire des résultats quantitatifs et qualitatifs.

6.1 Méthodologie

Le chapitre précédent, relatif à l'étude préliminaire, a permis de valider les tâches de recherche proposées aux sujets. Chacune de ces 2 tâches permet de simuler de manière réaliste les activités auxquelles les utilisateurs sont fréquemment confrontés, que ce soit pour des besoins personnels ou professionnels : d'une part, chercher des documents visuels d'après des caractéristiques visuelles et thématiques peu ou mal explicitées, et d'autre part, retrouver des documents visuels familiers. Ces tâches impliquent exploration visuelle et navigation. Nous les avons reprises dans cette première étude, pour leur caractère complémentaire, réaliste et discriminant. Complémentaires, car elles sont indissociables et souvent entremêlées dans une recherche visuelle usuelle. Réalistes, car le parti pris de l'usage de descriptions textuelles est la solution la plus cohérente pour obtenir des résultats satisfaisants sur le plan de la validité externe. Enfin, discriminantes, car nous avons obtenu des résultats très différents d'un sujet à l'autre. Par exemple, la tâche 2, davantage centrée sur une recherche visuelle rapide, est susceptible de fournir des résultats sensiblement différents de ceux de la tâche 1 : les stratégies visuelles peuvent évoluer avec une vitesse de rotation plus importante. Il est donc envisageable qu'une vue soit plus adaptée à une tâche de recherche que l'autre, et inversement. Les résultats de l'étude préliminaire semblent aller en ce sens, bien que l'influence de la lenteur du zoom dans la vue interne ne soit pas négligeable.

Si l'apport de la troisième dimension a été favorablement accueilli par 7 des 8 participants à l'étude préliminaire, le caractère immersif de la vue interne n'a pas été perçu par l'ensemble des participants avec la force que nous aurions pu attendre. Cela peut s'expliquer par la taille de l'écran utilisé dans l'étude préliminaire, trop limitée pour susciter ce sentiment. Dans l'étude que nous détaillons dans les paragraphes suivants, nous avons fait le choix de suggérer plus fortement la perception de l'immersion en augmentant fortement la taille de l'écran (voir le paragraphe 6.1.6).

6.1.1 Présentation générale

Dans l'étude présentée ici, nous conservons le scénario conçu pour l'étude préliminaire :
- la recherche d'une photographie non familière, dont la position est inconnue : les caractéristiques de cette photographie sont données sous la forme d'une description textuelle qui la caractérise de manière non ambiguë par rapport aux autres photographies ;
- la recherche d'une photographie familière dont la position est connue : cette photographie est présentée, isolée, au centre de l'écran pendant 3 secondes.

Les photographies utilisées dans la seconde tâche sont identiques à celles utilisées dans la première tâche.

La différence majeure par rapport à l'étude préliminaire est l'augmentation du nombre de photographies à chercher par vue et par tâche. Dans l'étude précédente, nous avons remarqué qu'il est difficile d'obtenir une difficulté homogène d'une cible à l'autre, malgré nos efforts. Les facteurs limitants sont tout autant la saillance visuelle (intrinsèque et/ou relative par rapport

aux photographies contiguës de la collection) que la variabilité culturelle des participants, même si un soin particulier a été apporté aux choix de ces derniers. Afin de pouvoir réaliser une étude susceptible de fournir des résultats statistiquement significatifs, nous avons choisi d'utiliser 10 cibles par vue, et non plus 6 comme dans l'étude préliminaire. Ce nombre représente un compromis entre validité statistique et temps d'expérimentation, que nous avons souhaité limiter à 2 heures 30. Une durée supérieure aurait risqué de fatiguer les sujets, ce qui aurait pu avoir une influence négative sur leurs performances et leurs préférences.

6.1.2 Terminologie

Voici les différents acronymes et termes que nous allons utiliser dans la suite de ce chapitre :
- collection : ensemble de photographies, au nombre de 869, plaquées sur la face (interne ou externe) du cylindre ;
- cible : photographie particulière donnée à chercher aux participants. Chaque cible fait partie de la collection courante ;
- recherche : phase de l'expérimentation qui correspond à la recherche visuelle d'une cible dans une collection ;
- tâche : ensemble des 10 recherches de cibles correspondant au premier scénario (recherche d'une photographie non familière, dont la position est inconnue) ou au second scénario (recherche d'une photographie familière dont la position est connue) ;
- vue : soit la face intérieure du cylindre (vue interne), soit la face extérieure (vue externe).

Nous remplacerons vue interne par **VI**, vue externe par **VE**, tâche 1 par **T1** et tâche 2 par **T2**.

6.1.3 Hypothèses à tester

Cette étude a pour but premier d'évaluer comparativement les deux vues du cylindre vertical, autant en termes de performances que de préférences. Il s'agit également de mesurer l'influence des performances sur les préférences et, plus généralement, de caractériser quels peuvent être les facteurs d'influence sur les préférences. Enfin, nous souhaitons valider les hypothèses de l'étude préliminaire, concernant la supériorité de la vue interne sur la vue externe en termes de confort d'utilisation. Nos hypothèses sont les suivantes :
- hypothèse A : la vue externe permet d'obtenir de meilleures performances que la vue interne, car les distorsions générées par la perspective induisent un meilleur guidage du regard dans la vue externe. En effet, les deux vues, bien qu'affichant le même nombre d'objets simultanément, offrent des distorsions non équivalentes : dans la vue externe, une partie de l'écran plus réduite que dans la vue interne présente une forte saillance (les colonnes centrales), et la différence de taille entre colonnes centrales et colonnes latérales est plus importante ;
- hypothèse B : même si les deux vues ont le potentiel de faciliter la mémorisation spatiale, la métaphore de la locomotion et le sentiment d'immersion induits par la vue interne favorisent davantage la mémorisation de la position des cibles sur le cylindre, sur la base d'études démontrant que les personnes établissent des relations spatiales entre les objets (*object-to-object*) d'un environnement réel inconnu qu'elles découvrent en s'y déplaçant

[Mou *et al.*, 2006] ;
- hypothèse C : les caractéristiques géométriques des vues sont de nature à influencer de manière non négligeable, outre les performances des utilisateurs (hypothèse A), leurs préférences et leurs stratégies d'exploration visuelle ;
- hypothèse D : la vue interne est plus adaptée à la tâche 2, qui demande un examen rapide des photographies affichées à l'écran, d'une part, parce qu'elle offre une meilleure vue globale et d'autre part parce que les déformations entre les colonnes centrales et les colonnes latérales sont moins marquées.

Dans les trois sections suivantes, nous présentons les modifications apportées à l'interaction avec les deux vues, et décrivons brièvement l'implémentation et les développements logiciels.

6.1.4 Matériel visuel et représentation

Comme pour l'étude préliminaire, la représentation choisie est un cylindre vertical, offrant deux vues distinctes, interne et externe. La vue interne est plus à même de favoriser un sentiment d'immersion que la vue externe, tandis que celle-ci permet de renforcer l'impression de manipulation directe, suivant le même schéma que les représentations 2D actuelles. A ce titre, nous avons conservé la même disposition, à savoir 11 lignes et 80 colonnes, soit au total 869 photographies, car nous laissons une colonne vierge de toute photographie pour donner au cylindre un « début » et une « fin ». Cela permet, tout en conservant le bénéfice de la continuité du défilement horizontal des vues, d'offrir un point de repère global facile à appréhender. Comme cela a été précisé dans le chapitre précédent, la courbure de la partie visible du cylindre a été définie empiriquement, de façon à obtenir un compromis satisfaisant entre deux contraintes contradictoires : d'une part, rendre perceptible les déformations dues à la perspective (i.e., la troisième dimension), d'autre part, limiter ces déformations (taille et forme) pour éviter d'augmenter sensiblement la difficulté de la recherche visuelle par rapport aux vues 2D. Le choix du nombre de photographies affichées simultanément a été dicté par la volonté de donner une taille raisonnable à l'ensemble des photographies visibles à chaque instant[14].

Le matériel visuel est composé des mêmes photographies que dans l'étude préliminaire : des documents numériques libres de droit, provenant de sites web. Ces photographies sont classées en thèmes, eux-mêmes organisés en deux catégories, objets et paysages[15]. Les cibles ont été sélectionnées parmi ces milliers de photographies.

Comme cela a été expliqué précédemment, nous nous appuyons sur des photographies, plutôt que tout autre document visuel, car c'est le média le plus fréquemment manipulé par le grand public. L'usage de photographies offre de plus un cadre réaliste et particulièrement bien adapté aux tâches de recherche que nous proposons.

[14] cf. paragraphe 6.2.5
[15] cf. paragraphe 4.2.1

6.1.5 Interaction avec la représentation

Les fonctionnalités de l'interface utilisateur sont reprises de l'étude préliminaire, car elles offrent les moyens d'intraction nécessaires au parcours de la collection dans des conditions optimales. Les études antérieures ont montré que les utilisateurs n'utilisent que les fonctionnalités les plus simples [Rodden et Wood, 2003], ce qui fournit une première justification à ce choix. La seconde justification s'appuie sur les commentaires des participants de l'étude préliminaire : mis à part quelques adaptations ou changements mineurs sur un aspect particulier (dispositif de zoom, défilement discontinu à vitesse fixe), aucun participant n'a demandé de fonctionnalités supplémentaires.

Avant de présenter les modifications apportées à l'interaction, rappelons brièvement les fonctionnalités retenues :

- déplacement : le déplacement doit pouvoir se faire dans deux dimensions. D'une part, le déplacement latéral à distance constante de la paroi du cylindre, c'est-à-dire une rotation par rapport à l'axe (vertical) du cylindre. D'autre part, un déplacement perpendiculaire au plan de l'écran (éloignement ou rapprochement de la vue), afin de moduler la distance entre le cylindre et le point de vue de l'utilisateur. Le mouvement de rotation correspond au défilement produit par les ascenseurs dans les interfaces 2D classiques : il permet l'exploration de la collection. Le déplacement perpendiculaire au plan de l'écran permet de faire varier continûment le niveau de détail et de ce fait le nombre d'images affichées simultanément. Il correspond au facteur de zoom, que l'on retrouve par exemple dans les logiciels *Microsoft Office* ou *Adobe Acrobat*, et il permet d'adapter l'interface aux capacités visuelles de chacun. Comme nous l'avons précisé précédemment, nous préférons « brider » le second type de déplacement, quitte à perdre une partie de l'intérêt de la 3D, pour rendre meilleure la compréhension de l'interface et faciliter son utilisation. Le déplacement vertical, parallèlement à la paroi du cylindre, est quant à lui rendu nécessaire lorsque la distance au cylindre empêche de visualiser l'ensemble des 11 lignes simultanément ;
- zoom : dans le cadre d'une recherche et/ou d'une exploration, il peut être nécessaire de moduler le niveau de détail d'une photographie particulière ;
- dispositif de repérage : la colonne sans photographie indique le parcours d'un tour complet. Cette information précieuse est pourtant incomplète, car elle ne permet pas de connaître à tout instant la position des photographies visibles dans la représentation visuelle de la collection courante. Il est nécessaire de disposer d'un dispositif supplémentaire répondant à ce besoin.

Améliorations par rapport à l'étude préliminaire

Dans l'étude préliminaire, de nombreux participants ont jugé inadapté le dispositif de zoom de la vue interne, car trop lent et fastidieux à mettre en œuvre. Les sujets auraient préféré un dispositif de zoom identique à celui de la vue externe. Même si le défilement discontinu a sans doute pesé fortement sur les jugements négatifs portés sur le zoom dans VI, il fallait revoir cet aspect. Le second point sur lequel nous avons porté notre attention a trait à la commande des déplacements, en particulier du défilement latéral. Le mode d'interaction proposé pour VE, qui

obligeait l'utilisateur à faire un *drag* sur la paroi du cylindre en dehors des photographies, s'est révélé peu pratique à utiliser en raison de la précision qu'il exige. Enfin, la discontinuité du mouvent de rotation, associée à l'impossibilité d'en moduler la fréquence, avait posé problème aux participants de l'étude préliminaire. Les raisons techniques qui avaient imposé ce choix ont pu être éliminées grâce à l'utilisation d'un nouvel atelier de réalité virtuelle.

Nous allons maintenant présenter le dispositif d'interaction de chaque vue.

Déplacements : interacteurs

L'idée que nous avons suivie est « d'unifier » les interacteurs utilisés dans chaque vue, en tenant compte des points à améliorer (précédemment exposés) et du respect de ces deux contraintes :

- faire cohabiter dans le même objet graphique commande et informations sur la position courante. C'est-à-dire se rapprocher du modèle de la barre de défilement (*scroll-bar*), pour que l'utilisateur puisse se focaliser uniquement sur la recherche, sans avoir à regarder à de multiples endroits pour réaliser une action ;
- adapter cette commande à la spécificité de la vue : déplacement « dans » la vue interne et déplacement « de » la vue externe.

Les premières ébauches (figure 6.1) n'ont pas été retenues, car leur compréhension n'était pas évidente, compte tenu de la taille nécessairement réduite de cet interacteur qui faisait également office d'indicateur de position. En ce sens, nous rejoignons le point de vue exprimé dans [Trumbo, 1998], à savoir que des objets sous forme de pictogrammes sont plus rapidement reconnaissables, comparés à des objets avec un niveau de détail élevé. Il était donc préférable de réaliser un interacteur simple, permettant un usage immédiat. L'idée première était de rendre évident le point de vue de l'utilisateur sur le cylindre par l'emploi d'un avatar simplifié. Nous avons conservé

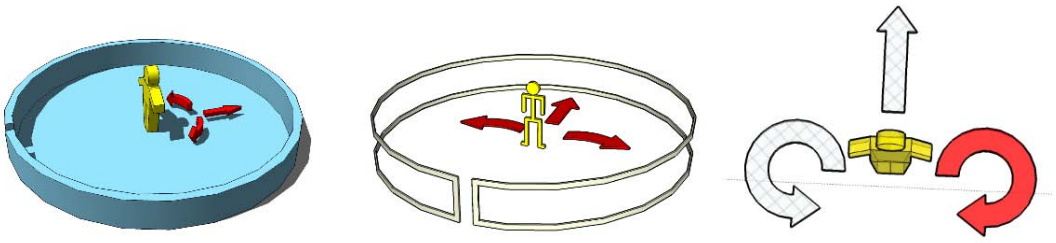
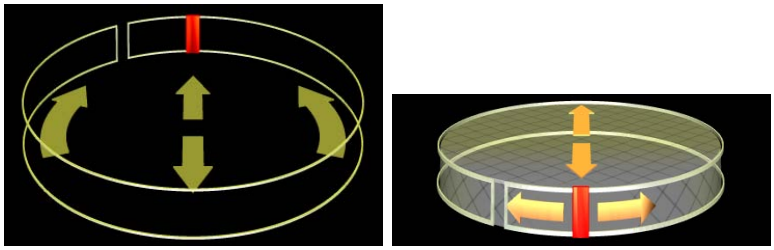


FIG. 6.1 – Trois projets non retenus d'interacteurs

cette idée mais sans recourir à une représentation de l'utilisateur sous forme d'avatar afin de matérialiser en quelque sorte son point de vue courant sur la visualisation. Nous avons repris l'interacteur utilisé lors de l'étude préliminaire, qui associe une direction de déplacement à un bouton en forme de flèche pointant dans la direction correspondante. Ce dispositif avait été apprécié pour son caractère intuitif.

Pour contrôler la rotation, le rapprochement et l'éloignement des vues cylindriques, l'utilisateur dispose donc, en bas de l'écran, d'une reproduction miniature (homothétique) du cylindre. La différenciation entre les deux vues se retrouve dans la position des boutons-flèches sur ce

cylindre. La rotation du cylindre est activée par un clic gauche sur l'une des deux flèches horizontales « collées » sur les parois du cylindre (VE) ou placées dans le cylindre (VI) (voir figure 6.2). Il suffit d'effectuer un clic droit sur n'importe quelle partie de l'affichage pour arrêter le mouvement. Les deux autres flèches, sur le « couvercle » (VE) ou sur le « fond » (VI) du cylindre



(a) Vue interne (b) Vue externe

FIG. 6.2 – Interacteurs des vues interne et externe

miniature, permettent d'ajuster la distance entre l'utilisateur et la représentation, pour permettre l'adaptation des visualisations aux capacités visuelles de chacun. La position de l'interacteur (en bas de l'écran, au centre) a imposé une adaptation supplémentaire dans la vue externe, car il se trouve « devant » la représentation et masque de ce fait quelques photographies. Ainsi, la paroi translucide n'est visible qu'à l'arrêt du cylindre, pour mieux discerner les flèches. Dès le début de la rotation, les flèches supérieures et la paroi disparaissent : seuls le cylindre et les flèches latérales subsistent, ce qui limite le phénomène de masquage. Dans la vue interne, du fait de sa forme concave, le cylindre miniature, qui est également au premier plan, ne masque pas la visualisation car il est situé dans une zone de l'affichage vierge de toute photographie.

Déplacements : vitesse

Compte tenu de la levée des contraintes techniques, le défilement peut être continu, pour un meilleur confort visuel. Dans l'étude préliminaire, certains sujets ont observé que la vitesse de défilement était trop faible dans la tâche 2, mais trop rapide dans la tâche 1. Pour tenir compte de la variabilité inter-individuelle et intra-individuelle, nous avons décidé de laisser aux utilisateurs la liberté de moduler la vitesse, selon 12 pas. Par défaut, la vitesse initiale correspond au premier pas, celui où la vitesse est la plus faible. C'est la molette de la souris qui permet la modulation de la vitesse de défilement, une fois la rotation lancée par la sélection de l'une des flèches.

Zoom

Un dispositif classique de zoom est proposé. Il est activé par un clic sur l'image à agrandir ; l'image agrandie apparaît au centre de l'écran et disparaît lorsque l'utilisateur clique sur elle. Ce dispositif est utilisé dans les deux vues, car il surclasse actuellement tous les autres dispositifs, et l'utilisation exclusive de la souris comme périphérique d'interaction se prête mal à la mise en œuvre d'autres possibilités. Cette solution a le grand avantage d'être immédiatement assimilée par la grande majorité des utilisateurs.

Système de repérage

Un dispositif permettant un repérage de la position courante est également proposé. La colonne affichée au centre de l'écran est matérialisée sur le cylindre miniature par une barre verticale rouge. Les mouvements de la représentation et ceux du cylindre miniature sont synchronisés. Ainsi, l'utilisateur peut connaître à chaque instant la position de la colonne centrale de l'affichage par rapport à la colonne vierge.

Discussion

Nous avons constaté dans l'étude précédente que les caractéristiques géométriques des vues semblent exercer une grande influence sur les préférences des participants. L'interaction, même si elle a suscité de nombreuses remarques, semble avoir joué un rôle secondaire. C'est pourquoi nous pouvons prendre le risque de modifier le mode d'interaction relatif aux deux métaphores, en allant vers un système mixte, identique pour les deux vues, qui cumule les avantages des solutions adoptées pour l'étude préliminaire tout en supprimant leurs inconvénients. Ainsi le zoom de la vue externe de l'étude préliminaire a été généralisé à la vue interne. Il en est de même pour l'interacteur de la vue interne. Les interacteurs des vues interne et externe sont légèrement différents les uns des autres pour être en cohérence avec chacune des métaphores qu'ils implémentent respectivement. Cependant, leurs fonctionnalités sont strictement identiques, ce qui garantit la validité des comparaisons entre les performances et jugements subjectifs des participants lors de l'utilisation de chacune d'elles pendant T1 et T2.

6.1.6 Implémentation et développements logiciels

L'implémentation des vues, ainsi que l'application permettant le déroulement de l'expérimentation (passage d'une recherche et d'une tâche à l'autre), a été réalisée en C++, en utilisant l'atelier de réalité virtuelle *AReVi* [Harrouet *et al.*, 2006]. D'un point de vue pratique, cela s'est traduit par une augmentation des performances brutes (fréquence de rafraîchissement des affichages), ainsi que des possibilités offertes. D'une part, nous n'étions plus limités pour la qualité des photographies : ainsi les photographies plaquées sur le cylindre sont toutes d'une définition de 256x192 pixels, utilisée également pour le zoom. D'autre part, *AReVi* permet l'importation de modèles *3dsMax*, ce qui permet d'utiliser des objets 3D complexes et offre, de ce fait, une plus grande liberté pour la conception des interacteurs. Le cylindre miniature ainsi que les flèches ont été réalisés avec le logiciel *SketchUp*, puis exportés au format *3dsMax* pour être utilisés dans l'application.

Nous avons utilisé un affichage mural d'une diagonale de 1,25 m (vidéo-projection), afin d'accroître les effets éventuels des métaphores d'interaction utilisées, manipulation d'un objet virtuel d'une part, et immersion dans un espace virtuel d'autre part. Les participants étaient situés à 2,20 m du mur. L'ordinateur utilisé était un Dell Precision 370, doté d'un processeur *Intel Pentium 4* à 3 Ghz, de 2 giga-octets de mémoire vive et d'une carte graphique *NVidia Quadro FX 3000*. La résolution d'affichage était de 1280x1024 pixels.

Les figures 6.3 et 6.4 présentent les vues interne et externe.



FIG. 6.3 – Vue intérieure du cylindre vertical

6.2 Protocole expérimental

6.2.1 Généralités

Nous utilisons les tâches de sélection de cibles visuelles en temps borné de l'étude préliminaire. La première tâche (T1) correspond à la recherche de cibles d'après une description textuelle, tandis que la seconde correspond à la recherche de ces mêmes cibles après une présentation visuelle de 3 secondes.

Nous avons retenu un total de 24 photographies pour servir de cibles, composées d'un nombre identique d'objets et de paysages. Nous avons conservé le principe des cibles d'entraînement pour permettre aux participants d'assimiler à la fois les tâches à réaliser et l'interaction avec chaque vue de la représentation cylindrique. Dans T1, les participants avaient à chercher 10 cibles *plus* 2 cibles d'entraînement par vue. Les mêmes cibles ont été utilisées à nouveau dans T2, placées de manière identique dans la même collection. Seul l'ordre de présentation des cibles était différent d'une tâche à l'autre, pour une même vue.

6.2.2 Choix des cibles et descriptions textuelles

La figure 6.5 présente les 10 cibles utilisées pour chaque vue. Les 4 cibles d'entraînement ne sont pas présentées, car elles n'ont pas été prises en compte dans l'analyse des données recueillies. Par rapport à l'étude préliminaire, certaines photographies, qui avaient engendré beaucoup d'échecs (comme par exemple le sentier s'enfonçant dans la forêt, ou l'éclair derrière le



FIG. 6.4 – Vue extérieure du cylindre vertical

poteau télégraphique), n'ont pas été réutilisées. Elles ont été remplacées par d'autres, puis nous avons complété cette sélection pour atteindre le nombre de 24 cibles (contre 16 précédemment). Une fois encore, les cibles ont été choisies pour leur saillance et leur facilité à être décrites verbalement.

La structuration des descriptions a également été légèrement revue, pour gagner en clarté et simplifier le « travail » du sujet. Ainsi, la description est maintenant décomposée en 4 parties, au lieu de 2 dans l'étude préliminaire :

1. indication du thème général de la photographie et de ses caractéristiques visuelles globales ;
2. informations sur le cadrage et le thème général de la photographie ;
3. informations sur la composition (arrière-plan, objets secondaires) ;
4. précisions complémentaires.

La première partie permet un pré-filtrage de la recherche, et ramène la décision à quelques photographies au maximum. Les trois parties suivantes détaillent et complètent la première. Elles sont destinées essentiellement à faciliter la construction d'une représentation mentale fidèle de la cible et, plus rarement, à la caractériser de manière unique afin de la distinguer des autres photographies avec lesquelles elle pourrait être confondue sans de telles informations (thème, couleurs, formes). La quatrième partie n'est pas systématiquement présente ; sa présence est fonction de la complexité de la cible. La figure 6.6 présente la description de la cible correspondant à la tête d'autruche.

Les 20 cibles choisies, ainsi que leurs descriptions, sont présentées dans l'annexe B.



(a) Vue intérieure



(b) Vue extérieure

FIG. 6.5 – Cibles

La tête d'une autruche

La photographie est un gros-plan sur la tête d'une autruche, vue de profil. L'autruche regarde vers la droite.

L'arrière plan est composé d'un ciel bleu uni, s'assombrissant légèrement vers le haut de la photographie.

La tête de l'autruche est de couleur gris-beige. Sous l'œil, une traînée rouge-orangée, en direction du bec.

FIG. 6.6 – Exemple de description

6.2.3 Constitution des collections

Comme pour l'étude préliminaire, nous disposions de 4 séries composées de 920 photographies (40 thèmes illustrés chacun par 23 photographies), comprenant autant d'objets que de paysages. Une série a été réservée aux recherches d'entraînement et nous avons divisé les 3 séries restantes en deux, de façon à disposer de bases de photographies différentes pour la vue interne et la vue externe. Soit potentiellement 1380 photographies par vue, avec lesquelles nous avons constitué les 10 collections, ce qui a donc imposé un renouvellement partiel d'une collection à l'autre. Un renouvellement intégral aurait nécessité de disposer d'un nombre de photographies trop important (*i.e.*, plus de 20 000). En plus du renouvellement partiel, la disposition des photographies est modifiée d'une collection à l'autre.

Chaque cible apparaît seulement dans la collection où elle est l'objet de la recherche, de même pour toutes les photographies du même thème. Ce qui implique que nous ne pouvons utiliser

effectivement que 1173 photographies pour constituer les 10 collections d'une vue auxquelles s'ajoutent, pour chaque collection, la cible et les 22 autres photographies du thème dont elle relève. Ainsi, pour constituer une collection particulière d'une des deux vues, nous avons opéré de la manière suivante :

1. utilisation des 23 photographies du thème de la cible (dont la cible) ;
2. tirage aléatoire de 423 photographies parmi les 575 photographies représentant des paysages ;
3. tirage aléatoire de 423 photographies parmi les 575 photographies représentant des objets.

Nous sommes sûrs ainsi que le sujet n'a vu aucune des photographies correspondant aux cibles, ni aucune des photographies thématiquement proches, avant de les rechercher dans la visualisation.

6.2.4 Placement des cibles

Afin de nous donner les moyens d'étudier la capacité des participants à mémoriser les cibles et leur position dans la visualisation, nous avons réparti les différentes cibles choisies pour une même vue, spatialement (verticalement et horizontalement) de manière non aléatoire. Disposant de 10 cibles par vue, nous les avons réparties verticalement sur la hauteur du cylindre(celui-ci étant composé de 11 lignes) et placées horizontalement toutes les 8 colonnes environ (voir figure 6.7). L'association position verticale / position horizontale s'est faite aléatoirement, tout comme le choix des cibles à placer, afin de maintenir la validité interne de notre expérimentation (voir paragraphe 6.3.3). Nous avons placé 2 des 10 cibles dans l'affichage initial (aux colonnes 1 et 7), en bas et en haut de la représentation (aux lignes 1 et 10). D'une part, cela nous permettait de vérifier facilement, lors de T2, la qualité de la mémorisation spatiale (en utilisant comme critères l'absence de rotation et le succès de la recherche). D'autre part, nous avions ainsi également la possibilité d'observer l'influence de la position verticale de la cible sur les performances des participants : en effet, lors de l'observation initiale, les sujets prennent davantage de temps pour observer les photographies, et sont donc plus attentifs ; l'influence de la position verticale, si elle existe, devrait donc être plus forte à ce stade de la recherche visuelle.

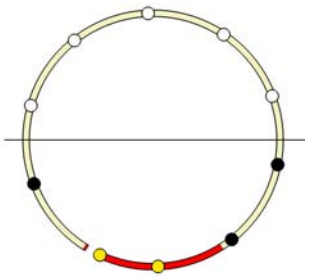


FIG. 6.7 – Position horizontale des cibles sur le cylindre - l'arc en rouge indique la partie visible initialement

6.2.5 Conditions initiales

Les conditions initiales sont les mêmes que pour l'étude préliminaire : la colonne sans photographies est affichée à gauche de l'écran ; l'une partie d'une colonne de photographies est visible à gauche de la fente. La distance à la paroi du cylindre permet la visualisation de 154 images simultanément, soit 18% de la collection. Cette distance a été fixée empiriquement ; c'est un compromis entre la courbure du cylindre, la taille et le nombre de photographies visibles simultanément, paramètres qui influencent, notamment, l'utilisation du zoom et la perception de la perspective, donc l'impresion d'immersion pour la vue interne. Les participants ont bien sûr toute latitude pour modifier cette distance (virtuelle) en cours d'expérimentation, en particulier au début de l'utilisation de chaque vue, pour l'ajuster à leurs capacités visuelles.

6.2.6 Choix des participants

Au total, 20 personnes, 10 hommes et 10 femmes, ont participé à l'expérimentation. Tous les participants étaient des utilisateurs expérimentés d'ordinateurs, avec un niveau d'étude BAC+2 minimum. Le nombre de participants étant relativement restreint, nous avons choisi de borner la plage d'âge (20 à 34 ans) pour limiter l'hétérogénéité du groupe en termes de performances et de préférences. La moyenne d'âge est de 25,2 ans (écart-type : 3,14). Aucun participant n'avait de problème visuel particulier d'après le test Bioptor (*Stereo Optical Company, Inc*).

Les données extraites du questionnaire initial sont synthétisées dans le tableau 6.1.

G1 : VI puis VE				**G2 : VE puis VI**			
n°	**Age**	**Sexe**	**Formation**	**n°**	**Age**	**Sexe**	**Formation**
1	26	F	M1 Histoire	**11**	30	H	Thèse Info.
2	23	H	M1 Sc. Cog.	**12**	22	H	Thèse Info.
3	25	H	Thèse Info.	**13**	21	H	M1 Génétique
4	25	H	Thèse Info.	**14**	27	H	Thèse Info.
5	28	H	Thèse Info.	**15**	24	H	Thèse Info.
6	26	H	Thèse Info.	**16**	24	F	Thèse Info.
7	24	F	Thèse Info.	**17**	24	F	M1 Info.
8	20	F	L2 Info.	**18**	27	F	M1 Info.
9	25	F	DEA Info.	**19**	25	F	M2 Info.
10	23	F	Thèse Info.	**20**	34	F	Thèse Info.

Tab. 6.1 – Identification des participants

6.2.7 Déroulement de l'expérimentation

Les participants ont réalisé les deux types de tâches de recherche visuelle présentées précédemment (T1 et T2). Dix recherches ont été effectuées pour chaque type d'activité, et pour les deux vues cylindriques (VI et VE). Les 20 recherches de type T1 ont été réalisées avant celles

de type T2. A ces 20 recherches, s'ajoutent 4 recherches préliminaires d'entraînement (2 par vue dans T1), afin que chaque sujet puisse se familiariser avec les interfaces et le travail à réaliser. Ajoutons qu'une phase de prise en main, au début de chaque vue dans T1, permet au sujet de s'approprier librement toutes les fonctionnalités d'interaction, sans contrainte de temps.

Le déroulement était le suivant :
- Questionnaire initial d'identification ;
- Test de vision ;
- Présentation générale de l'expérience ;
- Présentation de la structure des descriptions textuelles ;
- Tâche 1, pour chaque vue :
  - présentation de l'interface et des modalités d'interaction ;
  - entraînement (2 recherches) ;
  - réalisation des 10 recherches (à partir des descriptions textuelles) ;
- Tâche 2, pour chaque vue :
  - réalisation des 10 recherches (cibles affichées à l'écran pendant 3 secondes) ;
- Questionnaire d'évaluation des deux interfaces ;
- Entretien post-expérience.

Le test visuel avait pour rôle principal d'éliminer les participants qui avaient des problèmes visuels rédhibitoires pour notre expérimentation (comme le daltonisme). L'enchaînement des recherches et des tâches était automatisé, ce qui nous a permis d'homogénéiser les conditions de passation (voir paragraphe 6.3.3). L'expérimentateur était présent pour donner les explications nécessaires, et restait avec le sujet durant les recherches d'entraînement pour répondre à ses questions éventuelles. Il laissait ensuite le sujet seul, mais celui-ci pouvait le contacter à tout moment par téléphone via une ligne dédiée, pour signaler un problème technique, ou faire lever une ambiguïté éventuelle dans une description. Enfin, les instructions et les consignes étaient données sous forme papier au sujet, qui pouvait les consulter à tout moment. Ces documents sont également joints à l'annexe B. Les questionnaires initiaux et finaux sont identiques à ceux de l'étude préliminaire, et sont consultables dans l'annexe A.

6.3 Méthodes d'analyse

6.3.1 Données recueillies

Les données proviennent des mêmes sources que dans l'étude préliminaire, à savoir de trois origines différentes :
- les questionnaires initial et final : le questionnaire initial permet de disposer du profil général des participants (âge, sexe, cursus, degré de familiarité avec les ordinateurs et les logiciels courants). Le questionnaire final permet de recueillir les préférences et les jugements subjectifs ;
- l'entretien de débriefing : s'appuyant sur un schéma pré-défini, il permet aux participants de donner leur avis plus librement et, à l'expérimentateur, de se focaliser sur certains points précis, comme les stratégies de parcours visuel, par exemple en leur demandant

d'approfondir, compléter ou clarifier leurs réponses aux questions ouvertes du questionnaire final ;
- la trace des interactions : toutes les interactions avec la représentation sont sauvegardées. Elles permettent, d'une part, d'extraire les données permettant l'étude quantitative des résultats des sujets et, d'autre part, d'être en mesure de rejouer toutes les interactions des participants, en vue d'une analyse qualitative à faible granularité.

6.3.2 Variables et mesures

Les mesures suivantes ont été prises en compte, et ont servi de base à l'étude des performances des participants ; elles ont été calculées pour chaque sujet :
- temps de recherche **T** (en secondes) : temps cumulé de recherche des 10 cibles pour une vue (VI ou VE) et une tâche (T1 ou T2). Pour T1, le temps de recherche d'une cible débute lorsque le participant, après avoir lu la description, active l'affichage initial de la collection, et se termine soit après la sélection d'une image et la validation de cette sélection, soit lorsque la limite de temps est atteinte (180 secondes) ;
- nombre de succès **S** : nombre de cibles trouvées pour une vue et une tâche. Par conséquent, la valeur maximum est 10 ;
- vitesse de rotation **V** (en colonnes par seconde) : pour les 10 cibles recherchées au cours de T1 (ou T2) en utilisant VI (ou VE), la vitesse de rotation moyenne est le nombre de colonnes parcourues pour trouver les 10 cibles divisé par la somme des temps de recherche de ces cibles. Le temps de référence considéré exclut les intervalles de temps où le cylindre est immobile ;
- longueur du parcours **L** (colonnes) : cette valeur représente le cumul du nombre de colonnes parcourues pendant la recherche des 10 cibles relatives à une vue (VI ou VE) et à une tâche (T1 ou T2).

Ces données sont recueillies pour T1 et T2. Elles permettent de réaliser des comparaisons entre les vues, et également entre les tâches proposées. La position de la cible dans la collection est également prise en compte pour étudier son impact sur la mémorisation dans la visualisation (T2) et la capacité des participants à retrouver une cible en fonction de sa position. Pour T2, nous avons retenu comme indicateurs de la qualité de la mémorisation de la position des cibles, la durée de l'examen initial de la collection (il n'y a que 2 cibles par vue qui soient visibles initialement) et le choix du sens de rotation, particulièrement pour les cibles ayant des positions horizontales proches du point de vue initial.

6.3.3 Validité interne

Nous allons, dans cette partie, nous intéresser à la validité interne de notre expérimentation. Autrement dit, nous allons expliciter les différentes techniques mises en œuvre pour minimiser l'influence de certaines variables qu'il nous est difficile de contrôler parfaitement, comme par exemple le milieu social (niveau de pratique des participants, intelligence) ou encore leur habileté visuo-spatiale. Les variables parasites provenant du milieu physique (éclairage, bruit, température) sont quant à elles maîtrisées, l'expérimentation se tenant dans une salle unique, climatisée

et sans fenêtre.

Appariement

Etant donné que nous ne disposons pas d'un nombre très élevé de sujets, il nous est impossible d'affecter un groupe à l'évaluation de la vue interne, et un groupe à l'évaluation de la vue externe. Dans le but d'obtenir des résultats comparables entre les deux vues, chaque sujet évalue successivement l'une et l'autre vue (VI et VE), pour chaque tâche (T1 et T2). De cette façon, on contrôle les facteurs parasites, car ils sont constants tout au long de l'expérimentation.

Contre-balancement

L'appariement, s'il présente des avantages quant à la compacité de l'échantillon et au contrôle de certains facteurs, induit l'inconvénient d'introduire des effets d'apprentissage non contrôlés. C'est-à-dire qu'au fur et à mesure de l'expérimentation, chaque participant acquiert un certain niveau de pratique, et peut ainsi accomplir les tâches plus rapidement, par exemple en se construisant progressivement une stratégie de recherche performante. Cela implique que les résultats peuvent ne pas être équivalents entre le début et la fin de l'expérience (amélioration ou dégradation), et plus particulièrement d'une vue à l'autre. C'est ainsi que l'ordre de passation peut jouer un rôle non négligeable sur les performances, ainsi que sur les préférences (la motivation, comme la fatigue ne sont pas constantes).

C'est pourquoi nous avons défini deux groupes de sujets, **G1** et **G2**, pour contre-balancer l'ordre de passation. Ainsi, les sujets du groupe G1 (1 à 10) commencent par la vue interne dans chaque tâche, tandis que les sujets du groupe G2 (11 à 20) commencent par la vue externe. D'une tâche à l'autre, pour un même groupe, l'ordre de passation n'est pas modifié ; ainsi, la durée pendant laquelle les participants doivent éventuellement conserver en mémoire les positions des cibles est sensiblement la même pour les deux groupes. Chaque groupe comprend autant de femmes que de hommes (*i.e.*, 5).

Affectation aléatoire

Les participants ont été affectés de manière aléatoire à l'un ou l'autre groupe, de façon à ne pas favoriser involontairement VI ou VE. L'affectation des cibles à chaque vue, ainsi que leur placement sur le cylindre, ont également été réalisés aléatoirement.

Automatisation de la collecte des données

La présence de l'expérimentateur durant la passation est susceptible d'introduire un biais en gênant le participant. Le participant peut chercher à se rapprocher de ce qu'il croit être les attentes de l'expérimentateur, ou peut être perturbé par sa présence. L'expérimentateur peut également influencer le participant par son attitude. En dehors des phases d'explication et de prise en main, le participant est donc laissé seul, avec néanmoins un lien téléphonique avec

l'expérimentateur. La collecte des données est entièrement automatisée, ce qui facilite la mise en forme de celles-ci en vue de leur exploitation.

Pour ne pas introduire de variabilité dans les indications données aux sujets, les explications et instructions ont été mises sous forme écrite et imprimées. Ainsi, l'expérimentateur peut se borner à les commenter oralement, de façon semblable d'un sujet à l'autre. Ces documents sont par la suite laissés aux participants, qui peuvent s'y référer à tout moment durant l'expérimentation.

6.3.4 Validité externe

La validité externe permet de généraliser les observations particulières à un ensemble plus vaste d'observables. Il faut donc évaluer le degré de généralisation de notre étude. Concernant la validité échantillonnale, l'objectif est de trouver un bon compromis entre le nombre de sujets et la durée de l'étude expérimentale. Lors de l'étude préliminaire, nous disposions de 8 sujets, et les tâches comprenaient uniquement 6 recherches par vue. Pour cette étude, nous avons augmenté le nombre de cibles à chercher en le portant à 10, ce qui conduit à un temps d'expérimentation global maximum d'environ 2h30. Cette augmentation permet également de limiter l'impact de l'hétérogénéité des couples photographie/description. Le choix d'un effectif de 20 sujets est tout d'abord motivé par l'introduction du contre-balancement, et donc par la nécessité de disposer de deux groupes d'effectifs suffisants. Les sujets forment un groupe homogène pour les tâches considérées, de par leur âge, leur niveau et leur domaine d'étude : l'intervalle d'âge est restreint, et tous les sujets sauf un (le sujet 1) sont des informaticiens, présentant donc un niveau de pratique élevé de l'outil informatique.

Comme cela a été mentionné auparavant, nous avons défini un protocole expérimental avec un degré de réalisme élevé, autant par les tâches de recherche proposées que par le matériel visuel utilisé. Nous ne sommes pas dans le cas d'expériences de laboratoire artificielles. Cela assure la validité écologique de nos résultats. Au delà du protocole expérimental et du matériel visuel, le cadre dans lequel nous avons inscrit notre étude, celui des documents visuels, offre un large domaine d'application aux résultats que nous présentons dans les paragraphes suivants. En effet, leur portée ne se limite pas au domaine de la photographie numérique, mais peut s'étendre à tous les documents visuels et même dans certains cas aux documents textuels (voir paragraphe 2.1.1).

6.3.5 Tests statistiques

Nous avons réalisé différents tests statistiques, dont les résultats seront présentés dans la suite du chapitre. Ces tests vont maintenant être présentés, et les motivations de leur choix expliquées.

Comme point de départ, nous avons effectué des tests de normalité (Kolgomorov-Smirnov) sur chaque variable, pour chaque tâche, afin de savoir si nous pouvions recourir à des tests paramétriques ou non. De nombreuses variables parmi celles étudiées ne suivant pas une loi normale, nous avons décidé de nous appuyer uniquement sur des tests non paramétriques, pour une meilleure lisibilité et une meilleure compréhension des résultats. Cette approche est cohérente avec le nombre relativement limité de sujets par groupe (10).

Tout d'abord, étant donné que nous avons contre-balancé l'ordre de passation des sujets (et donc réparti les sujets en deux groupes G1 et G2), le premier aspect à évaluer est l'influence de l'ordre de passation sur les performances des participants, puisque cela va conditionner le choix des tests suivants. Les résultats du test de Mann-Whitney sur les variables retenues pour évaluer les performances des participants au cours de T1 ne permettent pas de conclure à un effet quelconque de l'ordre de passation sur ces performances, donc à la présence de phénomènes d'apprentissage pendant T1 ni, *a fortiori*, pendant T2. Par conséquent, nous ne tenons pas compte de l'ordre de passation dans l'analyse qui suit sur les performances des sujets par vue et par tâche, et considérons l'ensemble des 20 sujets.

Chaque participant travaillant successivement sur chaque vue et chaque tâche, nous avons choisi le test de Wilcoxon (données appariées). C'est à partir de ce test que nous établirons la signification statistique de nos résultats. Nous nous baserons sur le seuil classique d'une probabilité inférieure à 5 %, soit $p < 0{,}05$.

Enfin, pour évaluer l'interaction éventuelle entre type de vue, préférences et performances, c'est-à-dire l'influence des performances sur les préférences, nous ne pouvions utiliser de test non paramétrique. Le recours à une ANOVA s'impose, bien que la distribution de nos mesures de performances s'écarte de la normalité. Cela étant, nous nous appuyons sur la remarque de Winer [Winer, 1971] quant à la robustesse de l'ANOVA aux erreurs de type 1.

Les paragraphes suivants sont consacrés à l'exposé et aux commentaires des différents résultats.

6.4 Influence du type de vue sur les performances

Dans cette section, nous souhaitons mettre en évidence l'influence éventuelle du type de vue sur les performances, c'est-à-dire évaluer si une vue est meilleure que l'autre d'après les résultats quantitatifs obtenus. Ce qui nous permettra de répondre partiellement à la question soulevée par l'hypothèse A. Nous allons présenter les résultats de manière « brute », puis nous les discuterons globalement, afin d'avoir la possibilité de les synthétiser et, éventuellement, de mettre en relation certains d'entre eux lorsque cela s'avère pertinent.

6.4.1 Résultats globaux

Les résultats globaux pour chaque vue sont présentés dans le tableau 6.2. Pour faciliter la comparaison entre les deux vues, nous avons eu recours à une notation compacte, dont voici la signification :

- *Diff. X* : différence entre les valeurs de la mesure X pour les vues externe et interne, soit : valeurVE (X) - valeurVI (X) ;
- *Diff. Err.* : nombre d'erreurs ;
- *Diff. Tps* : temps cumulé de recherche des 10 cibles d'une vue (en %) ;
- *Vitesse* : vitesse de rotation moyenne pour les 10 cibles d'une vue (en %) ;
- *Longueur* : longueur de parcours cumulée pour les 10 cibles d'une vue (en %).

Les pourcentages sont calculés sur la valeur la plus faible du couple (valeurVE (X), valeurVI (X)). Pour mémoire, les temps sont exprimés en secondes et les distances en nombre de colonnes.

n°	**Diff. Err.**		**Diff. Tps (%)**		**Vitesse (%)**	**Longueur (%)**
	T1	**T2**	**T1**	**T2**	**T2**	**T2**
1	1	-	38	13	-	-
2	-1	0	-55	39	18	22
3	4	-1	33	-29	13	-87
4	1	0	-5	32	31	2
5	0	0	19	-41	-1	-44
6	-1	1	-5	7	-15	25
7	3	-1	57	-60	32	-127
8	3	0	9	49	14	46
9	0	0	-19	-9	20	-33
10	0	0	-11	-26	19	-54
11	1	2	-11	147	85	57
12	1	0	74	66	52	12
13	1	0	22	-29	42	-107
14	3	0	83	52	12	53
15	2	0	51	6	33	-34
16	0	0	-23	78	25	46
17	-1	-2	-51	-15	40	-76
18	2	0	104	12	26	-16
19	-1	1	15	20	32	-13
20	4	0	52	11	41	-29

Tab. 6.2 – Performances des participants : comparaison entre les vues VI et VE (les pourcentages sont calculés sur la plus faible valeur du couple VE/VI) pour le nombre d'erreurs, le cumul des temps de recherche, la vitesse moyenne, et le cumul des longueurs de parcours

Le tableau 6.3 présente quant à lui les données brutes, sujet par sujet. Les acronymes utilisés sont les mêmes que pour le tableau 6.2.

Ces tableaux nous permettent de constater la grande variabilité inter- et intra-individuelle. En effet, nous pouvons remarquer que pour un même sujet, les résultats sont très différents d'une vue à l'autre, mais également d'un type de tâche à l'autre. Par exemple, si l'on considère le sujet 7 pour la vue interne, nous pouvons constater qu'il fait nettement moins d'erreurs dans T1 avec VI qu'avec VE (3 erreurs en moins), alors que dans T2 la différence (-1) se réduit sensiblement et change de signe. La variabilité inter-individuelle peut, quant à elle, s'observer sur la comparaison de l'ensemble des sujets.

Les résultats des comparaisons entre les deux vues, pour T1 et T2, sont présentés dans le tableau 6.4. Les valeurs correspondent à des moyennes, calculées pour les 20 sujets. Ainsi, le

n°	T1				T2							
	Erreurs		Temps (s)		Erreurs		Temps (s)		Vitesse (col/s)		Longueur (col)	
	VI	VE	VI	VE	VI	VE	VI	VE	VI	VE	VI	VE
1	0	1	310	485		0	90	121	2,21	1,78	163	360
2	4	3	917	549	1	1	455	639	1,65	1,4	780	970
3	1	5	576	1032	1	0	414	245	2,13	1,88	905	485
4	5	6	698	743	1	1	373	477	1,76	1,34	703	717
5	0	0	659	686	0	0	308	143	1,47	1,49	505	351
6	5	4	1052	778	1	2	447	471	1,36	1,56	630	717
7	2	5	453	881	1	0	519	233	1,86	1,41	991	436
8	0	3	689	752	0	0	236	351	1,81	1,58	346	505
9	6	6	1039	798	2	2	638	510	1,84	1,53	1145	858
10	4	4	799	641	1	1	451	290	1,31	1,1	637	414
11	2	3	665	551	0	2	197	608	1,65	0,89	351	581
12	2	3	431	944	0	0	159	278	1,17	0,77	235	264
13	2	3	513	589	0	0	173	65	3,45	2,43	841	407
14	0	3	130	890	0	0	157	239	1,83	1,64	308	472
15	3	5	656	1242	1	1	453	438	1,58	1,19	731	547
16	2	2	475	316	0	0	110	219	2,09	1,67	338	495
17	3	2	1079	589	3	1	714	550	0,91	0,65	645	369
18	1	3	335	835	1	1	365	392	1,35	1,07	540	465
19	3	2	599	632	1	2	348	415	2,02	1,53	782	720
20	0	4	452	796	0	0	531	549	2,28	1,62	1500	1090

TAB. 6.3 – Performances brutes des participants : comparaison entre les vues VI et VE pour le nombre d'erreurs, le cumul des temps de recherche (secondes), la vitesse moyenne de défilement (colonnes/seconde) et le cumul des longueurs de parcours (colonnes)

temps **T** (secondes) correspond au cumul des temps de recherche des 10 cibles, la vitesse **V** correspond à la vitesse moyenne (colonnes par seconde) de défilement pour les 10 cibles, et la longueur de parcours **L** correspond au cumul du nombre de colonnes parcourues pour les 10 cibles. La variable **S** correspond au nombre de cibles trouvées sur les 10.

6.4.2 Discussion

Les deux tâches, T1 et T2, impliquent deux activités visuelles différentes, exploration *versus* navigation. Les performances des participants lors de la réalisation de chacune d'elles ont donc été analysées séparément. En particulier, deux mesures n'ont été prises en compte que pour T2 : la longueur du parcours et la vitesse moyenne de rotation. Deux raisons justifient ce choix. D'une part, lorsqu'ils réalisent T1, les sujets ignorent la position des cibles : la détermination du sens de rotation n'est donc pas discriminant ; alors que dans T2, la position de chaque cible est

		T1			T2			
		T (s)	**S**	**L (col.)**	**T (s)**	**S**	**V (col./s)**	**L (col.)**
Moyenne (E.T.)	**VI**	821 (206)	7,8 (1,9)	587 (226)	441 (172)	9,3 (0,8)	1,79 (0,53)	654 (336)
	VE	939 (193)	6,7 (1,6)	622 (233)	485 (164)	9,3 (0,8)	1,43 (0,41)	561 (219)
Wilcoxon	**z**	1,680	2,536	0,784	1,195	0,000	3,771	1,344
	p	0,093	0,011	0,433	0,232	1,000	<0,0005	0,179

TAB. 6.4 – Performances des participants : moyennes, écarts-types et tests statistiques pour le temps **T** (cumul pour les 10 cibles d'une vue, en secondes), le nombre de succès **S** (pour les 10 cibles d'une vue), la vitesse de parcours **V** (moyenne calculée pour 10 cibles d'une vue, en colonnes/seconde) et la longueur de parcours **L** (cumul pour les 10 cibles d'une vue, en colonnes)

censée être connue. D'autre part, nous leur demandions d'effectuer les tâches de type T2 le plus rapidement possible, alors que cette contrainte ne leur était pas imposée dans T1.

Au vu des résultats concernant le temps de recherche, T, et le nombre de succès, S, dans T1, VI paraît mieux adaptée que VE à l'exploration visuelle de grandes collections de photographies : les erreurs y sont significativement plus faibles ($p = 0{,}011$) et les temps de recherches sont plus courts (on observe une tendance : $p = 0{,}093$). Les résultats des autres comparaisons ne sont pas significatifs. Les analyses qualitatives indiquent que les participants n'ont pas trouvé 27 images sur les 200 avec VI, contre 43 sur 200 avec VE (temps limite dépassé). Pour 12 participants, on observe de grandes différences dans les temps de recherche (supérieures à 20 %) : 9 d'entre eux ont terminé les recherches de type T1 plus rapidement dans VI.

A la différence des résultats obtenus pour les tâches de type T1, les performances des participants sont similaires entre les 2 vues pour les tâches de type T2. Les différences entre les temps de recherche ne sont pas significatives, tandis que le nombre d'erreurs est trop faible pour être pris en considération. Les participants ont jugé T2 plus facile que T1 dans les questionnaires où ils avaient à évaluer la difficulté de chaque tâche à l'aide d'une échelle allant de 1 (très facile) à 6 (extrêmement difficile). La moyenne est de 3,35 pour T1, et de 1,76 pour T2 ($z = 3{,}72$; $p < 0{,}0005$). La vitesse de rotation pendant T2 est très significativement plus élevée avec VI ($p < 0{,}0005$) alors que la longueur de parcours (nombre de colonnes parcourues) est plus courte avec VE (faible tendance, p=0,179). Parmi les 20 participants, 9 ont des parcours plus courts pendant T2 avec VE contre 5 avec VI (écarts supérieurs à 20 %). Les effets de la vitesse de rotation et de la longueur de parcours ont pu se neutraliser, ce qui peut expliquer l'absence de différence entre les temps de recherche obtenus avec VI et VE durant T2.

La vue interne paraît donc offrir un meilleur support à l'exploration visuelle de collections non structurées de photographies durant T1, alors que les performances sont similaires lors de la recherche de photographies familières visuellement pendant T2. Les différences de distorsions entre VI et VE provoquées par la perspective, ainsi que les activités visuelles différentes entre T1 et T2, peuvent expliquer ces résultats. La vue interne offre une meilleure vision globale des photographies, alors que la vue externe favorise une observation détaillée des miniatures. En

effet, dans la vue interne, concave, la taille des images est plus importante sur les bords gauche et droit de l'écran, là où les déformations dues aux règles de la perspective sont les plus marquées, contrairement à la vue externe qui réduit la taille des images sur les bords latéraux de l'écran. En outre, la déformation des photographies n'est pas identique entre les deux vues. Ainsi, les colonnes d'images non déformées ou affectées de déformations légères sont bien plus nombreuses dans la vue interne que dans la vue externe, dans laquelle les images subissent des déformations et des réductions de taille qui croissent rapidement dès que l'on s'éloigne de la colonne centrale.

Ces caractéristiques géométriques font que la vue interne permet une vision plus globale que la vue externe. Or T1 nécessite une vision plus large que T2, puisque la recherche visuelle lors de T1 fait intervenir, outre les propriétés visuelles des cibles (comme pour T2), leur propriétés « sémantiques » (*i.e.*, thème, composition, cadrage). Les participants ont confirmé cette hypothèse, en indiquant avoir navigué dans les collections à la recherche de photographies avec des caractéristiques visuelles similaires et puis vérifié, pour chaque photographie candidate, la correspondance avec les détails donnés dans la description. Ces observations rendent compte de façon plausible de la plus grande efficacité des participants lors de l'utilisation de la vue interne pour T1. La plus grande vitesse de rotation dans T2 confirme également cette interprétation.

Concernant T2, la reconnaissance de photographies familières demande un examen plus détaillé des miniatures, ce qui devrait favoriser la vue externe, la colonne centrale offrant les photographies non déformées de plus grande taille. Les critères de performance, tels la précision et le temps de recherche, devraient donc valider l'hypothèse A. Pourtant, les distorsions ne semblent pas avoir joué un rôle déterminant, l'influence de ces dernières ayant sans doute été amoindrie par d'autres facteurs, comme le montre la grande diversité des performances inter-individuelles (voir les écarts-types, proportionnellement supérieurs à ceux observés pour T1, dans le tableau 6.4).

6.5 Préférences des sujets

Les préférences des participants ont été établies d'après les questionnaires et les entretiens de débriefing. Dans les questionnaires, l'évaluation a porté sur les points suivants : simplicité d'utilisation, confort, efficacité, précision du repérage. Les participants jouissant d'une plus grande liberté d'expression lors des débriefings, les réponses aux questions ouvertes des questionnaires ont été confrontées aux impressions exprimées oralement. Le tableau 6.5 synthétise les préférences pour la vue interne (VI) ou externe (VE), issues des questionnaires (Q) et des entretiens de débriefings (D). La marque « H » indique une hésitation entre les deux vues, tandis que « < » indique une préférence faible.

Tous les participants, à l'exclusion d'une personne, ont jugé positivement l'apport de la troisième dimension. Les vues 3D ont été préférées aux affichages 2D que les participants connaissaient par ailleurs (explorateur *Windows* ou logiciels de gestion de photographies, par exemple *Google Picasa*). En revanche, leurs impressions concernant les deux vues 3D proposées sont variées. Parmi les 20 participants, 14 ont exprimé une préférence pour l'une ou l'autre vue : 10 participants (dont 5 femmes) préfèrent la vue interne (groupe **GI**), et 4 participants (4 femmes) préfèrent la vue externe (groupe **GE**). Parmi les 6 autres participants (groupe **GH**), un parti-

VI puis VE			**VE puis VI**		
n°	**Q**	**D**	**n°**	**Q**	**D**
1	<I		**11**	I	
2	<E	H	**12**	H	
3	H	I	**13**	H	
4	H	E	**14**	<I	I
5	<I	I	**15**	<I	I
6	<I		**16**	E	
7	E		**17**	<I	
8	I		**18**	<E	H
9	E		**19**	<I	
10	E		**20**	I	

TAB. 6.5 – Préférences des participants en faveur de VI ou VE exprimées dans le questionnaire final **Q** et l'entretien de débriefing **D**. **E** (resp. **I**) indique une préférence pour la vue externe (resp. interne) ; < marque une préférence faible, et **H** indique une hésitation entre VI et VE

cipant apprécie les 2 vues de manière égale, et un n'en aime aucune. Les 4 participants restants ont exprimé des préférences différentes entre le questionnaire final et le débriefing. Si le groupe GI comprend autant de femmes que d'hommes, le groupe GE inclut uniquement des femmes. Mais l'influence du genre sur les préférences ne peut être déduit de cette distribution, en raison des effectifs trop faibles de ces groupes.

Les caractéristiques géométriques des vues ont été le plus souvent mises en avant pour expliciter les préférences lors des débriefings. D'un point de vue global, on remarque que les sujets appartenant au groupe GI évaluent positivement la vue interne, tandis que ceux du groupe GE évaluent positivement la vue externe. Les commentaires positifs des sujets du groupe GH se partagent entre les deux vues. D'un côté, la plus grande taille des photographies sur les colonnes centrales de VE a été appréciée par 6 participants (4 du groupe GE, 2 du groupe GI), car les images des colonnes centrales sont plus visibles, ce qui évite de rapprocher le cylindre et permet de focaliser le regard sur cette partie de l'écran.

> *Je suis plus à l'aise dans cette vue, car les images sont plus grandes.* `Sujet 7`
>
> *Dans la vue externe, les images sont plus visibles.* `Sujet 2`

D'un autre côté, 5 participants (4 GI, 1GH) ont jugé positivement la meilleure vue globale offerte par la vue interne, qui se caractérise par un grand nombre de photographies sans distorsion (ou avec de faibles déformations) au centre de l'écran, là où les colonnes sont les moins hautes, favorisant un balayage vertical rapide.

> *La vue externe n'est pas exploitable en raison de la déformation des images sur le côté du cylindre* `Sujet 11`
>
> *Je pense avoir mieux réussi dans cette vue car on visualise plus d'images.* `Sujet 13`

La plus grande taille des images sur les bords latéraux dans la vue interne à été également considérée positivement par 7 participants (6 GI, 1 GH).

La vue interne est plus intéressante car la taille des images entrantes est agréable. `Sujet 13`

Le caractère immersif de la vue interne ne semble pas avoir exercé une influence sensible sur les jugements subjectifs. Seuls 6 participants l'ont évoqué durant les débriefings, dont 5 positivement (4 GI ; 1 GH).

Je me sens plus à l'aise à l'intérieur du cylindre. `Sujets 5, 19 et 20`

Seul un sujet du groupe GE a mentionné qu'il n'était pas à l'aise dans la vue interne.

On a l'impression d'être dans le cylindre dans VI. Ce qui s'est vérifié car je me suis senti mal à l'aise dans cette vue. `Sujet 9`

Il y a donc un lien fort entre les préférences des participants et leurs commentaires relatifs à la géométrie des vues. Dans les questionnaires et les débriefings, leurs préférences sont justifiées principalement par la taille des photographies, plus importante au centre ou sur les côtés, et par le nombre, plus ou moins important, d'images affectées par les déformations induites par la géométrie cylindrique associée à l'angle de vue (intérieur ou extérieur). Enfin, dans la vue externe, si les photographies de grande taille ne sont pas déformées (contrairement à la vue interne), les images présentes sur les colonnes latérales sont très déformées et de taille très réduite, donc peu exploitables. Sur ce point, la vue interne garantit une déformation moindre et plus progressive. Les effets des métaphores d'interaction sont marginaux, malgré le choix de la projection murale.

6.6 Influence des performances sur les préférences

Comme le montre le tableau 6.2, les performances individuelles varient d'une vue à l'autre. Ces variations ont pu influencer les préférences des participants. L'examen des performances des sujets ayant exprimé clairement leur préférence pour l'une des deux vues conduit aux observations suivantes.

6.6.1 Observations

Pour T1, on observe des variations de performances importantes entre les deux vues pour 6 sujets. Quatre d'entre eux ont des performances en accord avec leur préférence (temps de recherche, nombre d'erreurs) : sujets `8`, `14`, `15` et `20` du groupe GI (VI préférée). Les deux autres ont des performances en désaccord avec leur préférence : sujets `7` et `17` des groupes GE (VE préférée) et GI, respectivement. Durant l'entretien de débriefing, ces deux sujets ont exprimé leur préférence pour une vue, tout en précisant qu'ils avaient été « meilleurs » dans l'autre vue. Concernant T2, 8 participants ont des performances qui diffèrent sensiblement d'une vue à l'autre. Les performances de 5 d'entre eux sont en accord avec leurs préférences (temps de recherche et longueur du parcours) : `S8`, `S11`, `S14` (groupe GI) et `S7`, `S10` (groupe GE). Les performances des 3 autres : `S5`, `S17` (groupe GI) et `S16` (groupe GE) sont en désaccord avec leurs préférences.

6.6.2 Analyse statistique

Nous avons effectué une ANOVA mixte à deux facteurs (un facteur « préférence », et un facteur répété « type de vue ») sur les mesures de performance utilisées pour T1 et T2. Cela nous a permis d'évaluer l'influence du facteur « type du vue », celle du facteur « préférence » et enfin l'interaction éventuelle entre ces deux facteurs. Les résultats de cette analyse ne mettent en évidence aucune interaction significative entre le type de vue, les préférences et les performances. Ils confirment ceux de l'étude qualitative, à savoir l'absence d'influence notable des performances sur les préférences.

6.7 Stratégies d'exploration visuelle

Les stratégies d'exploration visuelle ont été déterminées à partir des débriefings et des traces des interactions des participants. Nous avons confronté ces deux sources de données pour T1, la trace des interactions permettant l'analyse de la position à l'écran des images agrandies (par zoom) au cours de la recherche de la cible. Cette vérification n'a pu être effectuée pour T2, car la recherche des cibles familières pouvait s'effectuer facilement sans faire appel au zoom. Nous nous sommes principalement focalisés sur l'exploration visuelle pendant la rotation du cylindre, étant donné que les participants font un examen rapide de la vue initiale au début de chaque recherche, et arrêtent parfois la rotation pour observer plus en détail une photographie qu'ils pensent être la cible.

6.7.1 Détermination des stratégies

Nous avons réussi à identifier les stratégies d'exploration adoptées par 16 des 20 participants, cette identification ayant échoué partiellement et totalement pour les 4 autres (voir tableaux 6.6 et 6.7). Durant le parcours de la collection, les participants concentrent leur attention soit sur les colonnes centrales (**C**), soit sur les colonnes entrantes (**E**). Les tableaux précédents mettent également en relation l'acuité visuelle des participants, normale (**N**) ou faible (**F**). 6 participants mettent en œuvre des stratégies flexibles, c'est-à-dire changent de stratégie d'une vue à l'autre. 4 d'entre eux ont tiré parti des propriétés géométriques des vues, en particulier de la taille des images latérales dans la vue interne, et centrales dans la vue externe, pour améliorer leur confort visuel : ils ont utilisé la stratégie E pour VI et C pour VE. Les 2 participants restants ont réalisé l'inverse, peut-être pour réduire l'amplitude des mouvements oculaires verticaux pour explorer les colonnes. Enfin, les 10 autres participants ont conservé la même stratégie d'une vue à l'autre : 8 observaient les colonnes entrantes (6 GI, 1 GE, 1GH) et 2 observaient les colonnes centrales (2 GI).

6.7.2 Caractéristiques statiques et dynamiques

Les commentaires des participants pour justifier leurs préférences pour l'une ou l'autre vue apportent un éclairage supplémentaire sur les facteurs pouvant influencer les stratégies d'exploration visuelle. Des facteurs visuels autres que les distorsions peuvent avoir exercé une influence

Participant		**1**	**5**	**6**	**8**	**11**	**14**	**15**	**17**	**19**	**20**
Acuité visuelle		N	N	F	N	N	N	N	F	N	N
Stratégie	**VI**	E	E	E	E	E	C	C	E	E	E
	VE									C	?

TAB. 6.6 – Stratégies d'exploration visuelle des sujets pour le groupe GI. **N** (resp. **F**) indique une bonne (resp. faible) acuité visuelle ; **E** (resp. **C**) indique une stratégie « colonnes entrantes » (resp. « colonnes centrales ») ; **?** marque l'absence d'identification de la stratégie

Participant		**groupe GE**				**groupe GH**					
		7	**9**	**10**	**16**	**2**	**3**	**4**	**12**	**13**	**18**
Acuité visuelle		N	F	N	F	N	N	F	N	N	N
Stratégie	**VI**	E	?	E	C	C	E	?	?	E	E
	VE	C	C		E	E	C	?	?	C	

TAB. 6.7 – Stratégies d'exploration visuelle des sujets pour les groupes GE et GH (voir légende du tableau 6.6)

non négligeable sur les stratégies. Plusieurs d'entre eux ayant conservé la stratégie « colonnes entrantes » pour les 2 vues ont expliqué que le balayage des colonnes entrantes était plus facile et plus efficace que l'observation des colonnes centrales ; le bord de l'écran les aidait à contrôler plus facilement le balayage vertical des colonnes lors du défilement latéral de la vue ; l'observation des colonnes centrales était rendue plus difficile par l'absence d'éléments de repérage visuel vertical. Une facilité supérieure du balayage vertical, associée à une meilleure visibilité des photographies sur les colonnes entrantes peuvent expliquer pourquoi, d'après les tableaux 6.6 et 6.7, 13 participants sur 17 ont regardé les colonnes entrantes dans VI alors que seulement 7 participants ont observé les colonnes centrales dans VE. Ces caractéristiques peuvent justifier le fait que la majorité des participants qui se sont focalisés sur les colonnes entrantes dans les 2 vues (6 sur 8) ont préféré la vue interne, tandis que les préférences des participants qui ont utilisé les 2 stratégies sont plus variées.

Les distorsions engendrées par l'adoption d'une représentation cylindrique semblent avoir exercé une faible influence sur les stratégies visuelles des participants, en comparaison des caractéristiques dynamiques des deux vues. 2 sujets sur 3 ont conservé la même stratégie d'exploration visuelle dans les 2 vues. Les vues interne et externe implémentent des caractéristiques visuelles statiques différentes, mais des propriétés dynamiques identiques, le défilement latéral des photographies, dont l'influence peut avoir interféré avec celle des distorsions.

6.7.3 Acuité visuelle

L'acuité visuelle des participants aurait pu également exercer une influence sur les stratégies d'exploration visuelle. Comme les affichages faisaient l'objet d'une projection murale, nous avons retenu, dans le test *Bioptor*, l'évaluation de la vision binoculaire de loin qui fournit un score allant

de 1 (très faible) à 6 (très bon). Nous avons estimé qu'un score inférieur ou égal à 4 dénotait un défaut d'acuité visuelle. Selon ce critère, 5 participants, répartis comme suit, avaient une acuité plutôt faible : 2 dans le groupe GI (`S6`, `S17`), 2 dans GE (`S9`, `S16`), et 1 dans GH (`S4`). L'influence de l'acuité visuelle devrait se traduire par l'adoption d'une stratégie qui privilégie l'observation de la partie de l'écran où la taille des photographies est la plus importante (colonnes centrales dans VE et entrantes dans VI), ce qui implique un changement de stratégie entre VE et VI. Les 2 sujets du groupe GI ont conservé la stratégie « colonne entrante » pour les 2 vues, tandis que le sujet `16` a utilisé la stratégie opposée à celle qui devrait normalement être observée. Pour les 2 sujets restants, l'identification de la stratégie a échoué partiellement ou totalement. Ces observations, compte tenu du faible nombre de sujets disposant d'une acuité visuelle faible, ne permettent pas de conclure sur le rôle de l'acuité visuelle dans la détermination des stratégies d'exploration visuelle.

Synthèse

Les caractéristiques visuelles des deux vues du cylindre exercent une grande influence sur les préférences des participants, tandis que les métaphores ne jouent pas le rôle que nous attendions. Ces caractéristiques incluent, outre les distorsions dues à la perspective, le défilement horizontal qui interfère fortement avec le balayage vertical qu'exécutent les participants lors de la rotation du cylindre. Quatre stratégies ont été observées : concentration de l'attention visuelle sur les colonnes entrantes (E) ou centrales (C) pour les deux vues, ou bien sur les colonnes entrantes avec VI et les colonnes centrales avec VE (E-C) ou vice-versa (C-E). Elles illustrent la manière dont les participants tirent avantage des caractéristiques statiques et dynamiques des vues cylindriques pour améliorer les conditions d'observation des photographies et/ou optimiser les parcours oculaires, en vue d'accroître le confort et l'efficacité de la recherche visuelle. Cette diversité dans les stratégies utilisées semble être responsable des différences importantes de performances. En raison de leur importance éventuelle, il est nécessaire de mieux identifier et comprendre les stratégies d'exploration visuelle des utilisateurs, et donc d'approfondir l'analyse des comportements visuels lors de l'utilisation de chacune des deux vues, interne et externe, notamment en enregistrant et en étudiant systématiquement les mouvements oculaires lors de la recherche de cibles familières ou non, par exemple au moyen d'un oculomètre (*eye-tracking*). Une telle étude permettra de modéliser précisément l'influence respective des distorsions dues à la perspective et du défilement latéral sur les stratégies d'exploration visuelle des deux vues lors de la recherche d'une image cible ; ce qui permettra de confirmer ou infirmer l'hypothèse A.

6.8 Qualité du repérage spatial

A l'issue de chaque recherche de type T1, la cible à trouver, ainsi que sa position, sont indiquées au sujet, par l'intermédiaire d'une courte animation. Avant de commencer les recherches de la tâche 2, chaque participant a donc eu l'occasion de mémoriser la position de la cible dans la vue cylindrique dont elle fait partie. Rappelons que les cibles, dans T2, conservent la même position (dans la collection) que dans T1, seul l'ordre de présentation diffère. La géométrie

cylindrique permet de réaliser des défilements continus de l'ensemble de la collection, sans avoir à réaliser des retours-arrière, comme dans les interfaces matricielles classiques. Une conséquence importante de ce premier avantage est la possibilité de choisir le sens de rotation. Ainsi, lorsque l'on cherche un document connu, dont la position dans la visualisation est également connue, on peut donc minimiser la longueur de déplacement, le temps de recherche et la fatigue visuelle en choisissant le sens de rotation qui permet d'atteindre la cible le plus rapidement possible.

Comme pour l'étude préliminaire, nous pouvons utiliser plusieurs sources d'informations pour parvenir à évaluer la qualité du repérage spatial dans chaque vue :

- le sens de rotation choisi : pour les cibles en dehors mais proches de l'affichage initial, le choix du sens de rotation est déterminant pour limiter le temps de recherche ;
- la durée de l'examen initial : seules 2 cibles par vue sont présentes dans l'affichage initial. Pour atteindre les 8 autres cibles, une rotation est nécessaire. Par conséquent, pour chacune de ces 8 cibles, un examen détaillé (donc un temps élevé avant la première rotation) indique une quasi-absence de souvenir de la position de la cible ;
- le succès de la recherche.

Outre les 2 cibles présentes dans l'affichage initial, nous nous appuierons principalement sur les 3 cibles les plus proches de cet affichage, car ce sont les cibles pour lesquelles la différence de longueur de parcours entre les 2 chemins possibles est la plus importante (voir figure 6.7). Si les participants ont mémorisé la position horizontale de ces cibles durant T1, ils choisiront le sens de rotation optimal durant T2. Pour les 5 autres cibles, le choix du sens de rotation était plus difficile à mémoriser, la différence de longueur entre les deux chemins possibles étant moins sensible.

6.8.1 Examen initial

Lorsque le temps entre l'affichage initial de la collection et la première rotation dépasse 5 secondes, nous considérons que le sujet effectue un examen visuel des photographies affichées initialement. Pour 8 des 10 cibles, cet examen n'a pas lieu d'être puisque celles-ci ne figurent pas dans l'affichage initial. Pourtant, cet examen a lieu pour 24 cibles (cumulées sur l'ensemble des sujets) dans la vue interne, et pour 44 cibles dans la vue externe. Dans cette dernière vue, 6 participants (`S3`, `S11`, `S14`, `S15`, `S17` et `S20`) effectuent un examen prolongé de l'affichage initial pour au moins 3 des 8 cibles, contre 4 participants (`S4`, `S6`, `S14` et `S17`) dans la vue interne. De plus, nous observons que 4 sujets ont effectué un examen quasi-systématique de l'affichage initial pour plus de 4 cibles sur 8 dans VE (`S11`, `S14`, `S15`, `S17`) contre 2 dans VI (`S14` et `S17`) ; ces 4 sujets font partie du groupe GI.

6.8.2 Choix du sens de rotation

L'examen du choix du sens de rotation pour l'ensemble des recherches de chaque vue (pendant T2) montre que 14 sujets sur les 20 ont gardé un sens de rotation constant avec la vue interne et la vue externe. Cela peut s'expliquer par le fait que la consigne a été donnée aux sujets de réaliser les recherches de type T2 aussi vite que possible ; il n'a pas été demandé aux sujets de choisir le plus court chemin, car cette démarche aurait amoindri le réalisme de T2. Une vitesse de rotation

élevée peut gommer en partie un mauvais choix du sens de rotation. Les stratégies employées pour la mémorisation des positions des cibles dans les recherches de type T1 peuvent également justifier la conservation d'un sens de rotation constant. En effet, dans T1, les participants n'ont pas de raison de choisir un sens de rotation plutôt qu'un autre, et utilisent donc, pour la grande majorité, un sens de rotation constant. La mémorisation de la position de la cible se fait grâce à des points de repère (certaines photographies caractéristiques) tout au long du cheminement. D'après [Trumbo, 1998], les sciences cognitives distinguent 2 niveaux de mémoire : le premier est un processus primitif et passif tandis que le second est un processus actif et volontaire de conservation et de recherche de l'information dans la mémoire. En parcourant la collection, les participants mémorisent ainsi, sans effort, des photographies sur le chemin qu'ils empruntent pour atteindre la cible. Pour pouvoir retrouver celle-ci dans T2, il leur faut donc suivre le même chemin, afin de repasser par les mêmes points de repère. Cela est confirmé par certains sujets dans les débriefings :

> *Je ne me souviens pas de la position exacte de la cible, mais dès que je commence la navigation dans le cylindre, des éléments me permettent de me souvenir de la position approximative de la cible.* `Sujet 9`

Seuls 6 participants ont changé de sens de rotation au cours des recherches de type T2 avec l'une ou l'autre vue. Les traces des interactions suggèrent que ces 6 participants ont mieux mémorisé la position des cibles dans VI que dans VE : pour les 3 cibles proches de l'affichage initial, les chemins les plus courts ont été choisis en moyenne 2,5 fois dans VI, contre 1,75 fois dans VE. Ces résultats qualitatifs confirment l'hypothèse B. Ils sont différents des résultats de l'étude rapportée dans [Cockburn, 2004], mais il faut tout de même remarquer que les tâches proposées aux sujets dans cette étude étaient artificielles et basées sur des représentations 3D statiques, tandis que notre étude s'appuie sur des tâches réalistes avec des représentations 3D dynamiques.

Notons enfin que, lors des débriefings, les participants indiquent avoir davantage mémorisé la position verticale des cibles que leur position horizontale. Et les souvenirs, peu précis, correspondent à une zone de la représentation comprenant plusieurs lignes :

> *Je me souviens de la position en termes de hauteur : haut, milieu, bas.* `Sujet 14`
>
> *J'essaye de me souvenir de la position. Pour environ 5 cibles sur les 20, j'ai une estimation : haut, bas, gauche, droite.* `Sujet 18`

Si 8 des 20 sujets indiquent avoir mémorisé la position d'environ la moitié ou les 3/4 des cibles, cela ne s'est pas traduit par l'adoption d'une stratégie systématique visant à minimiser le parcours pour atteindre la cible. En plus des explications avancées précédemment, nous pouvons également supposer que la représentation (79 colonnes) n'a pas atteint la taille critique qui imposerait une optimisation des déplacements, ce qui, d'un côté, semble prometteur dans la perspective d'une augmentation éventuelle de la taille de la collection de documents visuels. D'un autre côté, cela nous empêche d'une certaine façon de conclure à la supériorité d'une vue sur l'autre concernant la qualité de la mémorisation et du repérage spatial, même si la vue interne semble présenter un avantage pour les sujets (un peu moins d'un tiers) qui n'ont pas conservé un sens de rotation constant durant les recherches de type T2 pour une même vue. Il est donc nécessaire de réaliser des études complémentaires pour déterminer si les environnements proposant la métaphore de la

locomotion comme mode d'interaction peuvent ou non rendre plus efficace la mémorisation des postions des items, par rapport à des représentations non immersives (3D ou 2D). Cela impose la conception d'une nouvelle méthodologie capable de mesurer de manière objective et fiable l'efficacité de la mémorisation dans le contexte d'activités et de situations réalistes.

6.9 Analyses complémentaires

6.9.1 Difficulté d'identification des paysages par rapport aux objets

Nous allons vérifier si les photographies de type « paysage » sont plus difficiles à trouver que les photographies de type « objet », et ainsi valider une des conclusions de l'étude précédente. Concernant T1, si l'on considère les photographies n'ayant pas été trouvées par au moins 5 sujets sur les 20, on remarque qu'elles représentent presque exclusivement des paysages ; nous les précisons dans l'ordre VI puis VE et par nombre croissant d'échecs :

- deux portes de maison, l'une verte et l'autre rouge (VI, C3) : 6 échecs ;
- une voiture de sport jaune (VI, C5) : 6 échecs ;
- un bouquet de palmiers se détachant sur un ciel rose violacé (VE, C4) : 6 échecs ;
- vue aérienne d'un îlot (VE, C5) : 9 échecs ;
- un vestiaire avec des casiers bleus (VE, C8) : 6 échecs ;
- un aqueduc à deux niveaux (VE, C10) : 11 échecs.

La voiture de sport exceptée, toutes les autres cibles mentionnées ci-dessus sont des paysages. Ce qui confirme la difficulté de la recherche de photographies représentant des paysages, structurellement plus complexes et faisant moins l'objet de stéréotypes, donc pour lesquelles les participants ont eu davantage de difficultés pour se créer une représentation mentale pertinente. A noter que la recherche de certaines cibles de la classe « objets » a entraîné des confusions avec d'autres photographies des collections, comme l'éclipse (VI, C8), la pyramide (VE, C3) ou encore la joueuse de tennis (VE, C7). Ce qui illustre la difficulté des tâches de type T1.

6.9.2 Influence de la position verticale des cibles sur le succès de la recherche

La mémorisation de la position verticale des cibles a pu faciliter leur recherche au cours de T2. Ainsi, un participant (`S14`) indique avoir eu plus de difficulté à retrouver les photographies lorsqu'elles étaient « en bas » de la représentation, et plusieurs participants affirment avoir mieux mémorisé la position verticale des cibles que leur position horizontale. La première remarque peut s'expliquer par le fait que la plupart des logiciels utilisent davantage le haut de l'écran (où sont généralement placés les menus) que le bas. La stratégie d'exploration visuelle des sujets, associée à la géométrie de la vue, a pu également influencer le succès de la recherche des cibles au cours de T2 mais aussi de T1. En effet, si l'utilisateur observe les colonnes centrales de la vue interne, les lignes basses et hautes de la visualisation sont éloignées des bords physiques de l'écran, ce qui facilite le parcours visuel vertical des colonnes centrales et le rend plus fiable, alors que les colonnes centrales de la vue externe occupent toute la hauteur de l'affichage. Toutefois, l'analyse des taux d'échec en fonction de la position verticale de la cible ne met en évidence aucune relation entre ces deux variables, pour T1 comme pour T2.

6.10 Commentaires sur l'interface : suggestions d'améliorations

Nous présentons dans cette section les commentaires des participants (issus des débriefings et des questionnaires finaux), sur la représentation, les fonctionnalités et les dispositifs d'interaction. Ces remarques se sont avérées utiles pour la conception de la dernière étude (voir chapitre suivant), en particulier pour améliorer les dispositifs d'interaction avec les deux vues, de manière identique.

6.10.1 La représentation 3D

L'ensemble des participants à jugé positivement le choix du cylindre comme représentation pour l'affichage de collections de photographies. Voici quelques commentaires des participants :

> *Le cylindre est un choix logique qui permet l'efficacité.* `Sujet 1`
>
> *Le cylindre est un bon choix pour les photographies.* `Sujet 2`
>
> *Le cylindre est une bonne représentation, c'est « fun ».* `Sujet 14`

Seule la fluidité de l'animation durant la rotation a été critiquée : à haute vitesse, le défilement devient haché et cela entraîne une fatigue visuelle. Ces saccades ont été mises en lumière par 14 des 20 sujets. Ce phénomène n'est pas lié au type de déplacement (la rotation), mais est à mettre sur le compte de faiblesses techniques. Il ne se manifeste qu'à vitesse très élevée, et est une conséquence du choix de permettre aux participants de faire défiler très rapidement le cylindre. Ce problème peut-être corrigé par une optimisation du code, mais l'investissement en développement serait disproportionné pour un prototype de recherche.

6.10.2 Dispositif d'interaction

Nous avons fourni aux participants deux dispositifs d'interaction fonctionnellement identiques et visuellement très proches bien qu'adaptés aux caractéristiques de chaque vue. Ils ont été positivement évalués par la moitié des sujets qui le considérent comme *simples* (`S9`, `S10`) ou encore *faciles à maîtriser* (`S5`). Les principales critiques se concentrent sur deux points particuliers :

- dans la vue externe, l'interacteur se trouve « devant » la représentation, et masque les images centrales situées sur les lignes inférieures, malgré nos efforts de le rendre « transparent » pendant la rotation ;
- dans les deux vues, le choix d'accompagner la rotation du cylindre par la rotation des flèches de l'interacteur a été contesté : le mouvement a été perçu comme *perturbant* (`S12`, `S14`) ou son interprétation jugée *difficile* (`S2`, `S16`).

Mis à part ces commentaires sur l'interacteur proposé, certains sujets auraient préféré une interaction différente, comme par exemple l'usage de la molette de la souris (`S2`, `S3`) ou encore une interaction directe par *drag* (`S6`). En outre, le fait de devoir cliquer (clic droit, n'importe où à l'écran) pour arrêter la rotation a été jugé peu naturel par seulement deux participants (`S7`, `S14`).

6.10.3 Dispositif de repérage spatial

Mis à part le problème de la superposition du dispositif d'interaction-repérage avec la vue externe, ce dispositif a été jugé positivement par les sujets :

Le système de repérage est bien. `Sujets 6, 15, 17`

La plupart des participants n'a cependant pas donné d'avis particulier sur ce dispositif, en particulier sur son utilité pour la recherche d'images. Il pourrait donc ne pas être affiché en permanence, mais uniquement à la demande.

6.10.4 Eloignement-rapprochement

Les participants pouvaient moduler la distance entre la caméra virtuelle et le cylindre pour choisir une distance qui leur permette une vision confortable. Cette fonctionnalité n'a pas été utilisée par au moins 5 des 20 sujets. Cinq autres sujets indiquent avoir modifié la distance uniquement pour la première recherche, ce qui correspond à l'usage que nous avions prévu. Le faible taux d'utilisation explique l'absence de commentaire sur cet aspect de l'interaction.

6.10.5 Zoom

Le zoom permet aux utilisateurs d'agrandir une photographie afin d'obtenir un niveau de détail plus élevé. T1 impose un recours au zoom plus important, afin de vérifier la correspondance entre une photographie et la cible décrite textuellement. La seconde tâche rend l'utilisation du zoom moins nécessaire, tout du moins dans le cas d'une bonne mémorisation de la position de la cible. La majorité des sujets a jugé le zoom utile et satisfaisant. 8 sujets affirment ne pas s'en être servi, jugeant les photographies suffisamment visibles. Le zoom est néanmoins une fonctionnalité qu'il faut implémenter, dans l'objectif de satisfaire les besoins du plus grand nombre d'utilisateurs, d'autant plus que sa présence est « invisible », et n'occasionne pas de gêne pour les personnes qui la jugent inutile. La principale critique concerne le nombre de clics nécessaires pour l'activation et la désactivation du zoom d'une image.

Le zoom est ennuyeux à cause du nombre de clics nécessaires. `Sujet 13`

Le zoom est à refaire en raison du clic nécessaire pour dézoomer. `Sujets 3 et 9`

La sélection par clic est intuitive et naturelle pour la grande majorité des participants, même s'il est vrai que cela nécessite attention et précision dans le mouvement. Deux améliorations formulées par les participants sont à prendre en considération. D'une part, la possibilité de moduler le niveau de zoom (`S5`, `S14`, `S18`), toujours dans l'optique d'être au plus proche des capacités de chacun. D'autre part, outre le fait que le nombre de clics est jugé trop important, le sujet `15` aurait jugé préférable de faire un clic droit pour dézoomer, plutôt que de devoir cliquer sur l'image agrandie. Cette amélioration présente deux avantages : le premier, d'être cohérent avec l'action à réaliser pour arrêter un mouvement ; le second, de limiter la précision demandée en autorisant un clic sur n'importe quelle partie de l'affichage.

6.10.6 Discussion

Comparée à l'interface réalisée pour l'étude préliminaire, l'interface de l'étude décrite dans ce chapitre a suscité beaucoup moins de commentaires critiques de la part des sujets. Nous sommes donc parvenus à une sorte de point d'équilibre où seules de petites retouches sont nécessaires. En particulier, faire disparaître le dispositif d'interaction de l'affichage pendant la rotation, pour éviter tout phénomène de masquage. La présence permanente du dispositif de repérage ne s'avérant pas utile, il sera nécessaire de déterminer comment représenter graphiquement les « boutons » de lancement de la rotation différemment, en minimisant la précision demandée. Pour autant, il apparaît que le clic de sélection reste le moyen le plus intuitif. Les fonctionnalités proposées sont quant à elles suffisantes pour réaliser, sur les visualisations que nous proposons, les actions élémentaires nécessaires à la recherche visuelle de photographies dans de grandes collections. Elles ont toutes été jugées positivement. L'intégration des observations des sujets doit permettre de s'approcher d'une interface simple à utiliser, ne détournant pas les utilisateurs de leurs activités de recherche et/ou d'exploration. L'interaction que nous avons mise en œuvre dans le prototype utilisé pour l'étude présentée dans ce chapitre est donc mature et adaptée aux représentations cylindriques. Elle constitue une base satisfaisante pour la conception de l'étude présentée dans le chapitre suivant.

6.11 Conclusions générales

Nous avons présenté une étude empirique visant à déterminer l'utilité, le confort et l'efficacité d'un cylindre virtuel interactif pour représenter de grandes collections non structurées de documents visuels, des photographies au cas particulier. Nous avons implémenté et évalué deux métaphores d'interaction avec cette représentation : « manipulation » d'un objet virtuel 3D, constitué par une vue externe du cylindre, ou interaction avec les éléments d'un espace virtuel immersif, constitué par une vue interne du cylindre. Dans la vue externe, les distorsions dues à la perspective canalisent le regard vers les colonnes centrales, tandis qu'elles permettent une meilleure vue globale des colonnes centrales dans la vue interne. L'étude préliminaire conduite précédemment a permis de mesurer l'intérêt de 8 utilisateurs potentiels pour ces représentations. Les résultats indiquaient une légère supériorité de la vue interne sur la vue externe, en termes de performances et de préférences. En tenant compte des améliorations proposées par les sujets et de leurs commentaires, nous avons corrigé les défauts des vues interactives utilisées lors de l'étude préliminaire et amendé, en particulier, l'interaction avec ces vues. Dans l'objectif de réaliser une analyse statistique des performances, nous avons augmenté, dans le cadre de l'étude présentée dans ce chapitre, l'effectif des participants (20), ainsi que le nombre de cibles à chercher (10) pour chaque vue et chaque tâche (T1 et T2) de recherche visuelle, l'une portant sur des photographies non familières (caractérisées de manière unique par une description textuelle), l'autre portant sur des photographies familières, dont les positions au sein des collections sont connues. Nous résumons dans la suite de cette section les principaux résultats de cette première étude qui fait suite à l'étude préliminaire et à laquelle est consacré le présent chapitre.

6.11.1 Synthèse des résultats

S'appuyant sur les mêmes fonctionnalités, les 2 vues du cylindre vertical possèdent des modalités d'interaction identiques, bien qu'adaptées à chaque métaphore d'interaction. L'étude préliminaire a montré le faible impact de ces métaphores sur les préférences et les performances. Ainsi, nous avons convergé vers un dispositif d'interaction « mixte », reprenant des éléments de chaque métaphore d'interaction pour rendre l'interface plus simple à utiliser, mais aussi plus efficace. Toutes les commandes de déplacement ont été regroupées dans un seul interacteur qui sert également de dispositif de repérage ; l'agrandissement d'une photographie se fait par un simple clic. La sélection par clic reste aujourd'hui la méthode la plus intuitive car la plus couramment utilisée, même si elle n'est pas exempte de défauts, entre autres la précision du geste requise.

Contrairement à nos attentes, les performances des participants diffèrent relativement peu entre les deux vues : cela est principalement dû aux importantes différences inter-individuelles observées. Les résultats obtenus sont en contradiction avec l'hypothèse A, la vue interne permettant d'obtenir des résultats significativement meilleurs durant T1 (temps de recherche, nombre de succès) que la vue externe. Dans la tâche T2, plus simple à réaliser, les performances des participants sont similaires, même si nous remarquons que la vitesse de rotation est plus élevée dans VI, tandis que la longueur de parcours est plus faible dans VE. Les performances globales étant très proches, ces deux variables se sont vraisemblablement neutralisées. La plus grande vitesse de défilement dans VI peut être mise en relation avec les meilleures performances dans T1 pour VI : la vue interne, par les déformations moindres des photographies qu'elle implique du fait de sa concavité, apporte une meilleure vision globale des photographies affichées. Elle est donc plus adaptée à des tâches de navigation qui demandent un examen rapide de l'ensemble des photographies de la collection, ce qui confirme l'hypothèse D.

Tous les participants, sauf un, ont jugé positivement l'apport de la troisième dimension, comparativement aux représentations 2D couramment utilisées (explorateur de fichier *Windows*, par exemple), ce qui confirme un des résultats de l'étude préliminaire. Sur les 20 participants, 14 ont exprimé une préférence marquée pour l'une des deux vues 3D : 10 participants préfèrent VI, alors que 4 participants seulement expriment une préférence pour VE. Les 6 autres participants n'ont pas de préférence particulière, ou ont exprimé des avis divergents entre le questionnaire final et le débriefing. Si les performances n'ont pas influencé les préférences des participants, les caractéristiques géométriques des vues ont joué vraisemblablement un rôle non négligeable ; elles sont en effet souvent mises en avant pour justifier les préférences. Ainsi, VE est préférée pour la taille des photographies dans les colonnes centrales, tandis que VI apporte une meilleure vue globale avec davantage de photographies faiblement déformées, et de plus grande taille sur les « bords ». Ces résultats sont partiellement en accord avec l'hypothèse C. Bien que renforcé par la projection murale, le caractère immersif de VI n'a été évoqué que par 5 sujets, dont 4 positivement et 1 négativement.

Les débriefings et les traces des interactions ont permis de déterminer 4 types de stratégies d'exploration visuelle : l'observation des colonnes centrales (C) ou entrantes (E) pour les deux vues, ou E pour VI et C pour VE ou *vice-versa*. Sur les 16 participants pour lesquels l'identification des stratégies visuelles a été possible, 10 conservent une stratégie constante d'une vue à

l'autre. Quant aux 6 autres, ils utilisent des stratégies flexibles. 4 d'entre eux ont tiré parti des caractéristiques géométriques des vues en adaptant leur stratégie aux distorsions (stratégies E pour VI et C pour VE). Nous pouvons néanmoins remarquer que 10 sujets ont une stratégie E dans VE, ce qui suggère une influence prépondérante des caractéristiques dynamiques des vues sur les stratégies visuelles, par rapport à leurs propriétés statiques spécifiques : le renouvellement des photographies se faisant latéralement, le bord de l'écran peut apporter un guidage naturel, ce qui favorise une stratégie « colonnes entrantes » même si, pour VE, les images latérales sont très déformées et de taille très réduite, alors que pour VI elles sont de taille importante et relativement peu déformées. VI et VE implémentant la même caractéristique dynamique, le défilement latéral, les deux vues tendent à favoriser une même stratégie visuelle (*i.e.*, E plutôt que C) puisque les propriétés dynamiques des vues semblent influencer davantage les stratégies d'exploration visuelle que leurs propriétés statiques. Cette observation peut expliquer pourquoi les performances entre VI et VE ne diffèrent pas davantage, les participants tirant au mieux parti des caractéristiques visuelles dynamiques identiques des deux vues et, à un degré moindre, des caractéristiques statiques différentes de chacune pour améliorer l'efficacité de leur recherche. Ces résultats ne contredisent pas l'hypothèse C, même s'ils imposent une lecture légèrement différente de cette hypothèse, dans la mesure où, dans sa formulation initiale, elle tenait compte exclusivement de l'influence éventuelle des caractéristiques statiques spécifiques de chacune des deux vues sur les stratégies d'exploration visuelle.

Pour évaluer la qualité du repérage spatial, nous nous sommes appuyés sur les 2 cibles visibles initialement et sur les 3 cibles les plus proches de l'affichage initial. Ces cibles peuvent conduire à des temps de recherche très différents, selon le sens de rotation choisi. Pour les 8 cibles nécessitant un déplacement, 4 sujets ont effectué un examen inutile de l'affichage initial dans VE, contre 2 seulement dans VI, ce qui traduit une mauvaise mémorisation de la position des cibles. 14 des 20 sujets gardent un sens de rotation constant durant toutes les tâches de type T2. Ce qui peut s'expliquer par la consigne de réaliser T2 aussi vite que possible, ainsi que par la stratégie adoptée pendant T1 (points de repère le long du chemin menant à la cible). Pour les 6 sujets changeant de sens de rotation au cours de T2, on observe que, pour les 3 cibles proches de l'affichage initial, le sens de rotation correct est choisi en moyenne 2,5 fois avec VI, contre 1,75 fois avec VE. Contrairement à [Cockburn, 2004], et en accord avec [Robertson *et al.*, 1998], l'analyse qualitative des comportements des participants suggère que la vue interne a le potentiel de favoriser une meilleure mémorisation spatiale que la vue externe. Ce qui contribue à confirmer l'hypothèse B. Néanmoins, l'adoption de l'hypothèse B nécessite la validation de ce résultat qualitatif par une étude qui implique davantage de sujets et fasse appel à des mesures plus précises tout en préservant le caractère réaliste des tâches proposées aux participants, ce qui impose la conception et la mise en œuvre d'une nouvelle méthodologie. Enfin, même s'il apparaît difficile de mesurer l'impact du caractère immersif de la vue interne sur le processus de mémorisation, il serait nécessaire de mener des études plus spécifiques sur ce point, en conservant le réalisme de l'étude présentée dans ce chapitre.

6.11.2 Conclusion

La vue interne semble prendre l'avantage sur la vue externe, au sens où elle permet de meilleures performances (dans T1), tout en garantissant une meilleure mémorisation de la position des photographies. De plus, compte tenu de l'influence prépondérante des caractéristiques visuelles dynamiques sur les stratégies d'exploration visuelle, VI semble plus adaptée au défilement latéral, les images entrantes étant agrandies dans cette vue, contrairement à la vue externe où les images entrantes sont de taille réduite. Enfin, rappelons que la moitié des participants donnent leur préférence à la vue interne, quand seulement 4 préfèrent la vue externe. Pourtant, compte tenu du partage des préférences, et des performances relativement proches obtenues avec les deux vues, il ne semble pas opportun d'exclure totalement la vue externe. Afin de répondre au mieux aux besoins et aux attentes des utilisateurs, une solution possible, à court-terme, est de proposer les deux vues conjointement. On se rapproche ainsi des solutions proposées dans [Vernier et Nigay, 1997] qui recommande la conception de représentations multiples. L'adoption de cette solution offrirait une meilleure adaptation à la diversité des tâches, à condition que les représentations interactives proposées vérifient les trois principes énoncés dans [Vernier et Nigay, 1997], à savoir, assurer la facilité du passage d'une représentation à l'autre à tout moment, la continuité visuelle lors du passage d'une visualisation à l'autre, et l'utilisation des mêmes outils de navigation dans les deux vues.

Il est aujourd'hui nécessaire de comparer de façon globale les vues de la représentation cylindrique aux structures matricielles 2D classiques pour déterminer avec précision quels peuvent être les apports de la troisième dimension par rapport aux visualisations 2D. Au vu de l'importance de la relation entre les stratégies d'exploration visuelle et les caractéristiques statiques et dynamiques des vues choisies, il est nécessaire d'observer plus finement le comportement des participants lors de leur utilisation, en ayant recours à des techniques avancées telles que l'oculométrie (*eye-tracking*). C'est l'objet de l'étude présentée dans le chapitre suivant.

Chapitre 7

Seconde étude

Cette étude expérimentale a pour objet la comparaison des vues 3D du cylindre vertical à une vue 2D où les photographies sont disposées sous forme matricielle, ce qui la rend semblable aux interfaces de visualisation de nombreux logiciels de gestion de documents visuels. 20 participants (10 femmes et 10 hommes) ont pris part à cette expérimentation, en réalisant les deux types d'activité visuelle proposés aux participants lors de l'étude préliminaire et de la première étude. L'évaluation porte sur l'efficacité, le confort de la recherche et la facilité du repérage spatial. L'analyse des parcours oculaires, rendue possible par l'usage d'un oculomètre, permet de comprendre les relations entre les caractéristiques géométriques des vues et les stratégies visuelles des participants en évaluant quantitativement et qualitativement le confort visuel et l'efficacité de chaque représentation.

L'étude précédente a permis une comparaison entre la vue interne et la vue externe : si la vue interne est préférée par 10 participants, contre 4 en faveur de la vue externe, les différences, en termes de performances, sont assez limitées. Ces deux vues pouvant être potentiellement proposées à des concepteurs de logiciels, il est donc important de les comparer conjointement à une vue 2D classique afin de déterminer quels peuvent être les apports de la troisième dimension à la recherche visuelle.

La majorité des 20 participants (14) donnent leur préférence à une des vues 3D ou aux deux, quand seulement 6 participants préfèrent la vue 2D. Si les performances n'influent pas sur les préférences, les caractéristiques géométriques ont majoritairement été mentionnées pour expliquer les préférences. Ainsi, la meilleure vue globale dans la vue interne et la taille plus importante des images centrales dans la vue externe ont été appréciées, tandis que le caractère familier et l'absence de déformation justifient les préférences pour la vue 2D. Les stratégies de parcours visuel (attention portée sur les colonnes entrantes ou centrales) semblent influencées davantage par les caractéristiques dynamiques (dues au défilement latéral des vues) plutôt que par leurs propriétés visuelles intrinsèques statiques. Enfin, les vues 3D garantissent un meilleur confort visuel, en offrant un meilleur guidage du regard (principalement la vue interne) et en minimisant les saccades désordonnées observées dans la vue 2D.

Après la présentation de la méthodologie et des hypothèses de travail, nous détaillons les représentations que nous proposons et les modalités d'interaction associées, en justifiant les mo-

difications apportées par rapport à l'étude précédente. Le protocole expérimental, les méthodes d'analyse et les données recueillies sont ensuite décrits. Ce chapitre se termine par la présentation et l'interprétation des résultats obtenus.

7.1 Méthodologie

Si les deux vues 3D ont été accueillies favorablement par les participants de l'étude préliminaire et de l'étude précédente, il nous est impossible d'affirmer leur supériorité sur les représentations 2D actuelles, et en particulier sur la disposition matricielle couramment utilisée dans les logiciels de gestion de collections de photographies, et même plus généralement dans les explorateurs de fichiers de nombreux systèmes d'exploitation. Le défilement horizontal et/ou vertical de ces représentations peut être mis en relation avec la rotation dans la représentation cylindrique. Il est donc possible de réaliser une comparaison 2D–3D en définissant des modalités d'interaction qui soient identiques dans les trois vues, tout en assurant une cohérence avec les métaphores d'interaction associées aux représentations 3D et une fidélité aux interfaces 2D classiques. Ces 3 vues ont été testées par l'ensemble des participants sur les deux tâches de recherche détaillées dans le chapitre 4, de façon à identifier les apports de chaque vue à la spécificité de chaque tâche.

7.1.1 Présentation générale

Nous conservons donc les tâches T1 et T2 présentées précédemment. En voici un bref rappel :

- T1 : recherche d'une photographie non familière dont la position dans la collection n'est pas connue. Une description textuelle contenant des éléments discriminants permet de distinguer sans ambiguïté la cible des autres photographies de la collection ;
- T2 : recherche d'une photographie familière dont la position est connue. La cible à chercher est présentée isolément au centre de l'écran durant 3 secondes.

Dans la première étude, nous avions augmenté le nombre de cibles par rapport à l'étude préliminaire, de manière à renforcer la signification des résultats. Dans cette étude, compte tenu de l'augmentation du nombre de conditions à tester, l'utilisation de 10 cibles par vue n'est pas raisonnable au regard du temps d'expérimentation. Ce dernier était déjà relativement important dans les deux études précédentes, et la contrainte de l'utilisation d'un oculomètre impose de limiter au maximum la fatigue visuelle. Nous avons choisi d'utiliser 6 cibles par vue, soit 18 cibles au total, pour limiter la durée de passation. Les 2 études précédentes ont souligné la difficulté d'uniformiser la difficulté des cibles à trouver, c'est pourquoi nous n'avons pas tenu compte prioritairement de la distinction objets/paysages. Nous avons pour cela sélectionné des triplets de photographies possédant des caractéristiques visuelles et sémantiques semblables. Néanmoins, l'impact de la position relative des cibles par rapport aux autres photographies des collections n'a pas été étudié, car cela aurait demandé un travail dépassant largement la portée des travaux rapportés ici. Les photographies sont placées aléatoirement dans la collection.

7.1.2 Terminologie

En plus des termes définis dans le chapitre précédent, à savoir « collection », « cible », « recherche », « tâche » et « vue », nous utiliserons les acronymes suivants pour désigner les trois vues :

- **VI** : vue interne du cylindre vertical,
- **VE** : vue externe du cylindre vertical,
- **V2D** : vue 2D.

7.1.3 Hypothèses à tester

L'étude précédente a permis d'étudier comparativement les vues interne et externe du cylindre vertical. Nous avons identifié deux stratégies d'exploration visuelle pendant le défilement du cylindre ; dans l'une, l'attention visuelle se porte principalement sur les colonnes entrantes, dans l'autre, sur les colonnes centrales. Les analyses qualitatives menées suggèrent que la vue interne a le potentiel d'offrir un meilleur support pour la mémorisation spatiale. Au delà de la comparaison 2D–3D en termes d'efficacité et de confort visuel, l'étude présentée dans ce chapitre doit permettre de confirmer les résultats de l'étude précédente, bien que le nombre de cibles y soit plus faible. Nos hypothèses de travail sont donc les suivantes :

- hypothèse A : les vues 3D, en raison de leur géométrie, apportent un meilleur confort visuel et permettent de meilleures performances de recherche visuelle que la vue 2D. En outre, la continuité de la représentation cylindrique est de nature à rendre le parcours de la collection plus rapide, durant T2. En effet, les vues 3D offrent deux sens de parcours de la collection à partir de l'affichage initial alors que la vue 2D ne permet qu'un seul sens de parcours ; par conséquent, les vues 3D offrent des possibilités d'optimisation de la recherche d'une cible lorsque sa position dans la visualisation est connue, contrairement à la vue 2D. Elles doivent également procurer un meilleur confort visuel ;
- hypothèse B : les vues 3D, et plus particulièrement la vue externe, permettent un meilleur contrôle des mouvements oculaires lors de l'exploration visuelle par rapport à la vue 2D, car elles canalisent le regard vers les zones de l'affichage où les miniatures ne sont pas déformées et/ou ont la plus grande taille ; en particulier, dans la vue externe, les miniatures agrandies et non déformées qui se situent au centre de l'écran, attirent le regard et le canalisent sur elles ;
- hypothèse C : la vue interne peut améliorer, par rapport à la vue externe, la réussite des tâches faisant appel à la mémoire spatiale (*i.e.*, les tâches de type T2), car la métaphore de la locomotion peut faciliter la mémorisation de la position horizontale des cibles, plus que la manipulation de la représentation, proche de la métaphore de la manipulation directe ;
- hypothèse D : plus que par les performances, les préférences des utilisateurs sont majoritairement influencées par les caractéristiques visuelles intrinsèques des vues, donc par le confort visuel procuré par chaque vue.

Ces hypothèses (principalement A, B et C), si elles sont confirmées, permettront de conclure sur la supériorité des vues 3D sur la vue 2D, reproduisant une disposition 2D matricielle classique.

Dans la suite de cette section, nous allons présenter la vue 2D, ainsi que les modifications apportées aux représentations 3D en vue d'obtenir des modalités d'interaction identiques avec les trois vues.

7.1.4 Matériel visuel : les trois représentations

Les deux vues 3D sont celles utilisées précédemment. Les photographies (869) de chaque collection sont plaquées sur la face interne (VI) ou externe (VE) d'un cylindre vertical ; elles sont disposées suivant 11 lignes et 79 colonnes et 20% d'entre elles environ sont visibles simultanément. Le défilement se fait par rotation de la gauche vers la droite ou inversement, et correspond au défilement horizontal des visualisations 2D classiques. La colonne sans image est conservée, car elle indique de manière simple un parcours intégral de la collection.

La représentation 2D (V2D) satisfait à deux contraintes principales : d'une part, être aussi proche que possible des visualisations 2D classiques et, d'autre part, être équivalente à VI et VE. Elle se distingue essentiellement de ces dernières par sa non-continuité : après un parcours complet de la collection, il faut inverser le sens de parcours. Enfin la disposition de V2D est différente, avec 66 colonnes et 13 lignes : cette modification était nécessaire pour que la taille des miniatures soit égale à la taille moyenne des miniatures dans VE et VI. Par conséquent, le nombre total de photographies dans chaque collection est de 858 en 2D contre 869 en 3D, et celui des miniatures visibles simultanément lors de l'affichage initial par défaut est de 169 (respectivement 176). Ce choix permet en outre de conserver l'aspect visuel des représentations matricielles classiques qui visent à une couverture maximale de l'écran. Les cylindres étant verticaux, la représentation 2D se présente donc comme une « bande » horizontale, afin que le déplacement soit identique pour les vues 3D et la vue 2D.

Comme pour les études précédentes, nous utilisons des photographies issues de bases de données libres de droit sur Internet. Nous conservons la classification en thèmes et la distinction entre objets et paysages, tout en raffinant celle-ci lors du choix des cibles pour homogénéiser leur difficulté, autant que faire se peut. Il est certain que notre parti pris d'étudier des situations naturelles et d'obtenir des données réalistes nous empêche d'atteindre une équivalence stricte entre les trois vues et entre les trois ensembles de cibles choisis pour les tester.

7.1.5 Interaction avec les représentations

Le dispositif d'interaction utilisé pour l'étude préliminaire a été modifié assez largement lors de la conception de la première étude décrite dans le chapitre précédent afin de le rendre plus simple à appréhender et à utiliser. Les dispositifs d'interaction avec VI et VE utilisés lors de la première étude, en devenant similaires, ont permis une comparaison non biaisée de l'utilisation des deux vues, au prix d'un respect moins strict des métaphores d'interaction propres à chaque vue. Dans l'étude présentée ici, il est nécessaire d'offrir aux 3 vues le même dispositif d'interaction.

Nous nous écartons de la méthodologie utilisée dans d'autres études, consistant à comparer des prototypes expérimentaux à des logiciels existants [Bladh *et al.*, 2004 ; Combs et Bederson, 1999]. Même si l'intérêt de ce type de comparaison est évident, il faut néanmoins assumer le

fait que les résultats obtenus dans un tel contexte ne peuvent prétendre à une totale objectivité. D'une part, les fonctionnalités et les propriétés ne sont pas totalement similaires et, d'autre part, l'expérience antérieure des participants joue un rôle prépondérant dans le cas de logiciels fréquemment utilisés par le grand public. Le code source de la plupart de ces logiciels n'étant pas publique, on peut envisager de récrire une partie du code pour avoir le contrôle de l'interface et des fonctionnalités, mais l'investissement requis par une telle entreprise la rend irréaliste au regard du temps demandé pour cette tâche. C'est pourquoi nous avons fait le choix de réaliser la vue 2D interactive destinée à être comparée aux vues 3D plutôt que d'utiliser un logiciel existant de gestion de collections de photographies. Ce choix de conception permet de contrôler entièrement les fonctions d'interaction avec la vue 2D et leur présentation graphique.

La principale exigence est de fournir des moyens d'interaction qui soient proches de ceux existants dans les représentations 2D classiques, sans dénaturer la troisième dimension apportée par les vues 3D. Avant de détailler les modalités d'interaction et les interacteurs retenus, nous allons, comme pour l'étude précédente, rappeler les fonctionnalités d'interaction que chaque représentation devra implémenter.

Fonctionnalités des représentations

Nous conservons les fonctionnalités d'interaction retenues pour l'étude précédente, car elles permettent de réaliser les tâches de recherche proposées, et correspondent aux opérations élémentaires les plus fréquemment utilisées par les utilisateurs de logiciels de gestion de collections de photographies. En voici un rappel, chaque fonctionnalité sera détaillée par la suite lors de la présentation des interacteurs.

- **déplacement** : deux types de déplacements sont proposés. Le premier permet de naviguer au sein de la représentation, puisque la taille de la collection dépasse la partie visible initialement. Compte tenu du choix des cylindres verticaux, ce déplacement est un défilement latéral à distance constante de la paroi du cylindre. Il en est de même pour la vue 2D. Dans les représentations matricielles classiques, il est courant de rencontrer un défilement vertical, parfois associé à un défilement horizontal. Ce dernier se retrouve dans des logiciels courants, comme par exemple *Microsoft Excel*. Le second déplacement est réalisé suivant un plan perpendiculaire à la représentation. Il permet d'adapter finement l'affichage aux capacités visuelles et cognitives de chacun, en faisant varier continûment la distance entre l'utilisateur et la représentation. Ce qui entraîne la modulation du nombre de photographies visibles simultanément, ainsi que celle de leur taille.
- **zoom** : agrandissement d'une photographie particulière de la collection. Cette fonctionnalité est proposée par la plupart des logiciels de gestion de collections de photographies. Lorsque la vue globale ne permet pas d'avoir un niveau de détail suffisant lors de l'observation d'une photographie, il doit être possible d'augmenter sa taille ainsi que sa définition.
- **dispositif de repérage spatial** : bien que cette fonctionnalité ait été peu utilisée par les participants de l'étude précédente, il est important de la conserver, car elle indique le point de vue courant de l'utilisateur sur la visualisation par la mise en évidence (en rouge) à tout instant de la colonne affichée au centre de l'écran ; cette indication complète celle fournie par la colonne sans image dans les vues 3D, ou par le « début » et la « fin » de la

collection dans la vue 2D. Nous tiendrons compte des remarques des participants de l'étude précédente quant à la forme et l'emplacement à l'écran de ce dispositif.

Déplacement latéral

Dans les visualisations 2D matricielles, les interacteurs utilisés sont majoritairement des ascenseurs (ou *scroll-bars*). Outre la navigation dans la collection, ils fournissent également une double information de repérage spatial : d'une part, la position du point de vue courant de l'utilisateur sur la collection et, d'autre part, le rapport entre la taille totale de la représentation et celle de la partie visible. L'intérêt supplémentaire des ascenseurs réside également dans leur grande diffusion et donc dans le degré élevé de familiarité des utilisateurs avec leur utilisation.

Si le choix d'utiliser un ascenseur horizontal peut être largement justifié pour la représentation 2D que nous proposons, il se révèle inapproprié pour les représentations 3D, du fait de leur continuité. On pourrait envisager, à la fin d'un tour du cylindre, de ramener automatiquement l'ascenseur au « début »(*i.e.*, à l'affichage initial). Cette solution n'est pas acceptable car elle est peu compréhensible par l'utilisateur, et impose des mouvements précis de la souris. Si ce dispositif peut éventuellement être utilisé de manière passive, comme indicateur de repérage du point de vue courant de l'utilisateur, il n'est pas adapté au défilement latéral du cylindre.

Dans l'étude précédente, nous avions pris le parti de regrouper toutes les commandes sur une seul interacteur. Si ce choix n'a pas été critiqué en tant que tel par les participants, deux remarques nous poussent à le reconsidérer. D'une part, la position de l'interacteur dans la vue externe a été critiquée car celui-ci masquait une partie des photographies situées dans le bas de l'écran. D'autre part, le nombre de clics nécessaires pour contrôler le défilement latéral et le zoom a été jugé trop élevé, donc nuisible à l'efficacité. Enfin, si l'idée de regrouper les commandes de déplacement dans un même interacteur est séduisante, sa mise en œuvre sous forme de flèches peut paraître complexe à l'utilisateur. Comme l'éloignement-rapprochement de la vue a été beaucoup moins utilisé que le défilement latéral, les flèches de contrôle de l'éloignement-rapprochement de la vue ont été remplacées par un bouton discrètement placé en bas de l'écran. Pour le défilement latéral plusieurs solutions ont été envisagées (voir figure 7.1).

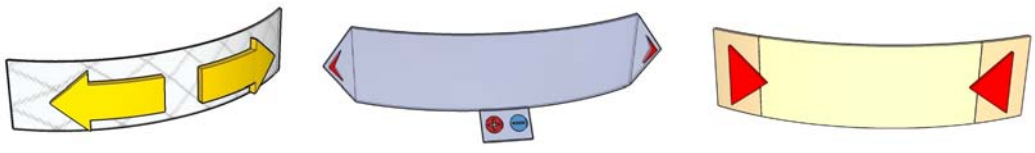


FIG. 7.1 – Trois projets non retenus d'interacteurs permettant l'activation du défilement

La solution retenue consiste à placer deux boutons « sensitifs » sur les bords latéraux de la représentation. Ces boutons présentent quatre avantages sépcifiques :

1. leur activation ne demande pas une précision importante de la part des utilisateurs, car ils font la hauteur de la représentation ;

2. ils ne masquent pas la représentation, étant situés sur les bords gauche et droit de l'écran et disparaissant lors du défilement latéral ;
3. ils sont adaptés aux vues 3D et 2D. De plus, pour les vues 3D ils suivent la courbure de ces vues, ce qui évite toute rupture avec l'impression de profondeur ;
4. le clic n'est pas nécessaire, il suffit de « passer » la souris sur ces boutons. Un court délai de réaction permet d'éviter une activation accidentelle.

Nous avons laissé aux participants la possibilité de changer, au début de l'expérimentation, l'affectation de ces boutons aux deux sens de défilement latéral possibles. En effet, certains participants avaient été déroutés, lors de l'étude précédente, par le sens de rotation affecté par défaut à l'une ou l'autre des vues 3D. C'est l'expérimentateur qui peut exécuter ce changement de sens, à la demande du participant. Cette adaptation permet à chaque participant de se trouver dans des conditions optimales d'utilisation des vues en lui évitant d'avoir à s'adapter à une situation peu naturelle pour lui. Par défaut, l'affectation des boutons proposée est en accord avec la métaphore de la manipulation : la représentation se déplace dans le même sens que le déplacement effectué par la souris pour atteindre le bouton.

Déplacement perpendiculaire au plan

Comme cela a été souligné dans le chapitre précédent, la faible fréquence d'utilisation de cette fonctionnalité autorise à lui associer un interacteur moins visible, même si la présence ou l'intérêt de cette fonctionnalité ne sont pas à remettre en cause. Nous avons choisi un bouton « sensitif » en forme de loupe pour activer le rapprochement ou l'éloignement de la représentation. En positionnant la souris sur cette loupe on fait apparaître deux boutons (voir figure 7.2). Un clic gauche sur le symbole + (respectivement -) permet de rapprocher (respectivement éloigner) la vue.

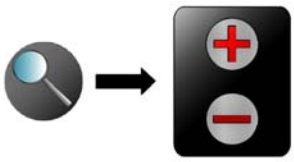


Fig. 7.2 – Dispositif de rapprochement-éloignement de la représentation : le survol du bouton sensitif à gauche fait apparaître la zone de droite comprenant deux boutons activables par un clic (« + » pour approcher, « - » pour éloigner)

Nature des déplacements

Les mêmes déplacements que dans les études précédentes sont reconduits, car ceux-ci n'ont pas été mis en cause par les participants. Le défilement latéral est activé lors du survol du bouton latéral approprié par la souris et commence à vitesse minimum. La vitesse peut être modulée selon 12 pas pour répondre aux exigences de chacun, en fonction des tâches de recherche proposées et de la nature du flux d'informations. La durée nécessaire au parcours intégral de la collection

(de manière identique pour les représentations 2D et 3D) peut ainsi varier entre 2 minutes 30 secondes et 10 secondes. Le déplacement de type éloignement-rapprochement s'effectue quant à lui à vitesse constante, faible, car il est utile pour une adaptation de la distance initiale à la vue de l'utilisateur, mais n'a pas vertu à être utilisé en remplacement du zoom.

Pour stopper le déplacement en cours, quel qu'il soit, il suffit d'effectuer un clic droit en un point quelconque de l'écran.

Zoom

Nous avons conservé le dispositif de zoom de l'étude précédente. Une courte animation a été ajoutée, pour faciliter la transition entre la position de la photographie dans la vue et sa représentation agrandie au centre de l'écran. Cela permet de renforcer la troisième dimension dans le cas des représentations 3D. De plus, un zoom progressif est de nature à aider les utilisateurs à conserver le contexte et la compréhension de l'action en cours [Shneiderman, 1996]. Pour éviter de compliquer inutilement l'interaction, le facteur d'agrandissement est fixe, suffisant pour réaliser les tâches de recherche proposées. L'image agrandie occupe une zone au centre de l'écran de 320x246 pixels (pour une résolution d'écran de 1280x1024 pixels).

Indicateur de repérage spatial

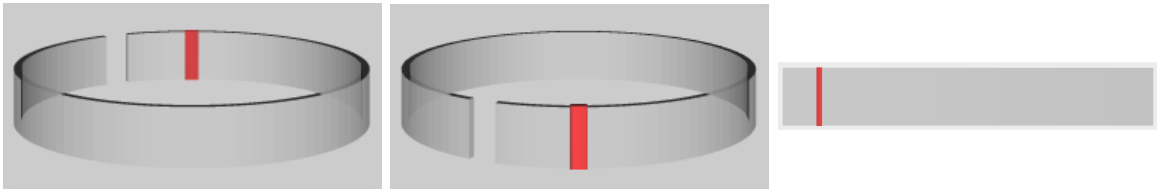
Le dispositif d'interaction de l'étude précédente incluait également le dispositif de repérage spatial, sous la forme d'un cylindre homothétique à la représentation. Bien qu'affiché de manière permanente à l'écran, il n'a été que très peu utilisé par les participants. Pour cette étude, deux options étaient envisageables : soit le maintenir de manière permanente à l'écran, mais en évitant qu'il ne masque même partiellement la représentation, soit le rendre visible à la demande de l'utilisateur. Pour la première option, l'utilisation d'une barre de progression a été envisagée puis écartée, car elle introduisait une rupture visuelle avec les collections de photographies (voir figure 7.3). Il s'avère difficile de réaliser un dispositif qui soit à la fois informatif et peu saillant. Nous avons donc jugé préférable d'afficher l'indicateur de repérage spatial sur demande, afin de ne pas détourner l'attention de l'utilisateur des tâches de recherche visuelle qui lui sont proposées.



FIG. 7.3 – Dispositif de repérage spatial non retenu : barre de progression graduée de 0 à 100 %

Le dispositif de repérage spatial prend la forme d'une mini-représentation homothétique à la représentation courante. Elle permet à tout instant un repérage précis de la position de la colonne affichée au centre de l'écran ; cette colonne est matérialisée par une barre verticale rouge (voir figure 7.4). Les mouvements des vues et ceux de la représentation miniature sont synchronisés. L'inversion de l'affectation des boutons au sens de défilement a également un effet sur l'indicateur de repérage spatial. Par défaut, pour correspondre avec l'affectation initiale des boutons de défilement, c'est la représentation qui se déplace tandis que la barre verticale rouge reste fixe.

Le changement d'affectation se répercute sur la représentation miniature : la mini-représentation reste fixe, tandis que c'est la barre rouge qui se déplace, rappelant la métaphore de la locomotion, dans laquelle c'est le point de vue qui est modifié. Le dispositif de repérage est affiché lors d'une



(a) 3D interne (b) 3D externe (c) 2D

FIG. 7.4 – Dispositifs de repérage pour les 3 vues : la barre rouge indique la colonne affichée au centre de l'écran

pression prolongée sur le bouton central (la roulette) de la souris. Le relâchement du bouton central fait disparaître le dispositif.

Discussion

Nous avons amélioré l'interaction avec la représentation grâce aux commentaires des participants aux études antérieures. Les modifications apportées ici sont relativement mineures et touchent principalement l'interface utilisateur. La représentation de la collection de photographies occupe quasi totalité de l'écran, et les dispositifs de commande et de repérage spatial n'interfèrent plus visuellement avec elle. Ces dispositifs sont maintenant dissociés, pour gagner en clarté et en simplicité. Si les boutons d'éloignement-rapprochement sont moins « visibles », ils n'en restent pas moins accessibles, et d'une utilisation aisée. Les efforts ont porté sur la réduction de la précision exigée des actions à la souris (*e.g.*, grands boutons sensitifs, placés de manière fonctionnelle), ainsi que sur l'élimination des erreurs de manipulation dues à la superposition du dispositif d'interaction sur la représentation de la collection.

7.1.6 Implémentation et développement logiciel

L'atelier de réalité virtuelle *AReVi* [Harrouet *et al.*, 2006] s'étant révélé pertinent pour l'étude précédente, nous l'avons utilisé à nouveau pour l'étude présentée dans ce chapitre. *AReVi* a été employé pour l'implémentation des vues, la réalisation du logiciel permettant le déroulement de la passation et les logiciels de rejeu des interactions. Les développements auxiliaires, comme ceux liés la constitution des collections, ou à l'analyse des traces d'interaction, ont été réalisés en *Java* et en *C++*.

Afin de garantir à la fois la fluidité de l'affichage et la précision des photographies agrandies, nous avons utilisé 2 jeux de photographies : les premières, d'une résolution de 190x142 pixels sont plaquées sur les représentations ; les secondes, d'une résoluton de 512x384 pixels, sont affichées lors de l'agrandissement d'une photographie. Les changements de résolution et de noms ont été automatisés avec le logiciel *XnView*. La figure 7.5 représente cette différence de résolution.



(a) 190x142 (b) 512x384

FIG. 7.5 – Différence de qualité entre les 2 résolutions utilisées pour les photographies sous forme de vignettes (à gauche) et agrandies (à droite)

La possibilté d'importer des modèles *3dsMax* a été exploitée pour la réalisation des boutons de défilement et d'éloignement-rapprochement, ainsi que pour le dispositif de repérage spatial. Tous ces objets graphiques ont été dessinés avec le logiciel *Sketchup*, puis exportés au format *3ds*. Les images associées à l'animation terminant chaque recherche dans T1 ont également été produites avec ce logiciel.

Le matériel utilisé était composé d'un écran LCD Dell de 20 pouces (fonctionnant à la résolution de 1280x1024 pixels), associé à un ordinateur Dell Précision 390, équipé d'un processeur *Intel Pentium 4 Dual Core* à 3,4 GHz, de 2 giga-octets de mémoire vive et d'une carte graphique *NVidia Quadro FX 3000*. L'oculomètre utilisé, à optique sur tête, est le modèle 501 de la société ASL[16] (enregistrement des données à 60 Hz). Le fonctionnement de ce dispositif sera détaillé dans le paragraphe 7.2.8.

La figure 7.6 présente les vues interne et externe du cylindre vertical, tandis que la figure 7.7 présente la vue 2D.

7.2 Protocole expérimental

7.2.1 Généralités

Afin que la durée totale de l'expérimentation ne dépasse pas 2h30, nous avons choisi un nombre de cibles, donc de tâches de recherche visuelle voisin de celui adopté pour l'étude précédente. Cette première étude faisait intervenir 24 cibles, 10 cibles + 2 cibles d'entraînement par vue ; l'augmentation du nombre de vues entraîne de fait une diminution du nombre de cibles par vue. Dans la seconde étude nous utilisons au total 22 cibles, soit 6 photographies par vue, 2 cibles d'entraînement pour la première vue testée, et une cible d'entraînement pour les deux autres vues dans T1. En effet, étant donné que l'interaction est identique dans les vues 3D et 2D, il n'est

[16] www.a-s-l.com



(a) vue interne (b) vue externe

FIG. 7.6 – Vues interne et externe de la représentation cylindrique



FIG. 7.7 – Représentation 2D

pas nécessaire de s'entraîner sur deux cibles, une fois que le fonctionnement de l'interface et les tâches à réaliser ont été assimilés. Dans T2, les mêmes cibles sont naturellement réutilisées, sans les cibles d'entraînement, car les représentations et les moyens d'interaction sont alors familiers aux participants. Si les cibles utilisées pour T1 et T2 sont identiques, leur ordre de présentation diffère d'un type de tâche à l'autre. L'ordre de présentation des vues est quant à lui identique, pour ne pas favoriser l'une ou l'autre vue.

7.2.2 Choix des cibles et descriptions textuelles

Les 6 cibles utilisées pour chaque vue sont présentées dans la figure 7.8 ; les 4 cibles d'entraînement ne sont pas reproduites, car elles n'ont pas été prises en compte dans l'analyse des résultats des participants. Nous avons conservé l'égalité entre objets et paysages, en faisant un



(a) 3D interne



(b) 3D externe



(c) 2D

FIG. 7.8 – Cibles utilisées pour chacune des trois vues pendant T1 et T2

effort supplémentaire pour l'homogénéisation de la difficulté. Si, pour une même vue, les photographies sont de difficulté variable, les niveaux de difficulté sont très proches d'une vue à l'autre. Mis à part pour la cible 2 des trois vues (voie s'étendant vers l'horizon), les photographies ont été choisies pour leur saillance visuelle, afin de limiter le temps de recherche et la fatigue des participants.

Ainsi, pour chaque vue, les 6 cibles sont choisies de la manière suivante :

1. une photographie présentant un contraste important entre l'objet au premier plan et l'arrière-plan (paysage) ;
2. une route s'étendant jusqu'à la ligne d'horizon (paysage) ;
3. un portrait en gros plan sur le visage d'une personne portant une tenue colorée (objet) ;
4. une photographie céleste, avec un élément saillant (paysage) ;
5. un gros plan sur un objet coloré (objet) ;
6. un gros plan sur la tête d'un animal, avec un contraste important entre celle-ci et l'arrière-plan (objet) ;

Comme dans l'étude précédente, les descriptions textuelles sont décomposées en 4 parties afin de faciliter le pré-filtrage et le filtrage des photographies de la collection courante lors de la recherche de la cible. La première partie indique le thème général et les caractéristiques visuelles globales, elle permet de construire une image mentale relativement précise de la photographie à chercher. L'impression en caractères gras met en valeur cette partie par rapport aux autres.

Les descriptions sont présentées dans l'annexe C.

7.2.3 Constitution des collections

Nous disposons pour cette étude de 4400 photographies, la base ayant été étendue par rapport à l'étude précédente pour tenir compte de l'ajout d'une condition dans l'expérimentation. Ces 4400 photographies sont réparties en 5 séries de 880, elles-mêmes classées thématiquement, chaque catégorie comprenant 20 images. Bien entendu, nous disposons d'autant d'objets que de paysages. Comme pour les deux études antérieures, nous ne disposons pas de suffisamment de photographies pour générer les 22 collections nécessaires, plus la collection permettant la prise en main de chaque vue. Notre stratégie est donc d'appliquer une disposition différente des photographies dans la représentation, ainsi qu'un renouvellement partiel des photographies.

Nous avons réservé une série de 880 photographies pour la phase de prise en main et les cibles d'entraînement. Les 4 autres séries ont été utilisées pour constituer les collections. Rappelons qu'une photographie utilisée comme cible ne peut apparaître que dans la collection où elle sert effectivement de cible, de même que les autres photographies du thème. De manière effective, nous disposons de 3160 photographies (4400-880-18*20) plus les 20 photographies du thème de la cible, pour constituer les 18 collections. Nous avons procédé de la manière suivante :

1. utilisation des 20 photographies du thème de la cible (dont la cible) ;
2. tirage aléatoire de 434 photographies parmi 1580 photographies représentant des objets (415 si la cible est un paysage) ;
3. tirage aléatoire de 434 photographies parmi 1580 photographies représentant des paysages (415 si la cible est un objet) ;
4. placement aléatoire des photographies sélectionnées ;
5. placement de la cible à l'endroit désiré dans la collection.

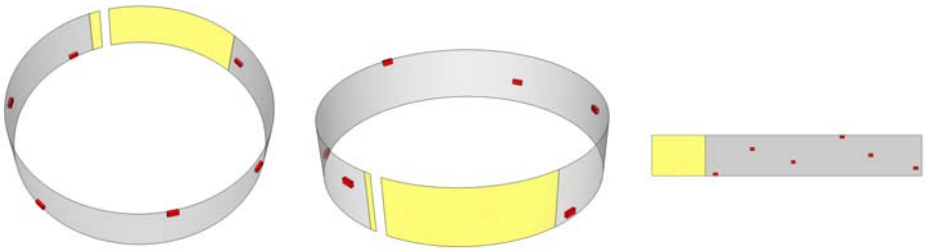
Pour la vue 2D, étant donné que le nombre de photographies composant la collection est légèrement différent (858 contre 869), nous utilisons les 20 photographies du thème de la cible puis 409 photographies appartenant à la même catégorie que la cible (objet ou paysage) et 429 photographies appartenant aux autres thèmes.

7.2.4 Placement des cibles

Les positions des cibles sont réparties régulièrement, horizontalement et verticalement, sur la surface de chaque représentation, afin d'étudier la mémorisation par les sujets des coordonnées horizontales et verticales de ces positions. Aucune cible n'est visible initialement. L'association entre les coordonnées verticales et horizontales des cibles est aléatoire, ainsi que l'ordre dans lequel elles sont placées dans les représentations. Rappelons que l'ordre de présentation des cibles aux participants est différent pour une même vue entre T1 et T2. La figure 7.9 indique la position exacte des cibles dans chaque vue.

7.2.5 Conditions initiales

Pour les vues 3D, les conditions initiales sont identiques à celles adoptées lors des deux études précédentes : 176 photographies sont visibles simultanément (en comptant la colonne sans



(a) 3D interne (b) 3D externe (c) 2D

FIG. 7.9 – Position des cibles : la partie de la vue visible initialement est représentée en jaune, les cibles en rouge

photographie), soit 20,2% de la collection. La colonne sans image est sur la gauche de l'écran, laissant visible une colonne non vierge à sa gauche, pour souligner la continuité du cylindre. Dans la vue 2D, 169 photographies sont visibles simultanément, soit 19,6% de la collection. L'affichage initial représente le « début » de la collection, en partant de la gauche, car le défilement de gauche à droite (ou de haut en bas) de l'affichage est celui que est le plus communément utilisé dans les logiciels grand public. Pour les vues 3D l'affectation par défaut des boutons de défilement latéral est en accord avec la métaphore de la manipulation : le sens du mouvement de la souris nécessaire pour atteindre le bouton est identique au sens de défilement de la représentation.

7.2.6 Choix des participants

20 participants volontaires, 10 hommes et 10 femmes, ont participé à l'expérimentation. Tous étaient des utilisateurs expérimentés d'ordinateurs, avec un niveau d'étude BAC+2 minimum. Le nombre de participants étant relativement restreint, nous avons choisi une tranche d'âge étroite (moyenne : 23,9 - E.T. : 2,31) pour limiter l'hétérogénéité du groupe en termes de performances et de préférences. Aucun participant n'avait de problème visuel particulier d'après le test Bioptor de vision monoculaire et binoculaire (Stereo Optical Company, Inc.). Comme T1 imposait la lecture de descriptions textuelles, les participants avaient tous le français pour langue maternelle.

Les données extraites du questionnaire initial sont synthétisées dans le tableau 7.1.

7.2.7 Déroulement de l'expérimentation

Le déroulement global est similaire à celui de l'étude précédente, mis à part l'ajout d'une condition et la réduction du nombre de cibles par vue : 6 recherches sont donc à effectuer, par type de vue et par tâche. Les 18 recherches de type T1 précèdent les 18 recherches de type T2. Dans T1, 4 recherches supplémentaires (entraînement) permettent aux sujets de prendre en main l'interface et de se familiariser avec le type de tâche à effectuer. Pour la première vue dans T1, chaque participant réalise 2 recherches d'entraînement ; pour les 2 vues suivantes, il ne réalise plus qu'une seule recherche d'entraînement, l'interaction étant identique dans les 3 vues. Ces recherches d'entraînement s'effectuent dans les conditions de l'expérimentation, à savoir

n°	âge	sexe	formation	n°	âge	sexe	formation
1	30	F	Thèse Info.	**11**	22	H	EXIA
2	26	H	Thèse Info.	**12**	21	H	Ingénierie Logicielle
3	27	H	Thèse Info.	**13**	22	F	M2 Info.
4	25	F	M2 Info.	**14**	23	F	Ingénieur
5	26	H	Thèse Info.	**15**	26	H	M2 PRIM
6	24	H	Thèse Info.	**16**	22	H	M2 PRIM
7	22	F	M1 Sc. Biologiques	**17**	23	H	Ingénieur ESSTIN
8	22	H	M2 PRIM	**18**	23	F	Ingénieur
9	22	F	M2 PRIM	**19**	22	F	M1 IST-IE
10	26	F	M2 PRIM	**20**	24	F	M1 IST-IE

TAB. 7.1 – Identification des participants (20) : n° de passation, âge, sexe et formation universitaire

avec un temps limité (3 minutes) et avec l'animation en fin de recherche présentant la cible à trouver et le parcours à effectuer pour l'atteindre, bien que les cibles d'entraînement ne soient pas à chercher lors de T2. Avant les cibles d'entraînement, chaque participant peut librement, et sans contrainte de temps, s'approprier les différentes fonctionnalités de l'interface, en présence de l'expérimentateur qui peut ainsi expliquer son fonctionnement et répondre aux questions éventuelles.

Le déroulement était le suivant :
- Questionnaire initial d'identification (formulaire en pdf) ;
- Test de vision ;
- Présentation générale de l'expérience ;
- Présentation de T1 (tâches à effectuer et structure des descriptions textuelles) ;
- Tâche 1, pour la première vue :
  - présentation de l'interface et des modalités d'interaction ;
  - entraînement sur 2 cibles ;
  - réalisation des 6 tâches de recherche ;
- Tâche 1, pour les deux autres vues :
  - présentation rapide de l'interface ;
  - entraînement sur 1 cible ;
  - réalisation des 6 tâches de recherche ;
- Pose de l'oculomètre et calibrage
- Présentation de T2 (tâches à effectuer) puis, pour chaque vue :
  - réalisation des 6 tâches de recherche ;
- Questionnaire d'évaluation de l'expérimentation et des 3 vues proposées (formulaire pdf) ;
- Entretien post-expérience.

Pour T1, au cours de la présentation de l'interaction avec chaque vue, l'expérimentateur propose au participant d'inverser l'affectation des boutons de défilement. Comme dans l'étude précédente, une animation clôt chaque recherche de type T1, pour montrer au participant quelle

était la cible à chercher, sa position dans la collection ainsi que, pour les vues 3D, le sens optimal de rotation pour l'atteindre. Néanmoins, à la différence de l'étude précédente, nous affichons à la fin de l'animation l'indicateur de repérage spatial à côté de la cible agrandie, afin de faciliter la mémorisation de la position de la cible dans la représentation et , pour les vues 3D, celle du sens optimal de rotation (voir figure 7.10).



FIG. 7.10 – Indications données à la fin de chaque animation dans T1 : la cible agrandie, sa position dans la vue (en rouge) et le « chemin » le plus court pour l'atteindre (en vert)

L'utilisation d'*AReVi* permet d'automatiser la passation. Ainsi, l'expérimentateur n'est pas présent avec les participants (un contact téléphonique est possible), hormis durant les phases de présentation des tâches et des vues (au début de T1), ainsi que lors du calibrage de l'oculomètre et la présentation de T2. Les instructions concernant les tâches et l'interaction avec les vues sont présentées oralement aux participants et sont mises à leur disposition sous forme textuelle pendant toute la durée de l'expérimentation. Les participants peuvent ainsi les consulter en cas de besoin. Les consignes générales propres au déroulement de l'expérimentation sont très similaires à celles de l'étude préliminaire, elles peuvent être consultées dans l'annexe B. Les instructions relatives à l'interaction avec les vues sont présentées dans l'annexe C.

7.2.8 Capture et suivi du regard pendant T2

Afin de valider les observations réalisées lors de l'étude précédente concernant les stratégies d'exploration visuelle des participants (observation des colonnes centrales *versus* observation des colonnes entrantes), nous avons décidé d'utiliser un *eye-tracker* dans cette étude pour enregistrer leurs mouvements oculaires. Les données oculométriques sont utiles également pour évaluer le confort visuel offert par chacune des vues et analyser finement la qualité de la mémorisation de la position verticale des cibles, à partir de l'examen des parcours oculaires. Même si l'utilisation d'un oculomètre pendant T1 aurait pu être intéressante, les activités impliquées dans la réalisation des tâches de type T1, notamment l'utilisation de descriptions textuelles sur papier, d'une part, et la durée de T1 (environ 1 heure), d'autre part, auraient entraîné une fatigue importante des utilisateurs, ce qui aurait pu nuire à la bonne réalisation de T2 et influencer négativement leurs jugements subjectifs.

Le LORIA dispose d'un *eye-tracker ASL* modèle *501* avec une fréquence d'échantillonnage modulable (60, 120 et 240 Hz), et d'un *head-tracker Flock of Birds* de marque *Ascension*. Le dispositif est fixé sur un serre-tête et se porte comme un casque (voir figure 7.11). Une source infrarouge illumine l'œil dont l'image est réfléchie sur un miroir translucide, le reflet est alors enregistré par une caméra infrarouge à l'avant du casque. C'est à partir de la détection de la

pupille, du reflet cornéen et de la distance de l'œil à l'écran que le logiciel calcule la position du regard sur l'écran. Cette distance est calculée à partir des données fournies par l'*head-tracker*, qui mesure la distance entre un capteur magnétique, également fixé sur le casque, et une source magnétique fixe. Le calibrage se fait via une mire à 9 points : l'utilisateur regarde ces points les

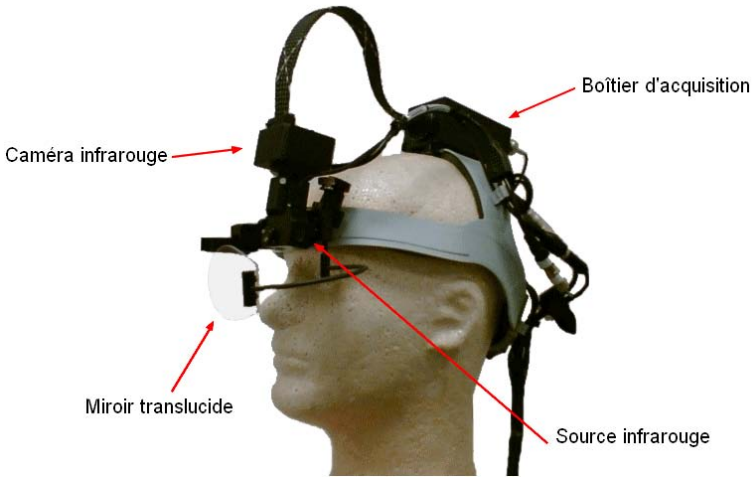


FIG. 7.11 – ASL modèle 501 et head-tracker magnétique

uns après les autres, ce qui permet de détecter la position de la pupille pour ces points clés (voir figure 7.12).

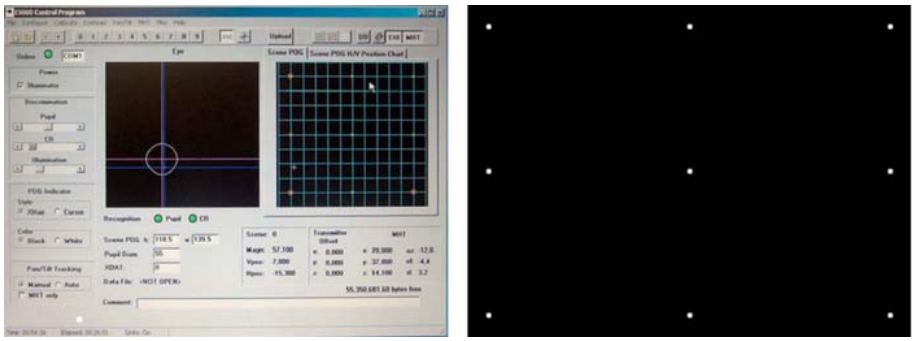


FIG. 7.12 – Interface de calibrage et mire à 9 points

Les données brutes provenant de l'*eye-tracker* (avec intégration des données de l'*head-tracker*) ont été enregistrées durant T2. Les recherches et les vues s'enchaînant automatiquement, une seule calibration avait lieu au début de T2. Le programme permettant l'acquisition et l'enregistrement des données brutes a été fourni par Daniel Gepner, doctorant au sein de l'équipe MErLIn. A l'issue de l'expérimentation, un programme fourni par le fabriquant permet de calculer les fixations, c'est-à-dire les zones sur lesquelles le participant a porté son attention. En plus des données brutes (position du regard, diamètre de la pupille, distance tête-écran), nous disposons donc des fixations (durée, position moyenne) et nous sommes capables d'identifier les saccades, c'est-à-dire les déplacements du regard entre deux fixations.

7.3 Méthodes d'analyse

7.3.1 Sources de données

Comme dans l'étude préliminaire et la première étude, nous disposons des sources de données suivantes :

- les questionnaires initial et final : le premier renseigne sur le profil des participants, et le second sur leurs préférences et jugements subjectifs. Ces questionnaires sont présentés dans l'annexe C ;
- l'entretien de débriefing : l'expérimentateur, en s'appuyant sur un schéma pré-défini, peut recueillir les commentaires des participants de façon moins formelle et plus libre que ne le permet le questionnaire final. En abordant certains points du questionnaire final (*e.g.*, stratégies visuelles, mémorisation spatiale) l'expérimentateur a la possibilité d'obtenir des précisions supplémentaires ;
- la trace des interactions : toutes les interactions sont enregistrées, ce qui permet de les « rejouer », de les annoter et de les analyser qualitativement.

Dans les questionnaires finaux, pour favoriser une étude quantitative et statistique, nous avons eu recours à des échelles de Likert [Likert, 1932] à 7 degrés (voir figure 7.13). Le terme négatif est toujours positionné à gauche pour simplifier la tâche des participants et ainsi favoriser la spontanéité de leurs réponses.

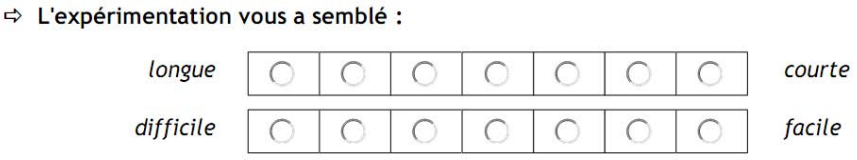


Fig. 7.13 – Exemple de question avec une échelle de Likert à 7 degrés

En plus de ces informations, nous avons utilisé les données fournies par l'oculomètre. Pour T2, le logiciel de « rejeu » des interactions superpose les fixations oculaires aux affichages en les synchronisant avec eux. Ce qui rend possible une étude qualitative fine des stratégies d'exploration visuelle des participants.

7.3.2 Variables étudiées

Pour chaque tâche et chaque vue, nous avons retenu les mesures suivantes :

- **S** : le nombre de cibles détectées avec succès. Le nombre maximum est 6 ;
- **T** : le temps de recherche moyen des cibles (secondes) ;
- **L** : la longueur moyenne de parcours (nombre d'images visualisées) ;
- **V** : la vitesse moyenne de rotation/défilement (images visualisées/seconde).

T1 et T2 impliquant des activités visuelles différentes, nous avons analysé séparément les performances des participants pour T1 et T2. Deux des mesures mentionnées ci-avant n'ont été prises en compte que pour T2 : la longueur de parcours et la vitesse de rotation/défilement. D'une part, lorsque les participants réalisent T1, ils ignorent la position des cibles, contrairement à T2.

D'autre part, ils devaient effectuer T2 le plus vite possible, alors que cette contrainte ne leur était pas imposée pour T1.

La longueur de parcours et la vitesse sont calculées en nombre de miniatures apparaissant à l'écran, et non en colonnes, en raison de la différence de disposition des miniatures sur les affichages des vues 3D et 2D (respectivement 79 colonnes/11 lignes et 66 colonnes/13 lignes). Comme dans l'étude précédente, nous tenons également compte de la position horizontale et verticale des cibles dans l'étude de la mémorisation spatiale. Dans les vues 3D, le sens de rotation est également étudié, principalement pour les cibles proches du point de départ dans T2, car il donne une information sur la mémorisation de leur position lors de T1. Si les données de l'*eye-tracker* peuvent être exploitées de manière qualitative par l'observation des « rejeux », il est également possible d'extraire de ces données des informations permettant une analyse quantitative :

- **NF** : le nombre de fixations ;
- **SP** : la longueur du *scan path*, en pixels, qui correspond à la distance cumulée de fixation en fixation (*i.e.*, la longueur cumulée des saccades sur la période considérée) ;
- **D** : la durée moyenne des fixations (en millisecondes).

Plus précisément, nous calculons le rapport NF/SP, pour obtenir l'amplitude des saccades sur l'intervalle de temps considéré et ainsi avoir une première évaluation du confort visuel. Le cadre précis dans lequel ces calculs vont être menés sera discuté dans la partie correspondante.

7.3.3 Validité interne

Comme pour l'étude précédente, nous allons présenter les différents moyens mis en œuvre pour assurer la validité interne de cette expérimentation. Si nous avons cherché à minimiser autant que faire se peut les variables liées à l'humain, nous avons maîtrisé les données du milieu physique, l'expérimentation ayant eu lieu dans une salle dédiée, climatisée, et dont les fenêtres étaient totalement occultées.

Appariement

L'effectif est similaire à celui de la première étude, à savoir 20 participants. Chacun d'eux a réalisé successivement 6 tâches de type T1 et T2, pour chaque vue (VE, VI et V2D), soit 36 tâches de recherche visuelle au total. Pour mémoire, toutes les tâches de type T1 ont été effectuées avant celles de type T2.

Contre-balancement

Nous sommes exactement dans la même situation que dans l'étude précédente, avec une condition supplémentaire à prendre en compte. Si l'appariement implique un certain apprentissage et une expérience grandissante en cours d'expérimentation, le contre-balancement est un bon moyen de limiter l'impact de ce facteur. Les différents ordres possibles de passation des différentes conditions doivent donc être répartis entre les sujets de façon équilibrée.

S'il suffisait d'avoir deux groupes de sujets dans la première étude puisque les participants expérimentaient deux vues seulement, l'étude présentée ici nécessiterait de définir 6 groupes de participants afin de tester tous les ordres possibles d'expérimentation des trois vues. Ce qui est impossible compte tenu de l'effectif. Nous avons donc décidé d'équilibrer uniquement l'ordre de passation des vues 3D, en gardant la vue 2D constamment en deuxième position. D'une part, puisque nous souhaitons opposer les vues 3D à la vue 2D, il est préférable de ne pas privilégier cette dernière par rapport aux vues 3D. D'autre part, il nous a semblé suffisant de varier uniquement l'ordre de présentation des vues les moins familières aux participants, vues pour lesquelles les phénomènes d'apprentissage risquaient d'être les plus importants. Nous avons donc défini 2 groupes de sujets, chacun comprenant 10 participants et autant de femmes que d'hommes (5). Dans un même groupe, l'ordre de passation est identique pour T1 et T2 afin de garantir une durée de mémorisation similaire pour l'ensemble des cibles.

Affectation aléatoire

Les participants ont été affectés de manière aléatoire à un groupe ou un autre. L'affectation et le placement des cibles ont été discutés précédemment.

Automatisation de la collecte des données

Le procédé (présence de l'expérimentateur, sauvegarde des traces d'interaction, normalisation des explications, questionnaires et débriefings) est strictement identique à celui détaillé dans le chapitre précédent.

7.3.4 Validité externe

Un effectif de 20 sujets nous a permis d'obtenir des résultats statistiquement significatifs dans la première étude. Il est certain qu'un plus grand nombre de sujets permettrait d'obtenir des résultats plus forts. Mais compte tenu de la durée de passation (autour de 2h30) et du temps nécessaire à l'exploitation des résultats, cette taille d'échantillon permet de concilier à la fois une validité minimum et un temps total d'analyse des données acceptable. Le choix d'un groupe de sujets homogène, c'est-à-dire constitué de participants volontaires qui possèdent des connaissances et des compétences similaires et dont les âges varient dans une plage restreinte, tend à compenser l'effectif limité.

Les tâches de recherche étant identiques à celles proposées lors des études exposées dans les chapitres précédents, nous conservons le même niveau de réalisme. D'une part, les activités proposées sont courantes pour le grand public ; d'autre part, nous disposons de documents visuels parmi les plus souvent manipulés, des photographies couleur disponibles sur Internet. Cela doit nous permettre d'obtenir des résultats facilement généralisables à d'autres catégories de documents visuels. Nous sommes ainsi à même d'adresser des recommandations directement exploitables aux concepteurs de logiciels de gestion de collections de documents visuels.

7.3.5 Tests statistiques

L'effectif limité, associé à l'augmentation du nombre de vues à tester, nous a imposé de diminuer le nombre de cibles à chercher, tout en ne constituant que 2 groupes de sujets, composés de 10 participants chacun. Comme dans l'étude précédente, nous avons utilisé des tests non paramétriques pour réaliser les analyses statistiques. Pour comparer les trois vues, nous avons appliqué le test de Friedman puis, lorsqu'il s'est avéré significatif, le test de Wilcoxon sur chaque couple de variables, en raison des différences inter- et intra-individuelles élevées.

7.4 Présentation des résultats

Après une présentation synthétique des résultats obtenus pour les trois vues en termes de performances, préférences et stratégies d'exploration visuelle, nous traiterons deux points particuliers :

- une comparaison entre les deux vues 3D, en vue de déterminer s'ils corroborent ou non les résultats présentés dans le chapitre précédent. Bien que nous ayons réduit le nombre de cibles à chercher, les résultats devraient être similaires, ou tout du moins nous devrions observer des tendances allant dans la même direction ;
- une comparaison entre les vues 3D et la vue 2D. L'idée est de choisir, pour chaque participant, la vue 3D pour laquelle il a obtenu globalement les meilleures performances et de comparer ces performances avec celles qu'il a obtenues avec la vue 2D. La première étude décrite dans le chapitre précédent montre en effet que les participants réussissent mieux avec l'une des deux vues 3D que l'autre, en fonction de leurs capacités de vision et de leurs stratégies d'exploration et de recherche visuelle qui sont, semblent-il, fortement influencées par les capacités visuelles. On n'introduit donc aucun biais en faveur des vues 3D en choisissant, pour les comparaisons avec la vue 2D, la vue 3D avec laquelle chaque participant a obtenu les meilleures performances ; on est dans une situation semblable à celle où pour une tâche manuelle on autorise les participants droitiers (respectivement gauchers) à utiliser leur main droite (respectivement gauche).

7.4.1 Préférences des participants

Dans le questionnaire final, les participants ont évalué chaque vue à l'aide d'échelles de Likert à 7 cases, numérotées de 1 (très mauvais) à 7 (très bon). Les questions portaient sur les points suivants :

- appréciation globale (**G**) : note globale donnée à chaque vue, reflétant la préférence de chaque participant pour l'une ou l'autre des trois vues ;
- facilité du repérage spatial (**RS**) ;
- confort visuel (**CV**) ;
- efficacité de la recherche (**ER**).

Pour les deux derniers critères, les participants ont évalué séparément T1 et T2, car nous avons vu auparavant que ces deux tâches supposent des activités visuelles différentes. Les jugements ont été recoupés avec ceux exprimés lors des entretiens de débriefing.

Le tableau 7.2 récapitule les notes données à chaque vue selon les critères présentés ci-avant.

		G	RS	T1		T2	
				CV	ER	CV	ER
2D	**m.**	4,95	4,65	4,65	4,95	5,00	5,45
	E.T.	*1,51*	*1,27*	*1,63*	*1,40*	*1,62*	*1,47*
VE	**m.**	5,68	4,60	4,80	5,60	5,50	5,85
	E.T.	*1,16*	*1,23*	*1,61*	*1,10*	*1,10*	*0,99*
VI	**m.**	5,05	4,80	4,80	5,30	5,10	5,60
	E.T.	*1,31*	*1,32*	*1,67*	*1,08*	*1,55*	*1,19*
Khi-deux		1,508	0,413	0,543	1,238	0,382	0,246
p		0,471	0,814	0,762	0,538	0,826	0,884

TAB. 7.2 – Evaluations des vues d'après les questionnaires finaux (20 participants) : note globale (**G**), facilité du repérage spatial (**RS**), confort visuel (**CV**) et efficacité de la recherche (**ER**). Tests statistiques de Friedman. En vert, les notes maximales pour chaque critère

Même si les vues 3D (et particulièrement la vue externe) semblent obtenir de meilleurs scores, aucune vue n'est statistiquement meilleure que les deux autres, d'après le test de Friedman. L'éventuelle supériorité des vues 3D sur la vue 2D sera discutée plus loin. Examinons maintenant les préférences globales, sujet par sujet (tableau 7.3).

n°	D.	Q.	n°	D.	Q.
1	E		**11**	E	>E
2	I	3D-I	**12**	2D	
3	I	>I	**13**	E	
4	2D		**14**	2D-I	>2D
5	3D		**15**	I	
6	I	2D-I	**16**	2D-E	
7	I	3D-I	**17**	E	2D-E
8	I		**18**	E	3D-E
9	2D	2D-E	**19**	2D	
10	2D	2D-E	**20**	E	3D-E

TAB. 7.3 – Préférences des participants (20) d'après le questionnaire final (**Q**) et le débriefing (**D**) pour la vue interne (**I**), externe (**E**), **2D** ou pour les 2 vues 3D en général (**3D**). Le signe « > » indique une faible préférence. **2D-I** ou **2D-E** indiquent une hésitation entre ces 2 vues

Sur les 20 participants, 14 préfèrent globalement l'une des 2 vues 3D, avec cette répartition :
- 6 participants préfèrent VI ;
- 7 préfèrent VE ;
- 1 participant apprécie de manière égale les deux vues 3D, qu'il préfère à la vue 2D.

Notons que les choix des femmes (9/10) se sont portés sur VE ou V2D, ce qui tend à confirmer les observations de la première étude, dans laquelle nous avions remarqué que le groupe des participants préférant VE n'était composé que de femmes. Même s'il nous est impossible de conclure statistiquement sur l'interaction entre le sexe et la préférence, ce double résultat a tout de même valeur de tendance, qu'il serait nécessaire de confirmer en augmentant sensiblement la taille de l'effectif.

7.4.2 Performances

Deux des mesures mentionnées dans la section 7.3.2 n'ont été prises en compte que pour T2, pour les mêmes raisons que dans l'étude précédente : la longueur de parcours et la vitesse de rotation/défilement. Les tableaux 7.4 et 7.5 synthétisent toutes les mesures de performances pour T1 et T2 respectivement.

n°	**S**			**T** (secondes)			**n°**	**S**			**T** (secondes)		
	VI	**VE**	**V2D**	**VI**	**VE**	**V2D**		**VI**	**VE**	**V2D**	**VI**	**VE**	**V2D**
1	2	3	5	101	67	109	**11**	2	4	3	125	80	112
2	3	4	5	87	71	95	**12**	1	5	6	45	103	61
3	4	4	6	79	78	79	**13**	3	6	4	116	73	81
4	4	3	4	128	84	111	**14**	6	4	5	75	95	80
5	3	4	4	36	68	45	**15**	4	5	6	90	96	72
6	4	4	5	55	51	84	**16**	3	5	5	100	85	116
7	5	5	5	41	40	57	**17**	3	5	5	124	68	100
8	5	4	6	40	71	45	**18**	5	5	5	80	80	67
9	5	4	6	64	59	63	**19**	4	5	5	115	99	113
10	5	6	6	82	68	50	**20**	5	4	3	42	44	73

Tab. 7.4 – Performances des participants pour T1 : comparaison entre les 3 vues **VI**, **VE** et **V2D** pour le nombre de succès **S** et la moyenne des temps de recherche **T** (secondes)

Pour T1, seul le nombre moyen de succès varie significativement entre les 3 vues, selon le test de Friedman (Khi-deux = 11,906 ; p = 0,003). D'après les tests de Wilcoxon menés sur les couples de variables (voir tableau 7.6), V2D s'avère la meilleure, suivie de VE puis de VI.

Pour l'ensemble des sujets, 89 photographies ont été trouvées sur les 120 avec VE (17 confusions, 14 temps dépassés), contre 76 avec VI (25 confusions, 19 temps dépassés) et 99 avec V2D (8 confusions, 13 temps dépassés). Le fait que nous ne disposions que de 6 cibles par vue nous empêche de conclure à la supériorité de la vue 2D dans T1, certaines cibles ayant occasionné de nombreuses erreurs dans VI (balle de paille, 18 échecs ; sentier, 14 échecs) et dans VE (canal et arc-en-ciel, 13 échecs). En particulier, selon les commentaires des participants, l'image mentale de la balle de paille qu'ils s'étaient faite ne correspondait pas à la cible. Le choix de cette photographie comme cible était sans doute une erreur car son contenu, du fait de l'angle de prise de vue notamment, n'était pas suffisamment stéréotypé. En outre, les participants abordaient l'uti-

n°	S			T (secondes)			L (images)			V (images/secondes)		
	VI	**VE**	**V2D**	**VI**	**VE**	**V2D**	**VI**	**VE**	**V2D**	**VI**	**VE**	**V2D**
1	4	4	5	62	77	82	308	407	343	5,9	5,9	4,6
2	6	5	6	36	43	37	466	447	572	16,8	14,0	18,8
3	6	6	6	27	34	29	510	497	472	24,6	21,1	20,7
4	6	6	6	25	64	48	304	486	648	22,1	12,3	21,6
5	6	5	6	22	19	33	528	411	826	28,4	23,2	29,3
6	5	5	6	47	35	35	458	387	546	16,6	13,4	19,4
7	6	5	5	27	31	25	523	684	473	22,3	29,9	23,8
8	6	6	6	19	17	24	471	279	550	24,8	18,7	25,7
9	5	6	6	41	33	21	1027	611	429	31,1	23,5	28,5
10	6	6	6	36	52	47	297	539	542	9,0	11,3	12,9
11	4	5	6	69	37	86	358	213	381	5,9	7,0	4,6
12	6	4	6	40	74	38	638	930	646	18,0	14,7	22,4
13	6	6	6	48	23	26	502	306	438	13,7	15,2	22,5
14	5	6	6	65	48	57	851	572	752	13,3	13,2	17,3
15	6	6	5	31	34	19	664	486	473	25,1	21,3	31,0
16	6	5	6	26	18	27	460	255	598	20,8	16,7	27,2
17	6	6	6	60	50	51	411	336	425	9,0	8,9	10,2
18	6	6	5	22	24	41	264	183	603	15,6	15,9	26,1
19	6	5	6	28	38	27	363	444	501	14,5	19,3	24,2
20	5	6	6	72	38	25	926	524	427	17,2	17,3	19,7

TAB. 7.5 – Performances des participants pour T2 : comparaison entre les 3 vues **VI**, **VE** et **V2D** pour le nombre de succès **S**, la moyenne des temps de recherche **T** (secondes), la longueur moyenne de parcours **L** (images) et la vitesse moyenne de rotation **V** (images/seconde)

lisation de V2D après celle de l'une des vues 3D ; ils étaient donc tous familiers des tâches de type T1, ce qui n'était pas le cas lorsqu'ils utilisaient les vues 3D en raison du contre-balancement de l'ordre d'utilisation des vues 3D. L'ordre de passation des vues peut donc contribuer à expliquer le nombre nettement moindre de confusions avec V2D qu'avec les vues 3D (8 *versus* 17 et 25).

Pour T2, seule la vitesse moyenne de défilement varie significativement d'une vue à l'autre, d'après le test de Friedman : Khi-deux = 16,900 ; $p < 0{,}005$ (voir tableau 7.7).

La vitesse de défilement supérieure observée avec V2D pendant T2 peut résulter de la familiarité des participants avec la recherche visuelle dans cette vue, puisqu'ils étaient tous des utilisateurs expérimentés de logiciels. En outre, le nombre supérieur de lignes dans la vue 2D ne semble pas avoir ralenti le balayage visuel vertical.

En conclusion, les différences de performances statistiquement significatives entre la vue 2D et les vues 3D sont peu nombreuses ; elles se limitent, pour T1, au nombre de succès et, pour T2, à la vitesse de défilement. La supériorité de la vue 2D pour ces deux critères peut s'expli-

	2D	VE	VI		2D-VE	VI-2D	VI-VE
Moyenne	4,95	4,45	3,8	**Z**	1,842	2,856	1,785
E.T.	0,95	0,83	1,28	**p**	0,066	0,004	0,074

TAB. 7.6 – Nombre de succès par vue dans T1 (20 participants) : moyennes, écarts-types et tests statistiques (Wilcoxon). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances

	2D	VE	VI		2D-VE	VI-2D	VI-VE
Moyenne (images/seconde)	20,5	16,1	17,7	**Z**	3,173	2,614	1,650
E.T.	7,5	5,9	7,2	**p**	0,002	0,009	0,094

TAB. 7.7 – Vitesse de défilement/rotation dans T2 (20 participants) : moyennes, écarts-types et tests statistiques (Wilcoxon) En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances

quer par la familiarité des participants avec cette vue. Plus globalement, Le faible nombre de différences significatives observées peut tenir également à l'importance des variations intra- et inter-individuelles ; un nombre plus élevé de participants et de tâches par vue auraient peut-être permis de mettre en évidence d'autres différences. Pour T2, la facilité et la brièveté d'exécution des tâches de recherche visuelle proposées ont contribué sans doute à masquer d'éventuelles différences.

7.4.3 Facteurs susceptibles d'influencer les préférences

L'analyse croisée des préférences et des performances des participants (nombre de succès, temps de recherche et longueur de parcours) montre que 9 d'entre eux ont des préférences en désaccord avec leurs performances pendant T1 et T2. La majorité d'entre eux justifient leur préférence pour l'une des 3 vues par ses propriétés géométriques qui, selon eux, contribuent à l'efficacité et/ou au confort de la recherche visuelle. Les 6 participants qui préfèrent V2D mentionnent l'absence de déformation (3/6) et le caractère familier (5/6) de cette vue ; en outre, ils soulignent négativement la faible taille des images centrales dans VI (5/6) et positivement la grande taille des images centrales dans VE (4/6). Les 6 participants qui préfèrent VI expliquent leur préférence en citant la meilleure vision globale offerte par VI (5/6 + le sujet qui préfère, sans les distinguer, les vues 3D à la vue 2D) qu'ils jugent particulièrement utile pour réaliser les recherches de T1. Les 7 participants préférant VE avancent la plus grande taille (5/7) et la plus forte saillance (3/7) des miniatures des colonnes centrales tandis qu'ils critiquent l'obligation, avec VI, de focaliser leur regard sur les colonnes latérales (5/7) et la couverture totale de la surface d'affichage par V2D (4/7). L'influence des métaphores d'interaction sur les préférences apparaît, en revanche, marginale : le sentiment d'immersion créé par VI est évoqué de façon neutre par 5 participants et l'impression, avec VE, de manipuler la collection comme un objet virtuel, par seulement 3 participants.

Les propriétés géométriques des vues 3D ont donc influencé les préférences des participants plus que les métaphores d'interaction associées à ces vues ou les performances obtenues avec elles, puisque ce sont ces propriétés qu'ils mentionnent pour justifier leur jugement sur le confort visuel et l'efficacité de la recherche procurés par chaque vue. VI est appréciée pour la vision plus large qu'elle offre, VE pour la taille et la saillance supérieures des images centrales, et V2D pour l'absence de déformation des miniatures et leur taille constante sur tout l'affichage. Les commentaires des participants confirment donc l'hypothèse D et, indirectement, l'hypothèse B concernant la canalisation du regard sur les colonnes centrales avec VE et sur les colonnes latérales avec VI.

Après l'étude des préférences et des performances des participants, nous examinons leurs stratégies d'exploration visuelle. L'utilisation des données oculométriques va nous permettre d'identifier et de caractériser avec précision les « comportements » oculaires individuels et leur évolution éventuelle en fonction des caractéristiques géométriques des trois vues. Les données oculométriques recueillies portent sur T2 exclusivement, c'est-à-dire sur la recherche de cibles familières visuellement et dont la position dans la visualisation est connue.

7.4.4 Stratégies d'exploration visuelle et données oculométriques pour T2

Durant T2, la grande majorité des participants effectue un balayage vertical lors du défilement latéral des vues. Deux stratégies principales ressortent de l'analyse des parcours oculaires : le regard se focalise soit sur les colonnes entrantes (E), soit sur les colonnes centrales (C), comme l'illustre la figure 7.14. Les tableaux 7.8 et 7.9 synthétisent les conclusions de l'analyse qualitative des stratégies d'exploration visuelle des participants. Ces derniers sont regroupés en fonction de leurs préférences : le groupe **GI** (respectivement **GE**) rassemble les participants préférant la vue interne (respectivement externe), et le groupe **G2D** ceux préférant la vue 2D.

<table>
<tr><th rowspan="2">Participant</th><th colspan="6">groupe GI</th><th colspan="7">groupe GE</th><th>3D</th></tr>
<tr><th>2</th><th>3</th><th>6</th><th>7</th><th>8</th><th>15</th><th>1</th><th>11</th><th>13</th><th>16</th><th>17</th><th>18</th><th>20</th><th>5</th></tr>
<tr><td>VI</td><td>?</td><td rowspan="2">C</td><td rowspan="3">E</td><td rowspan="2">C</td><td rowspan="3">E</td><td rowspan="3">C</td><td>E</td><td>?</td><td rowspan="3">?</td><td rowspan="3">C</td><td rowspan="3">C</td><td rowspan="3">E</td><td rowspan="3">E</td><td rowspan="3">E</td></tr>
<tr><td>VE</td><td>C</td><td>C</td><td>C</td></tr>
<tr><td>2D</td><td>?</td><td>L</td><td>E</td><td>E</td><td>L</td></tr>
</table>

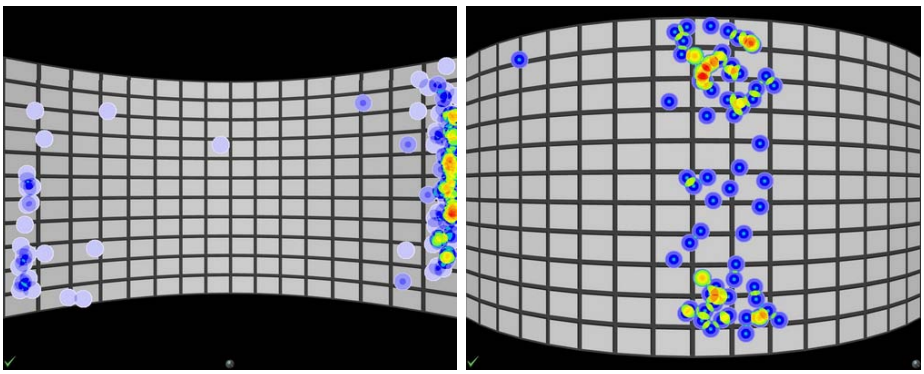
Tab. 7.8 – Stratégies d'exploration visuelle des sujets des groupes GI et GE, et du sujet préférant indifféremment l'une ou l'autre vue 3D : balayage des colonnes entrantes (**E**), des colonnes centrales (**C**), de la largeur de l'écran (**L**) et stratégie indéterminée (**?**)

Nous avons identifié, pour l'ensemble des trois vues, les stratégies de 17 des 20 participants : parmi ceux-ci, 13 gardent une stratégie constante d'une vue à l'autre (10 E et 3 C). Comme pour l'étude précédente, ces résultats confirment que les caractéristiques dynamiques des vues jouent un rôle déterminant dans la stratégie d'exploration visuelle. Parmi les 4 sujets dont la stratégie varie d'une vue à l'autre, la moitié seulement utilisent une stratégie appropriée aux caractéristiques géométriques de chaque vue 3D (C pour VE, E pour VI), ce qui confirme l'influence moindre des propriétés statiques des vues 3D par rapport à leurs caractéristiques dynamiques sur

Participant	Groupe G2D					
	4	**9**	**10**	**12**	**14**	**19**
VI	E	E	E	E	E	E
VE		C				
2D		E				

TAB. 7.9 – Stratégies d'exploration visuelle des sujets du groupe G2D : balayage des colonnes entrantes (**E**) et des colonnes centrales (**C**)

les stratégies visuelles des participants. La vue 2D, par son uniformité visuelle, guide faiblement le regard, ce que confirment nos observations : la vue 2D induit majoritairement (13 participants sur 18) une observation des colonnes entrantes, avec une tendance vers un parcours en largeur de l'affichage, particulièrement marquée pour 2 participants (`S3` et `S11`). Enfin, la distribution des groupes montre que les participants préférant la vue 2D ont une forte inclination à conserver, lors de l'utilisation des vues 3D, la stratégie « colonnes entrantes ». Cela peut être dû à la force de l'habitude, à un certain « conservatisme » dans les stratégies d'exploration visuelle, conservatisme que suggèrent également leur préférence pour la vue 2D et les justifications de cette préférence exprimées dans les débriefings. Les stratégies d'exploration visuelle sont beaucoup plus variées dans les groupes GE et GI où les résistances au changement sont moins fortes.



(a) stratégie *colonnes entrantes* (b) stratégie *colonnes centrales*

FIG. 7.14 – Cartes de température correspondant aux 2 stratégies d'exploration visuelle (1 sujet, 1 cible) : les zones colorées (du bleu au rouge) correspondent à la position des fixations ; les variations de couleur représentent la densité des fixations, croissante du bleu au rouge

L'analyse plus fine des parcours oculaires montre que les vues 3D canalisent davantage le regard que la vue 2D : ainsi, on observe des saccades sur la largeur de l'écran pour 9 participants dans V2D, pour 3 participants dans VE, et pour un seul dans VI. Les vues 3D, en canalisant plus fortement le regard que la vue 2D, sont susceptibles d'accroître le confort de l'exploration visuelle en limitant la longueur des parcours oculaires ; ce résultat contribue à confirmer l'hypothèse B.

7.4.5 Mémorisation de la position verticale et horizontale des cibles

Au delà de la distinction « colonnes entrantes / colonnes centrales », nous avons observé 4 tendances distinctes de « comportements » oculaires, qui traduisent indirectement la qualité de la mémorisation de la position verticale des cibles :

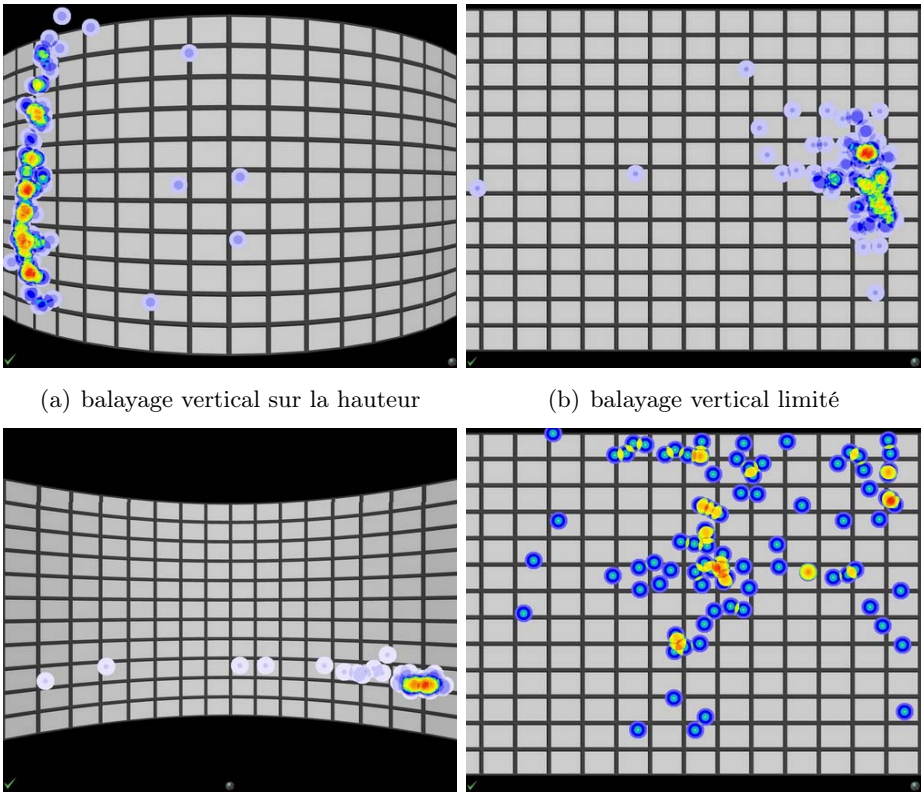
1. le sujet réalise un balayage vertical sur toute la hauteur des colonnes entrantes ou centrales lors du défilement de la représentation ;
2. le sujet réalise un balayage vertical d'une amplitude limitée à quelques photographies, centré sur la ligne où figure la cible ;
3. le sujet observe uniquement la ligne sur laquelle est positionnée la cible, lors du défilement. Ce qui se traduit par une sorte de suivi horizontal (des colonnes entrantes aux colonnes centrales) d'une image, suivi dont l'amplitude varie en fonction de la vitesse de défilement. Le sujet peut également réaliser un balayage horizontal sur la ligne de la cible ;
4. le sujet effectue des parcours visuels horizontaux et verticaux de grande amplitude, avec des fixations dispersées sur la quasi totalité de l'écran.

Si le premier comportement et le dernier semblent indiquer une absence de mémorisation de la position verticale de la cible, le deuxième et surtout le troisième dénotent, à l'inverse, une bonne mémorisation de la position verticale de la cible. Ces comportements sont illustrés par la carte de chaleur présentée dans la figure 7.15 : plus la couleur d'une zone est chaude (*i.e.* rouge), plus le nombre de fixations a été fréquent dans cette zone. Ne sont représentées que les fixations ayant eu lieu lors du défilement de la représentation pour un sujet et une cible.

Sur la figure 7.15 (a), on observe des fixations sur toute la hauteur de la colonne entrante, tandis que l'amplitude de celles-ci est limitée aux lignes 6 à 8 sur la figure 7.15 (b). La figure 7.15 (c) montre que le sujet s'est particulièrement focalisé sur la ligne 3, en observant surtout la colonne entrante, ce qui traduit la mémorisation de la position verticale de la cible. Enfin la figure 7.15 (d) indique une absence de souvenir de la position de la cible dans la mesure où on n'observe aucune focalisation sur une zone particulière de l'écran.

Si les stratégies « colonnes entrantes » et « colonnes centrales » sont utilisées de manière constante d'une recherche à l'autre pour une même vue, il n'en va pas de même pour les comportements présentés précédemment qui, d'une part, peuvent varier d'une recherche à l'autre en fonction de la qualité de la mémorisation et, d'autre part, au sein d'une même recherche. En effet, même si un participant pratique un balayage vertical de la colonne entrante, son attention peut être attirée par une image particulière de la collection, ce qui se traduit par une suite de fixations horizontales balayant la largeur de l'écran. Le tableau 7.10 présente la répartition des comportements par vue, cumulés sur l'ensemble des recherches et des sujets (soit 120 recherches de cibles par vue au total). Ce tableau résume l'analyse qualitative des rejeux, pendant T2, des interactions des participants avec les trois vues.

Le comportement visuel de chaque participant au cours de la recherche de chaque cible a été caractérisé « manuellement » comme suit. L'analyse a porté uniquement sur les moments où la vue défile. La caractérisation du comportement visuel d'un participant au cours de la recherche d'une cible s'appuie sur l'observation de la direction (verticale, horizontale, oblique) des saccades successives. Le comportement est étiqueté, en fonction de la répartition des directions



(a) balayage vertical sur la hauteur

(b) balayage vertical limité

(c) focalisation sur une ligne

(d) parcours irréguliers, horizontaux et verticaux

FIG. 7.15 – Cartes de température correspondant aux 4 comportements visuels identifiés (4 sujets, 4 cibles) : les zones colorées (du bleu au rouge) correspondent à la position des fixations ; les variations de couleur représentent la densité des fixations, croissante du bleu au rouge

des saccades en volume et dans le temps (nature des enchaînements temporels), comme suit : « balayage vertical », « balayage vertical limité », « point fixe » qui correspond à un balayage horizontal sur une seule ligne (à partir des colonnes entrantes), ou « grande amplitude » qui caractérise des parcours irréguliers combinant mouvements horizontaux, verticaux et en diagonale, ou encore « indéterminé » qui caractérise les parcours très brefs ou très variables dans le temps. Un traitement automatique des saccades n'aurait pas permis une caractérisation aussi précise ni aussi robuste des comportements. Les indications quantitatives reportées dans le tableau 7.10 représentent, pour les 120 recherches de cibles effectuées avec chaque vue par l'ensemble des participants, le nombre de comportements relevant de chacune des 5 classes distinguées.

Bien que le tableau 7.10 ne l'exprime pas explicitement, la vue 2D est la plus sensible à la « dispersion » des fixations oculaires sur l'affichage. Pour évaluer cette dimension des comportements oculaires, nous avons éliminé des recherches visuelles des trois premières colonnes du tableau, celles qui présentaient une certaine variabilité comportementale, par exemple pour les recherches de la première colonne (« balayage vertical »), celles qui incluaient des suites de fixations relevant d'autres catégories de comportements (*e.g.*, catégorie « point fixe »). Les recherches

	BV sur hauteur	BV limité	point fixe	grande amplitude	indéterminé
VE	29	30	34	23	4
VI	42	26	24	25	3
2D	34	21	35	18	12

TAB. 7.10 – Cumul des comportements visuels sur l'ensemble des recherches effectuées avec chaque vue par les 20 sujets, pendant T2

visuelles restantes sont qualifiées de comportements stables. Si l'on considère, pour chaque vue, les pourcentages de comportements stables calculés sur l'ensemble des comportements, on obtient 38 % (46/120) pour la vue 2D, 44 % (53/120) pour la vue externe et 47 % (56/120) pour la vue interne. La vue interne semble donc réduire sensiblement la dispersion des fixations oculaires observée avec la vue 2D et, à un degré moindre, avec la vue externe. En d'autres termes, les vues 3D semblent permettre un meilleur contrôle des mouvements oculaires que la vue 2D, ce qui contribue à confirmer l'hypothèse B. Cependant, la supériorité de la vue interne sur la vue externe contredit en partie les attentes exprimées dans cette hypothèse, dans la mesure où les caractéristiques géométriques de la vue externe (*i.e.*, la saillance des colonnes centrales) paraissent avoir eu moins d'influence sur les comportements visuels des participants que celles de la vue interne. Néanmoins, pour être validés, ces résultats qualitatifs doivent être confirmés par les conclusions d'une étude où le nombre de participants et le nombre de recherches visuelles de type T2 effectuées seraient nettement plus importants de façon à rendre possible une analyse statistique pertinente de leur signification, nécessaire pour envisager leur généralisation.

Le pourcentage de cibles dont la position verticale a été mémorisée (comportements de type « point fixe » et « BV limité ») varie également d'une vue à l'autre : de 42 % environ pour VI à 53 % pour VE. Si on ne considère que le souvenir précis de la position verticale (colonne « point fixe »), le pourcentage varie de 20 % pour VI à 30 % pour VE et V2D. Ces résultats suggèrent que VI favorise moins la mémorisation de la position verticale de la cible que VE ou V2D. Ce résultat peut s'expliquer par le fait que la taille des images non déformées des colonnes centrales est maximum dans VE et minimum dans VI ; les positions verticales des cibles se différencient donc moins les unes des autres dans VI que dans VE.

Quant à la mémorisation de la position horizontale des cibles, elle se mesure, pour les vues 3D, par le sens de rotation choisi pour les recherches de type T2, car les positions des 6 cibles dans chaque visualisation définissent des parcours de longueur très différente selon le sens de rotation adopté. 11 participants ont gardé, avec VI, le même sens de rotation pour les 6 cibles de T2, et 9 avec VE, ce qui restreint la portée de l'analyse et de ses conclusions. 2 participants sur les 9 qui ont changé de sens de rotation avec VI ont choisi le sens optimal pour au moins 5 des 6 cibles ; 6 participants sur les 11 qui ont changé de sens avec VE semblent également avoir mémorisé la position horizontale d'au moins 5 cibles sur les 6. Pour comparer les trois vues entre elles, l'indice essentiel est le pourcentage de cibles identifiées dès leur première apparition à l'écran : il est de 84 % pour VI, 87 % pour V2D et 90 % pour VE. Bien que les différences entre ces pourcentages soient relativement faibles, ces résultats ne sont pas en accord avec l'hypothèse C : aucune des trois vues ne semble faciliter sensiblement, par rapport aux autres, la mémorisation

de la position horizontale des cibles dans les visualisations des collections. En outre, VI qui met en œuvre la métaphore de la « locomotion » obtient le score le plus faible et VE le meilleur. Ces résultats incitent à poursuivre les recherches empiriques et expérimentales sur ce thème, en vue de mieux comprendre l'influence éventuelle des trois vues sur la mémoire spatiale et d'être en mesure de déterminer la validité exacte de l'hypothèse C. La difficulté majeure est la conception d'un protocole pertinent qui permette, en particulier, de mettre en œuvre des mesures objectives de la mémorisation de la position des cibles dans les visualisations qui soient (i) non intrusives (*i.e.*, qui ne perturbent pas l'activité des participants), (ii) applicables à chacune des tâches visuelles réalisées par chaque participant, et (iii) utilisables avec chacune des trois vues.

Pour compléter cette analyse de la mémorisation de la position horizontale des cibles avec chacune des trois vues, nous allons évaluer l'intérêt effectif de la « continuité » des vues 3D cylindriques, donc celui de l'utilisation du sens de rotation optimum, en comparant les temps de recherche obtenus avec chacune des vues pour les deux photographies proches de l'affichage initial. Etant donné qu'un seul sujet (S8) a trouvé les deux cibles de chaque vue (VI, VE et V2D) dès leur première apparition et en choisissant le sens optimum de rotation pour les vues 3D, nous avons choisi de calculer, pour chaque vue, la moyenne des temps de recherche de chacune de ces deux cibles vérifiant les contraintes suivantes : la cible a été trouvée dès sa première apparition à l'écran (pour les trois vues), en choisissant le sens de rotation optimum (pour les vues 3D). Les résultats sont présentés dans le tableau 7.11.

	n° cible	**position (col.)**	**moyenne (sec.)**	**E.T.**	**nb. rech**
VI	1	16	18,9	11,7	10
	4	76	10,7	4,6	12
VE	2	77	9,8	4	8
	4	16	16,5	7,2	14
2D	3	65	39,0	30,4	18
	4	16	17,4	5,2	10

Tab. 7.11 – Temps de recherche moyen (secondes), calculé sur l'ensemble des participants (20), pour les deux cibles les plus proches de l'affichage initial, trouvées dès leur première apparition à l'écran et, pour les vues 3D, en choisissant le sens de rotation optimal. « nb. rech » désigne le nombre de recherches visuelles satisfaisant ce(s) critère(s)

Lorsque la cible se trouve à la fin de la visualisation (voir les cibles 4 pour VI, 2 pour VE et 3 pour V2D), on observe que la « continuité » des vues 3D permet de diviser par trois, voire quatre, le temps de recherche moyen obtenu avec la vue 2D, ce qui représente un gain appréciable en efficacité. Ce résultat confirme une des attentes exprimées dans l'hypothèse A (la dernière mentionnée). A noter que, pour les vues 3D, les temps moyens obtenus pour les cibles en colonnes 76 et 77 sont plus courts que ceux correspondant aux cibles en colonne 16. Cette différence pourrait être liée aux gestes à effectuer à la souris pour visualiser la cible, le déplacement de la souris vers la droite étant plus rapide, plus précis et plus naturel que le déplacement vers la gauche. Pour expliquer cette différence, une étude complémentaire serait nécessaire, les données recueillies au cours de cette étude étant insuffisantes pour conclure.

7.4.6 Mesure du confort visuel pendant T2

Les données oculométriques recueillies au cours de T2 nous ont permis de calculer avec précision la position à l'écran des fixations oculaires et leur durée. Nous avons limité l'analyse aux périodes de défilement latéral des vues, car ce sont les moments les plus à même de nous renseigner sur le confort visuel lors de la recherche de photographies. Nous disposons donc, pour chaque sujet et chaque vue, des informations suivantes qui portent sur la durée du défilement de la visualisation pendant chaque recherche visuelle de T2 :

- le nombre de fixations (NF) ;
- la distance cumulée (en pixels) d'une fixation à la suivante pendant le défilement de la visualisation (SP) ;
- la durée moyenne des fixations en milli-secondes (D) ;

Le rapport NF/SP nous donne l'inverse de l'amplitude des saccades : plus ce rapport est grand, plus l'amplitude des saccades est faible. Les résultats sont synthétisés dans le tableau 7.12. Pour chaque mesure, les tests statistiques ont été effectués sur les 20 moyennes des valeurs obtenues pour les 6 recherches visuelles de T2 par chaque participant avec chaque vue. Les moyennes indiquées dans les colonnes de gauche du tableau ont été calculées, pour chaque vue, sur ces 20 moyennes individuelles.

		2D	**VE**	**VI**		**2D-VE**	**VI-2D**	**VI-VE**
NF/SP	**Moyenne**	0,0059	0,0052	0,0061	**Z**	1,9	0,86	3,25
	E.T.	0,0019	0,0010	0,0013	**p**	0,057	0,390	0,001
D	**Moyenne**	246	275	252	**Z**	2,24	0,33	2
	E.T.	54,2	58,6	80,3	**p**	0,025	0,741	0,045

TAB. 7.12 – Pour chaque vue, inverse de l'amplitude des saccades (**NF/SP**) et durée des fixations (**D**) pendant T2 : moyennes (20 participants), écarts-types et tests statistiques (Wilcoxon). NF désigne le nombre moyen de fixations et SP la longueur moyenne du parcours oculaire (en pixels), par recherche visuelle (pendant le défilement de la vue). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances

VI et V2D sont nettement supérieures à VE concernant l'amplitude des saccades et la durée des fixations ; en moyenne, les saccades (donc les parcours oculaires) sont plus courts et les fixations de durée inférieure avec VI et V2D qu'avec VE, et les différences sont statistiquement significatives. La vue externe est donc celle qui prodigue le moins bon confort visuel. Ce résultat est en accord avec les conclusions de l'étude des stratégies d'exploration visuelle des participants. En effet, celles-ci indiquent que la vue externe est la seule des trois vues dont la géométrie dessert la stratégie la plus fréquemment utilisée, à savoir l'observation des colonnes entrantes.

La portée de ces résultats sur l'apport respectif des trois vues au confort visuel de la recherche de cibles familières visuellement (recherches de type T2) dont la position dans la visualisation est connue peut être généralisée à des tâches de recherche visuelle intermédiaires entre celles réalisées pendant T1 et T2, c'est-à-dire la recherche de cibles familières visuellement mais dont la position dans la visualisation est inconnue. En effet, si on supprime les cibles pour lesquelles

les participants effectuent un balayage limité à quelques lignes et celles pour lesquelles les participants n'observent qu'une seule ligne, on obtient celles familières visuellement mais dont la position n'est pas connue. Le tableau 7.13 présente les données oculométriques correspondant à ces tâches de recherche visuelle qui constituent environ la moitié de celles réalisées pendant T2.

		2D	**VE**	**VI**		**2D-VE**	**VI-2D**	**VI-VE**
Nombre de cibles		56	64	50				
NF/SP	**Moyenne**	0,0051	0,0050	0,0056	**Z**	0,66	2,49	2,68
	E.T.	0,0008	0,0009	0,0009	**p**	0,509	0,0128	0,007
D	**Moyenne**	248	262	245	**Z**	1,05	0,61	1,57
	E.T.	55	53,3	75,1	**p**	0,294	0,542	0,116

TAB. 7.13 – Pour chaque vue, inverse de l'amplitude des saccades (**NF/SP**) et durée des fixations (**D**) pendant les recherches visuelles de T2 pour lesquelles la position de la cible n'a pas été mémorisée : moyennes (20 participants), écarts-types et tests statistiques (Wilcoxon). NF désigne le nombre moyen de fixations et SP la longueur moyenne du parcours oculaire (en pixels), par recherche visuelle (pendant le défilement de la vue). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances

Si la durée de fixation moyenne est globalement similaire entre les 3 vues, on peut remarquer que la vue interne est nettement supérieure aux deux autre vues concernant l'amplitude des saccades ; les saccades sont sensiblement plus courtes avec VI qu'avec VE ou V2D et ce résultat est statistiquement significatif. Le confort visuel procuré par la vue interne est donc supérieur à celui procuré par les autres vues. En revanche, bien que la durée moyenne des fixations soit plus courte avec VI qu'avec les autres vues, ce résultat n'est pas significatif. Ce résultat peut s'expliquer, comme précédemment, par les stratégies d'exploration visuelle des participants en majorité de type « colonnes entrantes » sous l'influence prépondérante de la dynamicité des vues. La vue interne et, dans une moindre mesure, la vue 2D, en offrant une bonne visibilité des colonnes latérales, contrairement à la vue externe, facilitent la mise en œuvre de cette stratégie. Ces résultats sont également en accord avec la stabilité supérieure des comportements visuels et le nombre limité de saccades horizontales observées pour les vues 3D, surtout la vue interne. Pour conclure, la vue interne offre donc un meilleur confort visuel que la vue externe et la vue 2D pour les tâches de recherche visuelle d'images familières visuellement.

Dans les deux paragraphes suivants, nous étudions les données recueillies selon deux angles particuliers, d'une part, en comparant uniquement les données obtenues pour les vues interne et externe, afin de confronter les résultats de cette étude avec ceux de l'étude précédente et, d'autre part, en comparant les données relatives à la vue 2D à celles, regroupées, relatives aux deux vues 3D.

7.4.7 Les vues 3D, retour sur les résultats de l'étude précédente

Dans la première étude, nous avons comparé la vue interne à la vue externe : nous allons donc pouvoir confronter les résultats de cette comparaison à ceux de l'étude présentée ici, en

nous focalisant sur les deux vues 3D et en ignorant la vue 2D. Contrairement à l'étude précédente, nous ne disposons que de 6 cibles par tâche et par vue. Ainsi, même si les résultats que nous allons présenter ne peuvent avoir la même portée, nous devrions obtenir des conclusions similaires. Toutefois, l'interface utilisateur ayant quelque peu été modifiée, certaines différences mineures peuvent apparaître. Nous abordons d'abord les préférences, puis les performances et, pour terminer, considérons les stratégies d'exploration visuelle et la qualité du repérage spatial.

Les préférences

Dans la première étude, 14 des 20 participants ont exprimé une nette préférence pour l'une ou l'autre vue 3D, dont une majorité (10) pour la vue interne. Dans la présente étude, les préférences entre les 2 vues sont plus également réparties, puisque 6 sujets ont donné leur préférence à la vue interne, contre 7 pour la vue externe ; à noter qu'un participant supplémentaire exprime sa préférence pour les vues 3D mais sans choisir entre elles.

La note globale donnée dans le questionnaire final de la présente étude est un bon indicateur de la préférence générale des participants : rappelons que cette note est comprise entre 1 et 7, la plus haute note (7) correspondant à une évaluation positive. Les notes sont synthétisées dans le tableau 7.14.

		VI	**VE**
Note globale	**moyenne**	5,05	5,68
	écart-type	1,31	1,16
Wilcoxon	**Z**	1,52	
	p	0,1285	

TAB. 7.14 – VI *vs* VE (seconde étude, 20 participants) : appréciations globales (questionnaire final) et tests statistiques (Wilcoxon). En vert la meilleure performance ; en orange les tendances

On peut observer une tendance assez forte en faveur de VE ($Z = 1{,}52$; $p = 0{,}1285$), ce qui semble en contradiction avec l'étude précédente. Si les préférences sont plus équilibrées dans cette seconde étude, cela peut être dû à la modification de l'interacteur qui, dans la vue externe, ne masque plus la bas de la représentation. Même si ce sont des considérations d'ordre géométrique qui motivent principalement les préférences des participants, les moyens d'interaction sont susceptibles d'exercer une influence importante sur les jugements subjectifs.

Une autre explication peut être avancée pour expliquer ces différences sur les préférences entre les deux études. Dans la présente étude, nous avons donné la possibilité aux participants de modifier l'affectation des boutons de défilement latéral aux sens de rotation lors de la prise en main de chaque vue, le choix par défaut étant conforme à la métaphore de la « manipulation » comme dans l'étude précédente où l'affectation de boutons de défilement était imposée. Or, 12 sujets ont changé ces affectations par défaut pour au moins une des deux vues 3D, dont 9 pour les deux vues. On ne peut donc exclure l'hypothèse que cette évolution de l'interface utilisateur entre les deux études ait exercé une influence sur les préférences pour l'une ou l'autre des deux vues 3D.

Enfin, si le caractère immersif de la vue interne avait été apprécié, dans l'étude précédente, par 5 participants et critiqué par un autre, il a été évoqué par 5 des 20 participants de la présente étude, mais de manière neutre. Cette différence est peut-être due à l'influence de la taille de l'écran sur le sentiment d'immersion : dans l'étude précédente, les vues étaient projetées sur un écran de 1 x 0,75 m, tandis qu'un écran plat de 20 pouces a été utilisé dans la présente étude.

Les performances

La vue interne s'était avérée plus efficace que la vue externe pour les recherches de type T1 dans la première étude : le temps moyen de recherche était plus faible, et le nombre moyen de succès des recherches visuelles plus important. En revanche, les performances obtenues avec VE et VI pendant T2 étaient similaires ; seule la vitesse moyenne de rotation était plus importante avec VI, et la longueur moyenne de parcours plus faible avec VE. Les résultats obtenus dans cette étude pour les mêmes variables suivent-ils ou non les mêmes tendances ? Ces résultats sont synthétisés dans le tableau 7.15.

		moyenne	**écart-type**	**Z**	**p**
Nombre de succès T1	**VI**	5,05	1,31	1,785	0,074
	VE	5,68	1,157		
Temps de recherche T1 (secondes)	**VI**	74,0	17,3	1,22	0,225
	VE	81,0	31,2		
Vitesse T2 (images/seconde)	**VI**	1,61	0,65	1,64	0,101
	VE	1,47	0,54		
Longueur de parcours T2 (images)	**VI**	46,9	19,2	1,33	0,1835
	VE	40,9	15,87		

TAB. 7.15 – VI *vs* VE (seconde étude, 20 sujets) : performances pour T1 (nombre moyen de succès et temps moyen de recherche) et pour T2 (vitesse moyenne de rotation du cylindre et longueur moyenne de parcours), tests statistiques (Wilcoxon). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances

Dans cette étude nous ne disposons que de 6 cibles par vue et par type de tâche, au lieu de 10 dans l'étude précédente. Cependant, on observe des tendances qui confirment les résultats de l'étude précédente, pour T2. Ainsi, la vitesse moyenne de rotation et la longueur moyenne de parcours sont plus élevés avec VI ; les temps de recherche, en revanche, sont similaires. Concernant T1, on n'observe pas la même concordance entre les résultats des deux études. La différence entre les temps moyens de recherche est en faveur de VI comme dans l'étude précédente, mais il s'agit d'une simple tendance statistique. Quant au nombre moyen de succès, il est plus élevé avec VE (forte tendance), contrairement à l'étude précédente. Ce résultat porte sur 6 cibles seulement au lieu de 10 ; en outre, les cibles utilisées avec VI et VE étaient différentes ; or, il est quasi impossible de contrôler parfaitement la difficulté de la recherche de cibles décrites verbalement. Le nombre restreint de cibles utilisées pour T1 dans cette étude peut donc expliquer la différence observée entre les deux études concernant le nombre moyen de succès pendant T1.

En conclusion, concernant les performances des participants, la plupart des résultats des deux études sont en accord, ce qui les conforte.

Les stratégies d'exploration visuelle

Dans la première étude, nous avons identifié les stratégies d'exploration visuelle de 16 des 20 participants : 10 conservaient une stratégie constante pour les deux vues (8 « E » et 2 « C ») quand 6 utilisaient une stratégie flexible. Parmi ces derniers, 4 utilisaient une stratégie « pertinente », c'est-à-dire en accord avec la géométrie de la vue : observation des colonnes entrantes avec VI et des colonnes centrales avec VE. Nous avions également observé l'influence prépondérante des caractéristiques dynamiques des vues cylindriques sur ces stratégies, 13 sujets sur 17 s'étant focalisés sur les colonnes entrantes avec VI, alors que seulement 7 participants sur 17 avaient centré leur attention visuelle sur les colonnes centrales avec VE.

Dans cette seconde étude, les résultats sont très similaires, puisque sur les 17 participants dont nous avons pu identifier la stratégie, 15 conservent la même stratégie (10 « E » et 5 « C ») quand 2 utilisent une stratégie pertinente (voir les tableaux 7.8 et 7.9). Si 12 des 17 participants observent les colonnes entrantes avec VI, seulement 9 (sur les 17) observent les colonnes centrales avec VE.

La conclusion de l'étude précédente concernant la prédominance de l'influence des caractéristiques dynamiques des vues sur les stratégies d'exploration visuelle, au détriment de celle de leurs caractéristiques géométriques, est donc confirmée.

La qualité du repérage spatial

Nous disposions, dans la première étude, de 10 cibles, dont 2 étaient visibles initialement, et 3 (non visibles initialement) étaient proches de l'affichage initial. Pour les deux premières, on suppose qu'une bonne mémorisation spatiale implique l'absence d'activation de la rotation, tandis que pour les trois autres, on fait l'hypothèse que les participants ayant mémorisé la position de la cible activeraient le sens de rotation minimisant la distance à parcourir pour atteindre la cible, et donc aussi le temps de recherche. Sur les deux cibles visibles initialement, les résultats étaient similaires entre les deux vues. Pour VE, 6 sujets (cible 2) et 20 sujets (cible 4) n'ont pas activé de rotation ; pour VI, on compte 3 sujets (cible 3) et 16 sujets (cible 4). Concernant les trois autres cibles, non visibles au départ de la recherche, l'analyse des traces des interactions suggère que les 6 participants ayant changé de sens de rotation au cours de l'expérimentation ont en majorité mémorisé la position des cibles : sur les 3 cibles considérées, le chemin le plus court a été choisi en moyenne 2,5 fois pour VI contre 1,75 fois pour VE.

Compte tenu du faible nombre de cibles utilisées dans la seconde étude, nous n'avons pas placé de cible directement dans l'affichage initial. On observe pour l'ensemble des 6 cibles les comportements suivants : 2 participants (`S5` et `S11`) sur les 9 qui ont changé de sens de rotation avec VI ont choisi le sens optimal pour au moins 5 des 6 cibles contre 6 participants (`S3`, `S5`, `S11`, `S15`, `S17` et `S18`) sur les 11 qui ont changé de sens avec VE. Ces résultats suggèrent une meilleure mémorisation avec la vue externe, et contredisent donc les résultats de l'étude précédente. Ils

sont en accord avec les observations des comportements visuels des participants : en considérant les comportements de type « balayage vertical limité » et « point fixe », on remarque que la vue externe est sensiblement meilleure (64 comportements de ce type identifiés avec VE, contre 50 avec VI). Nous pouvons de plus observer que les participants ayant mieux mémorisé la position des cibles donnent tous leur préférence à une des deux vues 3D.

Discussion

La comparaison entre les résultats de la première étude et ceux de la seconde étude qui portent sur les vues 3D conduit à nuancer les conclusions de chacune d'elles.

Les préférences sont davantage équilibrées entre les deux vues dans la seconde étude, alors que, dans la première, les participants préférant VI étaient nettement plus nombreux que ceux préférant VE. Cette différence est peut-être due à la modification de l'interacteur de défilement d'une étude à l'autre.

Concernant les performances, on observe des résultats concordants entre les deux études pour T2 : la vitesse de rotation avec VI est supérieure à celle obtenue avec VE, tandis que la longueur de parcours est inférieure avec VE par rapport à VI. Les temps de recherche similaires entre les deux vues peuvent résulter des effets contraires de ces différences. Concernant T1, on observe dans les deux études un temps de recherche visuelle plus faible avec VI qu'avec VE. Ce qui suggère que la vue interne se prête mieux que la vue externe à une exploration visuelle de type filtrage qui caractérise la recherche d'une image vérifiant un ensemble de propriétés visuelles et thématiques ; il semble en effet, d'après les débriefings, que VI facilite davantage que VE un survol global rapide d'une fraction importante des images affichées simultanément, alors que VE favorise plutôt un examen visuel détaillé d'un nombre limité de ces images. En revanche, les différences entre les deux vues quant au nombre de succès des recherches visuelles de type T1 s'inversent d'une étude à l'autre. Ce qui peut être dû à la quasi impossibilité de contrôler la difficulté des tâches de recherche d'une image à partir d'une description textuelle de ses caractéristiques visuelles et thématiques. Les résultats obtenus dans la première étude, à savoir la supériorité de VI paraissent plus sûrs que ceux de la seconde étude car ils portent sur un nombre de tâches nettement supérieur (10 contre 6), ce qui réduit les effets éventuels d'un contrôle imparfait de la difficulté des tâches de type T1.

Les résultats des deux études concordent parfaitement quant aux stratégies de parcours oculaire des visualisations mises en œuvre par les participants. Les données oculométriques de la seconde étude confirment les conclusions de la première étude issues de l'analyse des traces des interactions et des commentaires des participants dans les questionnaires finaux et lors des débriefings. Ces résultats représentent donc une réelle contribution scientifique. Ils mettent en évidence deux modes de parcours des visualisations lors de leur défilement latéral : focalisation de l'attention visuelle sur les colonnes entrantes (*i.e.*, celles qui apparaissent lors du défilement latéral) ou sur les colonnes centrales. Un même participant peut utiliser en alternance l'un et l'autre mode en fonction des caractéristiques géométriques de la vue (interne ou externe), ou conserver le même mode de parcours avec les deux vues. On obtient donc quatre stratégies de parcours oculaire possibles qui ont toutes les quatre été observées lors de l'analyse des données

des deux études. Une majorité des participants dont on a pu identifier les stratégies d'exploration visuelle conservent une stratégie constante d'une vue à l'autre. Les sujets qui changent de stratégie d'une vue à l'autre sont peu nombreux, et ceux qui adoptent une stratégie pertinente (colonnes entrantes avec VI et colonnes centrales avec VE) sont encore moins nombreux. Ces résultats montrent donc l'influence prépondérante des caractéristiques dynamiques des vues sur les stratégies d'exploration visuelle, par rapport à celle de leurs caractéristiques géométriques.

Pour déterminer l'apport des vues 3D et des métaphores d'interaction associées (locomotion *versus* manipulation) à la mémorisation de la position des cibles dans les visualisations 3D, nous avons distingué, par souci d'efficacité et de clarté, la mémorisation de la position verticale de la cible de la mémorisation de sa position horizontale dans la visualisation de la collection. Concernant la position verticale de la cible, les deux études aboutissent à des conclusions voisines. Les données oculométriques (seconde étude) confirment les conclusions de l'analyse des commentaires des participants à la première étude (questionnaires finaux et débriefings), à savoir que la position verticale des cibles a été souvent mémorisée avec plus ou moins de précision. Les données oculométriques montrent que, globalement, la position verticale de la moitié des cibles environ a été mémorisée ; à noter la supériorité de VE (53 % de positions mémorisées) sur VI (42 %), qui peut s'expliquer par les caractéristiques géométriques spécifiques des deux vues. Concernant la mémorisation de la position horizontale des cibles, les résultats diffèrent entre les deux études. Or, ils reposent sur deux mesures seulement : l'absence d'activation de la rotation pour les deux cibles de la première étude présentes dans l'affichage initial et, dans les deux études, le sens de rotation choisi pour la recherche des cibles non visibles initialement. En outre, la seconde mesure n'a pu être appliquée qu'à un nombre très limité de sujets. En effet, la grande majorité des participants aux deux études ont choisi le même sens de rotation pour toutes les cibles de T2, malgré la consigne (donnée oralement au début de T2) qui leur imposait d'effectuer les recherches le plus rapidement possible. Toutefois, le fait que la majorité des participants n'aient pas choisi le sens de rotation optimal pour atteindre les cibles ne signifie pas forcément qu'ils n'ont pas mémorisé sa position horizontale. Il a été montré que la mémoire spatiale s'appuie sur des repères (*landmarks*) [Mou *et al.*, 2006], ce qui impose, pour être en mesure d'exploiter ces repères lors de la recherche d'un objet dont l'emplacement est connu, de parcourir à nouveau le chemin/itinéraire emprunté initialement lors de sa découverte. Pour déterminer l'influence éventuelle, sur la mémoire spatiale, des deux vues et des métaphores d'interaction qu'elles suggèrent, il est nécessaire de concevoir des mesures fiables et non intrusives, c'est-à-dire des mesures qui permettent d'évaluer la qualité de la mémorisation de la position horizontale des cibles pour tous les participants en évitant de susciter de leur part un effort volontaire de mémorisation. Sans de telles mesures, il est impossible de déterminer l'assistance éventuelle que sont susceptibles d'apporter à la mémoire spatiale différentes visualisations interactives de collections de documents visuels, dans le contexte de tâches de recherche visuelle réalistes.

Enfin, les données oculométriques (voir la section 7.4.6) montrent que le confort visuel est meilleur avec la vue interne qu'avec la vue externe. En effet, l'amplitude moyenne des saccades et la durée moyenne des fixations oculaires sont plus courtes avec VI qu'avec VE (voir le tableau 7.12), et les différences sont statistiquement significatives. Ce résultat confirme les interprétations et conclusions qu'il est possible d'inférer des données de la première étude et qui ont été présentées dans le chapitre précédent ; voir la section 6.7.

7.4.8 Comparaison 2D-3D

Après avoir passé en revue les résultats de la comparaison des deux vues 3D, nous allons étudier l'apport éventuel spécifique de la troisième dimension, en comparant les vues 3D à la vue 2D à partir des données recueillies dans le cadre de la seconde étude qui est l'objet de ce chapitre. La méthodologie utilisée consiste à considérer, pour chaque sujet, la vue 3D pour laquelle les données quantitatives et qualitatives obtenues sont les meilleures. Dans la mesure où les jugements subjectifs et les performances sont de nature différente et ont été évalués par des mesures différentes dans cette étude, on peut, pour chaque vue, scinder les données recueillies pour un même sujet en deux ensembles, l'un comprenant les données relatives aux préférences, l'autre celles concernant les performances ; ensuite, on choisit pour les préférences (respectivement, les performances) l'ensemble de données relatives à l'utilisation de VI ou de VE qui contient les appréciations les plus positives (respectivement, les performances les plus élevées). On peut appliquer cette méthode au traitement des données de cette étude, sans risquer d'introduire un biais en faveur des vues 3D, car les appréciations des 20 participants sur les deux vues 3D se différencient nettement et présentent une cohérence interne forte, en accord avec leur préférence pour l'une ou l'autre vue ; de même, leurs performances varient d'une vue à l'autre dans le même sens.

Un protocole plus rigoureux aurait consisté à concevoir deux sessions faisant intervenir les mêmes participants. Au cours de la première session, ceux-ci utilisent les deux vues 3D. En fonction des résultats de l'analyse de leurs performances au cours la première session, ils utilisent, au cours de la seconde session, la vue 2D et soit la vue 3D où ils ont obtenus les meilleures performances, soit celle pour laquelle ils émis les avis les plus positifs. Ce protocole présente deux inconvénients. Il est lourd à mettre en œuvre car il est difficile de trouver des volontaires qui acceptent de participer à deux expérimentations ayant lieu à deux dates différentes (disponibilité) et portant sur des tâches semblables (motivation). En outre, on doit, soit éliminer (après la première session) les participants dont les préférences ne sont pas en accord avec leurs performances, soit solliciter leur participation à trois sessions, ce qui est difficilement réalisable.

Les deux paragraphes suivants décrivent les résultats de la méthode décrite ci-dessus, d'abord aux jugements subjectifs, puis aux performances des participants à la seconde étude.

Préférences des participants

Comme nous l'avons vu précédemment, la majorité des participants (14 sur 20, 5 femmes et 9 hommes) ont une préférence globale plus ou moins marquée pour l'une des deux vues 3D. Rappelons également que 9 participants de sexe masculin sur 10 préfèrent une des vues 3D, alors que 5 femmes sur 10 préfèrent la vue 2D ; cependant, les effectifs sont insuffisants pour qu'il soit possible de conclure à une influence quelconque du sexe sur les préférences des participants, même si la distribution observée ne semble pas être due au hasard.

Les appréciations exprimées dans les questionnaires finaux à l'aide d'échelles de Lickert à 7 degrés comprennent, pour chaque vue, outre une note globale (**G**), l'évaluation de la facilité du repérage spatial (**RS**), du confort visuel pour T1 et T2 (**CV**) et de l'efficacité de la recherche

pour T1 et T2 (**ER**). Pour 15 sujets sur les 20, les différences entre les notes qu'ils ont affectées à VI et à VE sont toutes de même signe. Pour 3 sujets (`S12`, `S13` et `S20`), une différence sur les 6 est de signe contraire aux 5 autres, pour le sujet `11` ce nombre passe à 2, et pour le sujet `14`, il passe à 3. Pour les sujets `11`, `12`, `13` et `20`, les données de la vue 3D comprenant le plus d'appréciations positives ont été utilisées ; pour le sujet `14`, la vue pour laquelle la somme des notes est la plus élevée a été choisie.

		G	**RS**	**T1**		**T2**	
				CV	**ER**	**CV**	**ER**
2D	**m.**	4,95	4,65	4,65	4,95	5,00	5,45
	E.T.	1,51	1,27	1,63	1,40	1,62	1,47
3D	**m.**	6,05	4,85	5,70	5,90	5,55	6,10
	E.T.	0,91	1,31	0,98	0,85	1,43	0,91
Z		2,295	0,60	2,131	2,57	1,38	1,927
p		0,022	0,548	0,033	0,010	0,168	0,054

TAB. 7.16 – Jugements subjectifs sur la vue 2D et la vue 3D (préférée), d'après les questionnaires finaux : note globale (**G**), facilité du repérage spatial (**RS**), confort visuel (**CV**) et efficacité de la recherche (**ER**). Moyennes (20 participants), écarts-types et tests statistiques (Wilcoxon). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances

Sur tous les critères d'évaluation subjective proposés, la vue 2D est perçue comme inférieure à la vue 3D que les participants ont préférée : les tests sont statistiquement significatifs, sauf pour l'efficacité de la recherche dans T2 (tendance forte) et la qualité du repérage spatial (non significatif). Ces résultats montrent que les participants attribuent à la troisième dimension la capacité d'améliorer sensiblement, par rapport aux vues 2D classiques, l'efficacité de la recherche et le confort visuel des tâches de type T1 qui impliquent un survol de la visualisation pour détecter les images susceptibles de correspondre aux descriptions textuelles. Concernant les tâches de type T2, l'apport de la troisième dimension à l'efficacité de la recherche est jugé moindre (simple tendance) ; quant à sa contribution au confort visuel, elle paraît inexistante (résultat non significatif). La facilité relative des tâches de type T2 par rapport aux tâches de type T1 peut expliquer ces différences d'appréciation. Enfin, la troisième dimension ne permet pas, selon les participants d'améliorer la qualité du repérage spatial. Ces résultats sont en accord avec les jugements subjectifs exprimés spontanément au cours des entretiens de débriefing.

Performances

La méthode de regroupement des données proposée ne peut être appliquée aux mesures des performances. En effet, 8 participants sur les 20 ont des performances nettement supérieures avec une des deux vues 3D pour les tâches de type T1 et avec l'autre vue pour les tâches de type T2 ; 5 d'entre eux obtiennent des résultats nettement supérieurs avec VE pendant T1, et avec VI pendant T1 ; les résultats des comparaisons en terme d'efficacité s'inversent pour les 3

autres. Parmi les 12 autres participants, 6 sont plus efficaces avec VE pendant T1 et T2, et un seul avec VI ; quant aux 5 autres, leurs performances sont, soit identiques avec les deux vues pour les deux types de tâche (1 participant), soit supérieures avec l'une des vues pour l'un des types de tâche et identiques avec les deux vues pour l'autre type de tâche. Le faible nombre de recherches qui accentue les effets des différences de « difficulté » entre les cibles à rechercher avec les trois vues peut expliquer, mais en partie seulement, cette variabilité des performances individuelles. La question du choix d'une visualisation de collections de documents visuels qui accroisse sensiblement l'efficacité de la recherche visuelle d'un ou plusieurs items particulier(s) dans de telles collections reste ouverte.

7.5 Commentaires des participants sur l'interface : suggestions d'améliorations

A l'issue de l'étude préliminaire, nous avons modifié sensiblement l'interface utilisateur des visualisations, afin d'offrir aux participants des moyens identiques d'interaction avec les trois vues (interne, externe et 2D) qui, en outre, soient compatibles avec les métaphores suggérées par les deux vues 3D, locomotion et manipulation, respectivement ; c'est essentiellement le sens de rotation affecté par défaut aux vues 3D, différent d'une vue à l'autre, qui assure cette compatibilité. Les études rapportées dans ce mémoire sont centrées sur l'apport de la troisième dimension aux activités d'exploration et de recherche dans des visualisations 3D et 2D de grandes collections d'objets visuels. L'évaluation de l'interface d'interaction avec les visualisations étudiées est d'une importance secondaire, quoique non négligeable, par rapport à cet objectif. Cette interface doit être « transparente » au sens où elle ne doit pas interférer avec les stratégies de recherche et d'exploration visuelle des participants. En outre, sa qualité ergonomique doit être suffisante pour que les concepteurs de logiciels de gestion de collections de photographies destinés au grand public envisagent et soient en mesure d'implémenter les vues externe et/ou interne de façon efficace et conviviale. C'est pourquoi nous passons en revue dans cette section les différents commentaires des participants (dans les questionnaires et au cours des débriefings) sur l'interface utilisateur commune aux trois vues, et discutons les modifications qu'ils ont suggérées pour améliorer son ergonomie.

7.5.1 Représentations visuelles des collections

En plus des observations expliquant leurs préférences pour l'une ou l'autre vue, les sujets ont également formulé des remarques d'ordre plus général sur les représentations 3D des collections.

Le choix même de la représentation cylindrique est approuvé par l'ensemble des participants. Deux d'entre eux, en particulier, apprécient le fait que, contrairement à la vue 2D, la vue interne et, à un degré moindre, la vue externe n'occupent pas toute la hauteur de l'écran (`S5` et `S6`). Néanmoins, le participant `6` aurait souhaité une vue trapézoïdale (avec les deux bases placées verticalement) en lieu et place de la représentation cylindrique, afin d'avoir un niveau de détail croissant latéralement. Une telle vue peut être obtenue aisément à partir de la vue externe : il suffit de n'afficher qu'une des deux « moitiés » de VE. Ainsi, le niveau de détail serait faible sur

les colonnes entrantes, favorisant un examen rapide, tandis qu'il serait plus élevé sur les colonnes sortantes, permettant un examen plus détaillé des photographies. L'avantage de cette solution serait un guidage du regard sur la totalité de la largeur de l'écran ; cela étant, il serait plus difficile de mettre en œuvre un déplacement latéral de la vue qui assure la « continuité » de la représentation et soit compréhensible facilement.

Certains participants ont regretté, lors de l'interaction avec la vue 2D, l'absence de continuité de la représentation, continuité qui caractérise les vues cylindriques :

> *On ne peut pas aller vers la fin depuis le début sans avoir à parcourir tout la collection, contrairement aux vues 3D.* `Sujet 9`

La continuité est une propriété dont l'application à des représentations planes ne peut qu'être artificielle et, de ce fait, risque de dérouter l'utilisateur et de lui poser des problèmes d'utilisation. C'est pourquoi elle n'a pas été mise en œuvre dans la seconde étude.

La réorganisation de la disposition des items dans l'affichage et la modification du nombre d'items affichés simultanément ont été souhaitées par plusieurs participants. Ces suggestions portent principalement sur la vue 2D pour laquelle nous avons choisi de couvrir la totalité de l'écran. La réduction du nombre de photographies présentes dans un même affichage favoriserait un meilleur confort visuel en réduisant la couverture de l'écran. A l'inverse, augmenter le nombre de photographies affichées simultanément en réduisant leur taille offrirait une vision plus globale de la collection. Ces souhaits contradictoires sont surprenants car les participants pouvaient rapprocher ou éloigner d'eux la vue 2D de la même manière que les vues 3D. Le fait que les vues 2D classiques des logiciels courants n'offrent pas cette possibilité peut expliquer pourquoi ils n'ont pas songé à l'utiliser.

7.5.2 Interaction avec les visualisations

Quatre sujets (`S6`, `S12`, `S15` et `S19`) ont critiqué la lenteur de réaction des boutons de défilement latéral :

> *Le temps de latence est trop important pour lancer la rotation, ce n'est pas assez réactif.* `Sujet 6`

La solution adoptée visait à éviter des activations involontaires et donc intempestives de ces boutons, d'une part, en introduisant un délai de 300 millisecondes entre le passage de la souris sur l'un des deux boutons et l'activation de la rotation, et d'autre part, en exigeant un balayage du bouton par la souris d'une amplitude d'au moins 1 pixel. Il faudrait rendre ces mesures de précaution moins perceptibles par les utilisateurs : la suppression de la seconde contrainte permettrait de gagner en fluidité, car c'est l'obligation de déplacement, mal comprise, qui induit l'impression de latence.

Une autre suggestion porte non pas sur l'amélioration de l'interacteur existant, mais sur son remplacement par un contrôle complet du défilement (vitesse, lancement et arrêt) par la roulette de la souris :

> *J'aurais aimé pouvoir contrôler tout le défilement avec la molette, cela aurait évité de distraire le regard.* `Sujet 18`

> *Pouvoir faire tourner le cylindre avec la molette serait plus intuitif que d'utiliser les boutons latéraux, car cette action demande un déplacement de la souris.* `Sujet 7`

L'utilisation de la roulette de la souris pour contrôler le défilement est séduisante, surtout pour les vues 3D en raison de l'analogie évidente entre la forme de la roulette et le cylindre qui sert de support aux vues 3D. Pourtant, cette solution a été écartée pour plusieurs raisons. D'une part, elle implique des actions peu commodes (*i.e.*, l'enfoncement par pression) sur la roulette de la souris lors des phases d'arrêt et de redémarrage de la rotation ; la répétition de ces actions peut provoquer une fatigue musculaire non négligeable. D'autre part, elle complique la gestion du retour à la vitesse nulle : si l'utilisateur souhaite arrêter un défilement à vitesse élevée, il lui faudra faire preuve de rapidité et de précision dans l'anticipation et la synchronisation du geste de pression sur la roulette pour obtenir l'arrêt du défilement sur les photographies qu'il souhaite examiner plus longuement.

Enfin, deux sujets (`S11` et `S14`) auraient préféré une activation du défilement par clic, quand deux autres sujets (`S3` et `S13`) auraient apprécié de pouvoir manipuler le cylindre à la souris en utilisant l'action de *drag*. Le choix de l'activation par survol a été discuté au début de ce chapitre. Quant à la manipulation directe de la représentation (*i.e.*, le recours au *drag*), elle est incompatible avec la métaphore de la locomotion associée à la vue interne, ce qui explique pourquoi elle a été abandonnée après l'étude préliminaire.

Il est facile, en interaction homme-machine, d'imaginer de nouvelles solutions conçues pour remplacer ou améliorer l'existant. Cependant si la plupart d'entre elles paraissent séduisantes au premier abord, rares sont celles qui s'avèrent satisfaisantes au terme d'un usage prolongé en situation réelle d'utilisation. On peut offrir le choix à l'utilisateur entre plusieurs alternatives mais, en procédant ainsi, on augmente la complexité de l'interaction et de la prise en main du logiciel, ce qui va à l'encontre des souhaits de la majorité des utilisateurs qui visent une maîtrise facile et rapide de l'interaction avec un logiciel qu'ils découvrent. Même si le délai de réaction des boutons d'activation du défilement peut irriter l'utilisateur, la solution mise en œuvre actuellement pour contrôler le lancement, l'arrêt et la vitesse du défilement représente un compromis acceptable entre les différentes contraintes à satisfaire : les boutons ne masquent pas l'affichage, le contrôle du défilement s'effectue par des gestes simples donc faciles à apprendre qui exigent peu d'effort et de précision.

7.5.3 Dispositif de repérage spatial

Sur les 20 participants, 18 n'ont pas du tout utilisé l'indicateur de repérage spatial durant l'expérimentation ; ils l'ont simplement testé lors de la phase de prise en main pendant laquelle ils expérimentaient chaque fonctionnalité de l'interface. Les deux autres participants ont utilisé cet indicateur, mais rarement. La principale explication avancée par les participants au cours des entretiens de débriefing est que la fente verticale dont les vues 3D étaient dotées, leur suffisait pour se repérer lors de l'exploration des collections, compte tenu de la taille de celles-ci, de l'ordre d'un millier d'items. D'autres justifications ont été également présentées, comme la peur de perdre du temps dans les tâches de type T2 où la consigne donnée aux sujets imposait d'optimiser le temps de recherche (`S14`), ou bien l'inutilité de cet indicateur pour les tâches de type T1 qui portaient

sur des cibles dont la position dans la visualisation était inconnue *a priori* des participants, comme pour celles de type T2 en raison d'une bonne mémorisation de la position des cibles dans les collections (sujet `S13`).

Deux améliorations ont été proposées concernant l'indicateur de repérage spatial. La première est de le rendre actif, c'est-à-dire de donner à l'utilisateur la possibilité de déplacer la colonne rouge à la souris pour effectuer des déplacements de grande amplitude. La seconde proposition est de remplacer le cylindre miniature par des indicateurs visuels présents en permanence dans l'affichage ; cette proposition est analogue à la solution adoptée pour les deux études présentées dans les chapitres précédents. Regroupés, ces deux commentaires décrivent les fonctionnalités de la barre de défilement des interfaces graphiques actuelles, qui permet des déplacements rapides dans la représentation d'un ensemble d'informations et indique à chaque instant la position du contenu courant de la fenêtre d'affichage dans cette représentation. Ces propositions sont à étudier, même si nous avons jugé préférable de réduire au maximum la présence des interacteurs à l'écran. Dans une version plus évoluée de l'interface utilisateur, il serait envisageable d'ajouter des *widgets* contrôlables par la souris (clics et *drags*) ; mais cette évolution nécessite une évaluation ergonomique et une étude d'usage. A l'heure actuelle, compte tenu de la taille des collections utilisées, la présence d'un dispositif de ce type ne s'impose pas, comme le suggère le faible taux d'utilisation.

7.5.4 Eloignement-rapprochement des vues

Comme dans les études précédentes, nous avons laissé la possibilité à l'utilisateur de moduler la distance entre la caméra virtuelle et la représentation, ce qui a pour effet d'agir de manière conjointe sur le nombre de photographies affichées à l'écran et sur leur taille. Comme celles offertes par l'indicateur de repérage spatial, cette possibilité a été relativement peu exploitée en dehors de la phase de prise en main des visualisations. Deux propositions ont été formulées : d'une part, rendre possible l'utilisation de la molette pour contrôler ce mouvement (`S11` et `S14`) et, d'autre part, associer l'éloignement-rapprochement à une nouvelle disposition des items affichés pour éviter de « perdre » des photographies verticalement et horizontalement (`S15`).

Comme pour le contrôle de la rotation, l'utilisation de la roulette de la souris est la solution privilégiée proposée par les participants, ce qui est compréhensible vu son utilisation intensive par les logiciels grand public actuels, que ce soit pour faire défiler des pages web ou contrôler le niveau de zoom de documents visuels. Pourtant, il est préférable de restreindre son usage à des interactions spécifiques peu fréquentes, afin de limiter à la fois la fatigue musculaire (concernant le défilement) et le risque d'opérations accidentelles (concernant l'éloignement-rapprochement). D'autre part, un changement de disposition des photographies (c'est-à-dire du nombre de colonnes et de lignes) conduirait à une perte totale du contexte et des points de repère acquis durant le cheminement vers la cible. Un compromis envisageable serait de limiter ce type d'action aux premières interactions, afin de permettre à l'utilisateur d'adapter une fois pour toutes l'interface à ses capacités visuelles.

7.5.5 Zoom des photographies

Le dispositif de zoom n'a pas évolué par rapport à la première étude : le coefficient d'agrandissement est fixe, la photographie sélectionnée dans la visualisation est encadrée en rouge, et une courte animation fait le lien entre la photographie encadrée et son agrandissement qui vient se placer au centre de l'écran. Cette animation est destinée à faciliter la compréhension de l'action réalisée (*feedback* visuel). Sur les 20 participants, 16 n'ont pas formulé de commentaires particuliers sur l'implémentation du zoom ; ils ont jugé la taille de l'image agrandie suffisante et l'animation utile. 3 participants seulement (`S6`, `S11` et `S12`) auraient apprécié de pouvoir moduler le facteur d'agrandissement ; l'un d'eux a proposé l'usage de la molette de la souris pour le faire (`S12`). Enfin, un participant a trouvé l'animation trop longue (`S19`).

Il semble que la mise en œuvre du zoom soit parvenue maintenant à un stade où elle répond de façon satisfaisante aux besoins des utilisateurs potentiels. En particulier, elle est parfaitement adaptée aux tâches de recherche et d'exploration proposées dans T1 et T2. L'affichage en haute résolution pourrait éventuellement faire appel à un logiciel externe, dédié.

7.5.6 Discussion

Au delà de la représentation 3D, figée dès l'étude préliminaire, nous avons fait sans cesse évoluer l'interaction. A l'heure actuelle, nous avons atteint, nous semble-t-il, le meilleur compromis possible pour le contrôle des déplacements (*i.e.*, défilement latéral et éloignement-rapprochement). Si les boutons de défilement latéral gagneraient à être plus « réactifs », la solution retenue pour cette étude nécessite une précision moindre que les précédentes, la présence des interacteurs est discrète : ils ne masquent pas la visualisation, disparaissent lorsqu'elle est mouvement et s'intègrent visuellement à elle par leur forme. Les boutons permettant l'éloignement-rapprochement sont quand à eux toujours présents à l'écran, mais peu visibles, ce qui est approprié à la faible fréquence de leur utilisation. Enfin, l'indicateur de repérage spatial est maintenant affiché à la demande.

Les fonctionnalités de base que nous avions sélectionnées, et qui correspondent aux besoins fondamentaux de la navigation et de la recherche dans des représentations visuelles de collections de photographies, sont maintenant dans une version que l'on peut considérer comme aboutie. L'interface utilisateur qui les regroupe peut donc être proposée conjointement avec les deux vues de la représentation cylindrique aux concepteurs de logiciels orientés grand public. Certaines fonctionnalités complémentaires (*e.g.*, dispositif de repérage spatial actif) et améliorations de fonctionnalités existantes (*e.g.*, modulation du niveau de zoom) peuvent être implémentées en fonction du public visé et de ses besoins spécifiques.

7.6 Conclusions générales

Cette étude avait pour but de comparer les deux vues de la représentation cylindrique que nous proposons pour visualiser des collections de documents visuels, à une représentation 2D classique où les documents sont disposés sous la forme d'une matrice à deux dimensions. Cette

représentation matricielle est aujourd'hui la plus utilisée par les logiciels grand public, et même par les systèmes d'exploitation, comme l'explorateur de fichiers de *Windows*. Si la large diffusion de ces visualisations 2D interactives facilite leur prise en main immédiate par l'utilisateur *lambda*, ce type de représentation n'est pas exempt de défaut. D'une part, la non-continuité de la représentation impose aux utilisateurs des mouvements supplémentaires lorsqu'ils souhaitent revenir de la « fin » au « début » de la visualisation. D'autre part, l'uniformité de la taille des images/documents affichés simultanément implique un faible guidage du regard en raison de l'absence de zone(s) saillante(s). Ces deux points faibles, pris en compte conjointement, sont de nature à limiter le confort visuel et l'efficacité des recherches. Les deux implémentations du cylindre vertical considérées se proposent de remédier à ces défauts, en favorisant de plus la mémorisation spatiale de la position des photographies dans les collections dont elles font partie.

Afin de réaliser une comparaison objective des deux vues 3D considérées avec une représentation 2D « classique », nous devons contrôler tous les paramètres d'interaction et d'affichage. C'est pourquoi nous n'avons pas utilisé un logiciel de visualisation 2D existant (commercial ou non), mais implémenté une représentation dont les caractéristiques sont semblables à notre représentation cylindrique, à l'exception des affichages qui reproduisent les visualisations matricielles classiques. En particulier, nous nous sommes attachés à garantir une interaction homme-machine identique avec les trois vues testées, qui soit à la fois adaptée à la troisième dimension et paraisse familière, naturelle, aux utilisateurs expérimentés d'ordinateurs. Le défilement latéral induit par la rotation du cylindre étant parfaitement compatible avec le défilement horizontal fréquemment implémenté dans les interfaces classiques, nous avons choisi une interaction simple basée sur des boutons sensitifs pour l'activation du défilement latéral des visualisations. Les seules différences entre les représentations 3D et 2D résident dans la non-continuité de la représentation 2D, et dans le nombre légèrement différent de photographies affichées simultanément. L'objectif de cette seconde différence est que les photographies entrant dans la composition des affichages des trois vues soient de même taille et qu'elles couvrent l'écran de la même manière dans l'affichage initial. Les fonctionnalités d'interaction sont les mêmes que dans la première étude, seules la position, la visibilité et les modes d'activation des différents interacteurs ont été revus. Les tâches de recherche sont celles décrites dans le quatrième chapitre. 20 participants, 10 hommes et 10 femmes, d'une moyenne d'âge de 24 ans et d'un niveau de connaissances générales et informatiques similaire, ont pris part à cette expérimentation, en réalisant 6 recherches de cibles pour chaque type de tâche (T1 et T2) et chaque vue, soit 18 cibles au total et 36 tâches de recherche visuelle. Durant T2, nous avons équipé les participants d'un *eye-tracker* afin, grâce à l'enregistrement de leurs mouvements oculaires, de pouvoir analyser finement leurs stratégies d'exploration visuelle, construire une taxonomie des comportements d'exploration et de recherche visuelles, et étudier l'influence éventuelle des deux métaphores d'interaction mises en œuvre (locomotion, manipulation) sur la qualité de la mémorisation spatiale.

Les hypothèses de travail que cette étude visait à tester, au nombre de quatre, se résument ainsi :

- A. Les vues 3D, en raison de leur géométrie, apportent un meilleur confort visuel et permettent une meilleure efficacité de la recherche et de l'exploration visuelles que les vues 2D.
- B. Les vues 3D canalisent le regard vers les zones de la visualisation où les miniatures

sont agrandies et/ou non déformées. La vue externe, où les images agrandies sont non déformées, permet un meilleur guidage des mouvements oculaires que la vue interne ; elle conduit à des parcours visuels plus courts et plus systématiques.
- C. La vue interne, qui implémente la métaphore de la locomotion, peut améliorer le succès des tâches de recherche visuelle faisant appel à la mémoire spatiale.
- D. Les préférences des utilisateurs sont influencées principalement par les caractéristiques géométriques des vues 3D.

Les résultats de cette étude sont synthétisés dans la suite de cette section qui se termine par une conclusion générale.

7.6.1 Synthèse des résultats

Sur les 20 participants, 14 donnent leur préférence à une des deux vues 3D (ou aux deux vues, pour un sujet). Une très grande majorité de participants de sexe masculin donnent leur préférence aux vues 3D (9 sur 10), tandis que 9 femmes sur 10 préfèrent la vue externe ou la vue 2D. Cette dernière observation tend à corroborer les résultats de l'étude précédente concernant la distribution des préférences des participants. D'après les commentaires exprimés dans les questionnaires finaux et les entretiens de débriefing, les participants justifient essentiellement leurs préférences par les caractéristiques géométriques spécifiques des vues qui contribuent, selon eux, au confort visuel de l'exploration des affichages et/ou à l'efficacité de la recherche des cibles. La vue 2D est ainsi préférée en raison de l'absence de déformations des photographies et de son caractère familier, habituel. Les participants préférant cette vue critiquent la vue interne pour la faible taille des images des colonnes centrales, et apprécient leurs grandes dimensions dans la vue externe. La meilleure vue panoramique procurée par la vue interne est appréciée par ceux qui la préfèrent aux deux autres. Quant aux partisans de la vue externe, ils justifient leur préférence par la plus grande saillance des photographies de la colonne centrale et critiquent l'uniformité de la présentation 2D des photographies et le fait que celles-ci couvrent la totalité de l'écran. Ces observations vérifient l'hypothèse D.

Les performances des participants varient peu d'une vue à l'autre. Pour T1, seul le nombre de recherches visuelles réussies diffère significativement entre les vues 3D et la vue 2D, au bénéfice de la vue 2D. Pour T2, seule la vitesse de défilement varie significativement d'une vue à l'autre, et là encore en faveur de la vue 2D. Cependant, le nombre limité de cibles (6 par vue et par type de tâche) nous interdit de conclure sur ces deux points. A noter que la continuité des vues 3D est un avantage certain pour l'identification des cibles situées à la fin de la représentation dans la vue 2D ; elle minimise le temps de recherche par rapport à la vue 2D. Ce résultat est une confirmation partielle de l'hypothèse A.

Comme pour l'étude précédente, nous avons identifié deux types de stratégies d'exploration visuelle différentes lors du défilement latéral des vues : les participants focalisent leur attention visuelle, soit sur les colonnes entrantes (*i.e.*, celles qui apparaissent à l'écran), soit sur les colonnes centrales. 13 des 17 participants dont la stratégie a pu être identifiée gardent un comportement constant d'une vue à l'autre ; la majorité d'entre eux (10 sur les 13) appliquent une stratégie de type « colonnes entrantes ». Ce résultat qualitatif indique que le caractère dynamique des visualisations (défilement latéral) exerce une influence prépondérante sur les comportements visuels par

rapport à leurs caractéristiques statiques (nature et position des déformations dues à la perspective). Avec la vue 2D, 13 sujets sur 18 gardent la stratégie « colonnes entrantes ». Les stratégies d'exploration visuelle identifiées sont donc conformes à celles observées dans le cadre de l'étude précédente. L'observation des mouvements oculaires permet en outre de valider l'hypothèse B : la vue 2D occasionne de nombreuses saccades horizontales sur la largeur de l'écran, contrairement aux deux vues 3D qui favorisent un balayage vertical régulier des colonnes centrales (pour la vue externe) ou entrantes (pour la vue interne). Quatre types de « comportements » visuels ont été observés. Ils fournissent des informations à la fois sur le confort visuel de l'examen de l'affichage et sur la qualité du repérage spatial de la cible : le regard peut se focaliser sur une colonne entière, sur quelques photographies d'une même colonne ou sur une seule photographie qu'il suit souvent horizontalement. Les deux derniers comportements indiquent une mémorisation de la position verticale de la cible dans la visualisation. Le quatrième comportement correspond à de grandes saccades sur toute la surface de l'écran ; il est plus fréquent avec la vue 2D qu'avec les autres vues. Si aucune des vues n'a d'influence significative sur la mémorisation de la position verticale des cibles, les vues 3D sont supérieures à la vue 2D pour la mémorisation de la position horizontale des cibles (notamment la vue externe). Ces résultats confirment partiellement l'hypothèse C, car nous attendions sur ce point une supériorité de la vue interne sur la vue externe plutôt que le contraire. La métaphore de la locomotion suggérée par la vue interne ne joue pas le rôle que nous avions escompté, ce qui est confirmé par le faible nombre de sujets ayant remarqué la caractère immersif de cette vue.

L'analyse des données oculométriques, principalement l'étude de l'amplitude des saccades et de la durée moyenne des fixations, nous permet de confirmer en totalité l'hypothèse A. En effet, la vue interne permet effectivement de réduire l'amplitude des saccades, améliorant ainsi le confort visuel des activités de recherche et d'exploration visuelles. Sur l'ensemble de T2, on peut observer que la vue interne et, à un degré moindre, la vue 2D favorisent des saccades d'amplitude nettement moindre que la vue externe. Ce résultat peut s'expliquer par l'influence prédominante du caractère dynamique des vues par rapport à leurs caractéristiques géométriques sur les stratégies de parcours visuel, et de l'adéquation entre les caractéristiques statiques et dynamiques pour la vue interne.

Concernant les vues interne et externe, les résultats de cette étude et ceux de la première étude, décrits dans le chapitre précédent, sont souvent en accord, ce qui renforce leur robustesse et accroît leur portée. Ainsi, concernant les performances, les deux études montrent que la vitesse et la longueur de parcours sont plus importantes avec VI qu'avec VE pendant T2. En revanche, on observe un plus grand nombre de participants qui accordent leur préférence à VE dans cette étude que dans la précédente ; en outre, la mémorisation spatiale semble supérieure avec VE alors que, dans l'étude précédente, les indices disponibles étaient en faveur de VI. Ces différences tiennent vraisemblablement à plusieurs facteurs dont la modification de l'interacteur de manipulation des vues qui ne masque plus, dans cette étude, les photographies en partie basse de l'écran dans VE, la possibilité offerte aux participants de changer l'affectation par défaut des boutons de défilement latéral aux sens de rotation possibles, ainsi que le caractère plus ou moins immersif de la vue interne en fonction de la taille de l'écran.

Enfin, concernant l'interface utilisateur, nous sommes parvenus à un compromis satisfaisant

entre les différentes contraintes que doit satisfaire l'implémentation des fonctionnalités d'interaction. Certains points de détail sont encore à améliorer, comme la réactivité des boutons de défilement latéral, mais ces modifications sont mineures. La solution adoptée pour l'activation du défilement paraît satisfaisante, car elle est simple et efficace ; en outre, elle requiert un geste de faible précision et peu fatigant, et ne nécessite aucun effort d'apprentissage particulier.

7.6.2 Conclusion

Cette étude nous a permis de démontrer la supériorité, en termes de confort visuel, des vues 3D sur les représentations 2D classiques couramment utilisées par les logiciels de gestion de documents visuels. Concernant les performances, on n'observe aucune différence statistiquement significative entre les trois vues, sauf en ce qui concerne le nombre de recherches visuelles réussies pendant T1 et la vitesse de défilement pendant T2, ces deux différences étant en faveur de la vue 2D. Compte tenu du manque d'expérience des participants avec les vues 3D par rapport à la vue 2D, ces résultats sont prometteurs. Une étude portant sur un usage prolongé des trois vues serait nécessaire pour déterminer avec certitude l'influence exacte de chacune d'entre elles sur l'efficacité de la recherche et de l'exploration visuelles de collections de documents visuels. Cette seconde étude fournit en outre des conclusions utiles pour faire évoluer la conception des visualisations interactives de collections de documents visuels destinées au grand public. S'il est peut-être prématuré d'envisager la mise en œuvre des vues interne et externe dans les logiciels de gestion de documents visuels actuels, les moyens d'interaction conçus et développés dans le cadre de cette études sont applicables dès maintenant aux interfaces utilisateur des visualisations 2D matricielles existantes, puisqu'ils ont été utilisés pour interagir avec la vue 2D et sont le résultat d'une démarche de conception itérative impliquant les participants aux trois études réalisées et présentées dans ce mémoire. Concernant les vues 3D, les trois études effectuées démontrent la nécessité d'implémenter à la fois la vue externe et la vue interne, en raison de la diversité des préférences, des performances et des stratégies visuelles individuelles. D'autres études sont nécessaires pour déterminer, notamment, la contribution de la troisième dimension à l'efficacité de la recherche visuelle et, surtout, son soutien, effectif ou non, à la mémoire spatiale.

Chapitre 8

Conclusions et perspectives

La croissance rapide de la capacité de stockage informatique d'informations se ralentit aujourd'hui. Actuellement, la grande majorité des utilisateurs disposent de suffisamment d'espace pour stocker plusieurs centaines de giga-octets de données multimédia comprenant de nombreux documents visuels. Parmi ces derniers, les photographies sont les plus fréquentes dans les espaces de stockage personnels. Si le volume n'est plus un facteur limitant, retrouver un document particulier devient une tâche critique, dont le succès n'est pas assuré.

Produits majoritairement par le grand-public, ces documents visuels sont faiblement annotés. L'ajout de méta-données est une activité à la fois fastidieuse et particulièrement difficile pour ce type de document car le langage se prête mal à la caractérisation d'informations visuelles qui sont de nature sémiotique, donc autre que sémantique. Pour cette raison, les outils d'indexation existants, principalement linguistiques, sont peu adaptés à la recherche de documents visuels, et les techniques de recherche par le contenu sont encore embryonnaires pour ce type de document. Par conséquent, la recherche visuelle est une étape nécessaire importante du processus de fouille dans ces données faiblement structurées. Aujourd'hui, le recours, pour la visualisation d'images, à des représentations matricielles que l'on peut survoler à l'aide d'ascenseurs horizontal et/ou vertical, est une pratique quasi systématique. Pourtant, de récentes études montrent que ce type de représentation est loin de garantir un confort visuel optimal. En tirant parti des progrès récents des technologies informatiques, il devient possible d'imaginer et de concevoir des représentations visuelles de grandes collections d'images non structurées qui soient plus efficaces que les visualisations matricielles classiques tout en implémentant des métaphores d'interaction évidentes et simples.

Le but de cette thèse était de proposer une représentation de collections non structurées de documents visuels, au cas particulier, des photographies, qui facilite la recherche visuelle de photographies familières visuellement ou vérifiant des critères exprimables verbalement. Plus précisément, l'objectif était triple : accroître le confort visuel, réduire le temps de recherche et favoriser la mémorisation des positions des photographies dans l'espace de représentation, par rapport aux logiciels actuels de gestion de collections d'images. A partir d'une représentation 3D d'une collection sous la forme d'un cylindre vertical, nous avons conçu deux visualisations correspondant à deux vues en perspective du cylindre qui suggèrent intuitivement deux métaphores

d'interaction différentes : une vue intérieure du cylindre favorisant la métaphore de la locomotion, et une vue extérieure évoquant plutôt la métaphore de la manipulation. Les documents visuels sont plaqués sur les faces interne ou externe du cylindre vertical. Au delà des métaphores d'interaction, les deux vues se distinguent par les déformations géométriques différentes induites par les règles de la perspective.

L'objectif principal était d'évaluer expérimentalement dans quelle mesure ces deux vues sont à même d'offrir un meilleur confort visuel que les représentations 2D classiques, tout en garantissant efficacité et simplicité d'usage. Nous souhaitions également déterminer l'impact éventuel de la troisième dimension sur la mémoire spatiale. Le programme de recherche, organisé en trois étapes de complexité croissante, a pris la forme de trois études expérimentales successives, s'appuyant sur une méthodologie semblable : la réalisation par les participants de deux tâches de recherche visuelle : la première consiste à rechercher une photographie particulière non familière visuellement, à partir d'une description textuelle suffisamment précise pour éviter toute confusion avec d'autres photographies de la collection, la seconde est la recherche d'une photographie familière visuellement dont la position dans la représentation est connue. Ces deux situations, parmi les plus communément rencontrées en recherche visuelle d'informations, permettent d'évaluer les vues 3D à la fois dans des activités de d'exploration et de navigation.

La conduite d'une étude préliminaire auprès de 8 sujets nous a permis d'évaluer l'implémentation des deux métaphores d'interaction, locomotion et manipulation, ainsi que leur pertinence pour la recherche de documents visuels, tout en estimant l'intérêt d'utilisateurs potentiels pour la troisième dimension et ces deux métaphores d'interaction. Si les performances se sont révélées similaires et les préférences partagées, la qualité du repérage spatial a semblé meilleure dans la vue interne. Le zoom de la vue interne a cristallisé de nombreux commentaires, car jugé trop lent et peu pratique. Le parti pris du respect des métaphores d'interaction au détriment de la stricte équivalence des vues en termes d'interaction, s'est révélé partiellement inadapté, principalement à cause de contraintes matérielles. Cette première étude nous a permis de valider le protocole expérimental, tout en confirmant l'intérêt de nos représentations ; elle a en outre guidé les choix de conception de la seconde étude, en particulier le choix de moyens identiques d'interaction avec les deux vues.

L'étude suivante, menée cette fois auprès de 20 sujets, avait pour but de conforter et de raffiner les résultats de l'étude préliminaire. Le protocole expérimental était identique, mais l'interaction a été modifiée et les performances d'affichage notablement améliorées. L'équivalence rigoureuse de l'interaction avec les deux vues a permis la réalisation d'une étude statistique sur les performances des participants. L'objectif principal de cette étude était d'identifier la « meilleure » vue, mais nous souhaitions également mettre en évidence les stratégies d'exploration visuelle des participants et leur lien éventuel avec les propriétés géométriques spécifiques des vues. Aucune des deux vues n'est vraiment « meilleure » que l'autre en termes de performances, pourtant les participants ont accordé majoritairement leur préférence à la vue interne, pour des raisons liées à ses caractéristiques géométriques. Deux stratégies d'exploration visuelle ont pu être mises en évidence : l'attention visuelle se focalise soit sur les colonnes entrantes, soit sur les colonnes centrales des vues. C'est le défilement latéral, plus que les caractéristiques géométriques statiques des vues, qui semble influencer de façon privilégiée le choix des participants en faveur de

l'une ou l'autre de ces stratégies. Enfin, la mémorisation de la position des cibles est légèrement meilleure dans la vue interne, en conformité avec nos hypothèses.

Si ces résultats sont en faveur de la vue interne, ils interdisent cependant d'exclure de notre champ d'étude la vue externe dans la mesure où celle-ci a recueilli les suffrages d'au moins 20 % des participants, tout en permettant des performances similaires à celles obtenues avec la vue interne. Dans la dernière des trois études présentées, nous avons donc comparé l'utilisation de ces deux vues à à celle d'une représentation 2D classique, en faisant tester par 20 participants les trois vues. Cette dernière étude nous a permis de mieux comprendre l'apport de la troisième dimension grâce à l'analyse de l'utilisation des deux vues en perspective de la représentation cylindrique. Afin d'évaluer de manière quantitative et qualitative le confort visuel, nous avons équipé les participants d'un oculomètre à optique de tête, ce qui a permis d'analyser leurs mouvements oculaires (saccades et fixations) pendant la réalisation des tâches de recherche de photographies qui leur étaient familières et dont la position dans la collection courante leur était connue. Cette étude est la pierre angulaire de notre travail, car elle a montré que la représentation cylindrique est suceptible de remplacer, avec profit, les représentations matricielles classiques. Une large majorité des particpants accordent leur préférence à l'une des vues 3D plutôt qu'à la vue 2D et, comme dans l'étude précédente, ils justifient cette préférence par les caractéristiques géométriques spécifiques de la vue 3D choisie. Concernant les performances, on n'observe aucune différence statistiquement significative entre la vue 2D et les deux vues 3D, sauf pour le nombre de succès de recherches de cibles inconnues (premier type de tâche) et la vitesse moyenne de défilement pendant la recherche de cibles familières (deuxième type de tâche) ; pour ces deux mesures, la vue 2D s'avère nettement supérieure aux vues 3D (résultats statistiquement significatifs). On peut raisonnablement supposer qu'avec une expérience de l'utilisation des vues 3D comparable à celle qu'ils possédaient de l'utilistion des affichages graphiques matriciels, les particopants auraient obtenu, pour les trois vues, des résultats similaires sur la totalité des mesures de performances. Concernant le confort visuel, les données oculométriques recueillies pendant les tâches de recherche de cibles familières montrent que la vue interne offre un confort comparable à celui de la vue 2D et nettement supérieur à celui de la vue externe, concernant l'amplitude des saccades et la durée des fixations (résultats statistiquement significatifs). Quant à l'analyse des parcours oculaires, elle montre une dispersion supérieure des points de fixation à la surface de l'écran avec la vue 2D ; c'est la vue interne qui canalise et guide le plus fortement le regard. Cette analyse met également en évidence l'existence de quatre comportements d'exploration visuelle différents lors du défilement des vues. Enfin, la comparaison des résultats de cette étude avec ceux de la précédente conforte la plupart d'entre eux, qu'ils portent sur les préférences, sur les performances ou sur les stratégies de parcours visuel.

Les hypothèses de travail que les trois études réalisées tentent de vérifier sont rappelées brièvement ci-après :

– **A**. Le recours à la troisième dimension accroît, par rapport à l'usage de visualisations 2D classiques, le confort et l'efficacité des activités de recherche et d'exploration visuelles, en canalisant et guidant davantage le regard vers les zones de la visualisation où les miniatures sont agrandies et/ou non déformées ; un parcours visuel plus systématique est de nature à réduire, d'une part, les échecs (par omission) de détection des cibles et, d'autre part, la longueur des parcours visuels.

- **B**. Si les vues 3D canalisent le regard vers les zones de la visualisation où les miniatures sont agrandies et/ou non déformées (hypothèse A), la vue externe où les images agrandies sont non déformées permet un meilleur guidage/contrôle des mouvements oculaires que la vue interne ; elle conduit à des parcours visuels plus courts et plus systématiques.
- **C**. La vue interne, qui implémente la métaphore de la locomotion, peut améliorer davantage, par rapport à la vue externe et aux vues 2D, le succès des tâches de recherche visuelle faisant appel à la mémoire spatiale.
- **D**. Les vues 3D sont susceptibles de recueillir majoritairement les préférences des utilisateurs, par rapport aux visualisations 2D classiques.

Les résultats majeurs obtenus sont synthétisés dans la suite du chapitre, et leur contribution à la validation ou à l'invalidation des hypothèses ci-dessus précisée.

Les préférences : attractivité des vues 3D Dès l'étude préliminaire, les participants ont jugé positivement l'apport de la troisième dimension dans les tâches de recherche et d'exploration visuelles de grandes collections non structurées de documents visuels. Les études suivantes confirment ces jugements subjectifs. Les préférences de la majorité des participants à chacune des trois études se partagent entre les deux vues 3D. La préférence pour la vue interne est justifiée par la meilleure vue globale qu'elle propose, qui facilite les tâches d'exploration. La plus grande taille des images sur les colonnes latérales, ainsi que les distorsions plus progressives sont également mises en avant très souvent. La vue externe est, quant à elle, préférée pour la plus grande taille et, de ce fait, la plus grande saillance visuelle, des photographies au centre de l'écran. Les deux principales justifications avancées en faveur de la vue 2D, par la minorité de participants qui l'ont choisie, sont son caractère familier et l'absence totale de déformation.

Le fait que les vues 3D soient plus appréciées que les vues 2D classiques donne déjà une première légitimité à la représentation que nous proposons, et contribue à valider l'hypothèse D. Globalement, les jugements subjectifs tendent à confirmer partiellement les hypothèses A et B, puisque les participants justifient exclusivement leur préférence pour l'une des vues 3D en mentionnant une, voire plusieurs, des caractéristiques géométriques spécifiques de celle-ci. En revanche, l'une des principales justifications de la préférence pour la vue 2D est son caractère familier, habituel. Cet argument relève d'un comportement, minoritaire ou non suivant les cas, qui caractérise les réactions de certains usagers vis-à-vis d'une innovation quelconque. Le temps alloué à l'utilisation des vues dans les trois études présentées était insuffisant pour neutraliser, chez tous les participants, les réactions négatives que les vues 3D pouvaient susciter en raison de leur caractère innovant, qui rompait avec les habitudes des participants ; ceux-ci, en effet, étaient tous des utilisateurs expérimentés de logiciels de gestion de documents visuels. Néanmoins, la simplicité d'utilisation de la représentation 3D proposée, le caractère intuitif de l'interface utilisateur des deux visualisations qui la mettent en oeuvre sont suffisants pour que de telles réticences disparaissent rapidement avec la pratique.

Stratégies d'exploration et confort visuel Un des apports majeurs de la dernière étude concerne l'évaluation du confort visuel de chacune des trois vues grâce aux « outils » d'observation et d'analyse des mouvements oculaires mis en oeuvre. Les données oculométriques ont

permis d'identifier quatre « comportements » visuels différents qui s'inscrivent dans deux stratégies principales d'exploration visuelle, mises en évidente de façon convergente par les trois études : l'attention visuelle des participants se porte, soit sur les colonnes centrales, soit sur les colonnes latérales lors du défilement des vues. La prédominance de cette dernière stratégie dénote une influence prédominante du caractère dynamique des vues par rapport à leurs caractéristiques géométriques, le regard étant naturellement « attiré » vers la zone de l'écran où apparaissent les photographies. L'analyse des saccades et des fixations montre que, dans ce contexte dynamique, l'examen des photographies s'effectue principalement par un balayage vertical. En outre, d'après l'analyse qualitative des mouvements oculaires, la vue interne est celle des trois vues qui limite le plus la dispersion des fixations sur la surface de l'écran, canalise et guide le plus fortement les mouvements oculaires et donc, permet le balayage le plus régulier et le plus systématique des colonnes « entrantes ». Ce résultat s'explique par l'adéquation entre la géométrie de la vue interne et le type de défilement choisi (*i.e.*, latéral) : du fait de la concavité de la vue, d'une part, et du renouvellement latéral des photographies, d'autre part, le regard est naturellement canalisé vers les colonnes entrantes, là où les images sont les plus grandes et où les déformations engendrées par la perspective sont moins gênantes que dans la vue externe. En outre, la vue interne et la vue 2D favorisent, par rapport à la vue externe, des saccades de plus faible amplitude et des fixations de durée moindre lors de l'examen des miniatures (résultats statistiquement significatifs), ce qui contribue au confort visuel en réduisant la fatigue induite par les mouvements oculaires et le maintien du regard sur une zone très réduite de l'affichage pendant les fixations. Par conséquent, si l'on tient compte de la dispersion des fixations, de nature à accroître la longueur des parcours visuels, de l'amplitude des saccades et de la durée des fixations, la vue interne apparaît comme celle qui assure le meilleur confort visuel pendant la recherche des cibles, suivie par la vue 2D, qui supplante de loin la vue externe. Les hypothèses A et B doivent donc être nuancées pour tenir compte des résultats médiocres obtenus avec la vue externe, contrairement à nos attentes, ainsi que des très bons résultats obtenus avec la vue 2D, contrairement à nos attentes également. Les deux hypothèses A et B tenaient compte exclusivement des propriétés géométriques des vues et ignoraient leur caractère dynamique qui favorise une stratégie d'exploration visuelle centrée sur les colonnes entrantes plutôt que sur les colonnes centrales ; or, dans la vue externe, les miniatures des colonnes entrantes sont de taille très réduite et leur contenu très déformé ; l'inexactitude des hypothèses A et B concernant la comparaison entre les vues 3D interne et externe résulte de cette erreur. L'expérience antérieure des participants, leur familiarité avec les vues 2D matricielles peut expliquer les résultats qu'ils ont obtenus avec cette vue et qui sont en désaccord avec l'hypothèse A. Cependant, on peut supposer qu'avec davantage de pratique, la vue interne pourrait offrir un meilleur confort que les vues 2D classiques quant aux activités de recherche visuelle.

Efficacité de la recherche Les valeurs des variables destinées à mesurer l'efficacité de la recherche, telles le temps moyen de recherche des cibles, la longueur moyenne de parcours de la collection ou le nombre de passages devant la cible recherchée, varient très peu d'une vue à l'autre, contrairement à nos attentes. La variabilité intra- et inter-individuelle des performances s'avère très importante en raison, notamment, de la diversité des stratégies de parcours visuel des participants aux trois études. Compte tenu de la nouveauté des vues 3D proposées par

rapport aux visualisations 2D matricielles utilisées généralement, ces résultats, même s'ils peuvent sembler décevants au premier abord, sont, en fait, plutôt prometteurs. En effet, les participants à la dernière étude ont réalisé avec les deux vues 3D qu'ils ont expérimentées brièvement, des performances comparables à celles obtenues avec la vue 2D dont ils ont une pratique et une expérience importantes.

Mémorisation spatiale Compte tenu de la disposition en grille des items dans les trois vues, l'analyse de la mémorisation de la position des cibles lors de la réalisation des tâches de type T2 peut être décomposée suivant deux axes, horizontal et vertical. Concernant la mémorisation de la position verticale des cibles, seule la dernière étude fournit des renseignements précis grâce à l'analyse qualitative des données oculométriques. Pour les deux premières études, nous disposons uniquement des témoignages des participants qui sont nombreux à affirmer, lors des débriefings, qu'ils se sont souvenus de la position verticale de la majorité des cibles, notamment celles situées en bas ou en haut de l'écran. Dans la dernière étude, l'analyse des parcours oculaires montre la supériorité de la vue externe par rapport aux deux autres vues : la hauteur de la cible dans l'affichage a été mémorisée en moyenne, avec plus ou moins de précision, pour plus de la moitié des cibles avec VE et pour 42 % d'entre elles avec VI, les résultats obtenus avec la vue 2D se situant entre les deux. Concernant la position horizontale des cibles, les résultats, qui portent uniquement sur VE et VI, varient d'une étude à l'autre : dans les deux premières études, VI semble favoriser davantage que VE la mémorisation de la position horizontale des cibles, alors que ce résultat s'inverse dans la dernière étude où VE se montre supérieure à VI. La mesure principale utilisée mesurer la mémorisation de la position horizontale des cibles est le choix du sens de rotation, optimal ou non, adopté pour atteindre la cible. Or, cette mesure n'est pas applicable à la vue 2D qui autorise un seul sens de défilement à partir du début de la collection. En outre, pour les vues 3D, la moitié au moins des participants à chacune des trois études adoptent un sens de rotation constant, quelle que soit la position de la cible, pour au moins une des deux vues 3D. Le faible nombre de participants pour lesquels l'analyse de la mémorisation de la position des cibles explique vraisemblablement les variations des résultats observées d'une étude à l'autre, et interdit toute conclusion sur l'influence des vues et des métaphores d'interaction qu'elles sous-tendent sur la mémoire spatiale. Pour établir avec certitude la validité de l'hypothèse C, il est nécessaire de concevoir un protocole expérimental qui permette de mesurer avec précision, pour tous les participants, l'influence des trois vues interactives sur la mémorisation de la position horizontale des cibles, dans le contexte de tâches réalistes comme celles de T2, et sans interférer avec la réalisation de ces tâches, c'est-à-dire sans inciter les participants à faire un effort volontaire de mémorisation. A noter qu'il n'est pas certain que l'hypothèse C s'avère pertinente dans la mesure où, dans les trois études, un quart environ des participants seulement évoquent dans leurs commentaires écrits ou oraux le caractère immersif de la vue interne. En outre, dans lse trois études, rares sont les participants qui utilisent, même occasionnellement, l'indicateur de repérage spatial, que celui-ci soit présent en permanence ou affichable à la demande.

Perspectives Cette recherche expérimentale ouvre de nombreuses perspectives. La comparaison entre les trois vues s'est avérée difficile, du fait du faible nombre de cibles utilisées dans l'expérimentation et de la grande variabilité inter-individuelle des performances ; en particulier,

les résultats variaient sensiblement, pour un même participant, d'une une vue 3D à l'autre, et d'un type de tâche à l'autre, vraisemblablement en fonction, principalement, de ses capacités visuelles. Le choix du nombre de tâches était un compromis raisonnable entre la durée de la passation et le volume de données quantitatives et qualitatives, nécessaire pour être en mesure d'établir des conclusions de portée générale. Pour déterminer avec précision l'influence éventuelle de la troisième dimension sur les performances, en particulier sur l'efficacité de la recherche visuelle d'images familières, il nous semble nécessaire d'effectuer une étude longitudinale qui permette d'acquérir des informations sur l'évolution des performances avec la pratique, la portée de la dernière étude expérimentale comparative que nous avons réalisée étant limitée par la familiarité des participants avec la vue 2D et leur inexpérience de l'utilisation des vues 3D, ainsi que par les variations des performances individuelles avec les deux vues 3D. Cette étude pourrait se dérouler en deux étapes. D'abord, une première expérimentation portant exclusivement sur les vues interne et externe du cylindre vertical permettrait de déterminer, pour chaque participant, la vue 3D avec laquelle il obtient les meilleures performances. Ensuite, une seconde expérimentation faisant intervenir les mêmes participants serait centrée sur la comparaison entre leurs performances avec la vue 3D sélectionnée lors de la première étape et celles obtenues avec une vue 2D classique. Une étude d'usage, qui nécessiterait l'utilisation prolongée de prototypes opérationnels, par un échantillon représentatif d'utilisateurs, offrirait également des informations précieuses sur l'attractivité et l'efficacité réelles de chacune des trois vues.

Quant à l'apport de la troisième dimension à la mémoire spatiale par rapport aux visualisations 2D, il reste à démontrer. C'est une entreprise difficile car, pour progresser sur cette question de recherche, il est nécessaire de concevoir une nouvelle méthodologie. En effet, les protocoles expérimentaux existants s'avèrent inappropriés pour l'étude, dans le contexte d'activités réalistes, donc complexes, de recherche visuelle et de navigation dans des collections d'images non structurées.

Enfin, nous envisageons d'étudier l'apport de la troisième dimension à la navigation dans des visualisations d'informations structurées hiérarchiquement. Si de nombreux travaux se sont intéressés à cette problématique, peu ont proposé des représentations facilement compréhensibles par le grand public ; et la plupart des visualisations sont en deux dimensions. Nous réfléchissons actuellement à la manière dont la troisième dimension pourrait être utilisée pour faciliter la compréhension des cartes arborescentes [Shneiderman, 1992] et les rendre accessibles au grand public. Notre intuition est que l'utilisation de cette dimension supplémentaire pourrait simplifier la compréhension des relations entre les niveaux successifs d'une hiérarchie d'informations. Nous envisageons de vérifier cette intuition prochainement.

La diffusion immédiate des deux vues 3D auprès du grand public est peut-être prématurée. Il est nécessaire auparavant, nous semble-t-il, de conforter davantage, parmi les résultats obtenus, ceux qui tendent à montrer l'influence des stratégies individuelles d'exploration visuelle et, à un degré moindre, celle des caractéristiques géométriques de chaque vue, sur les préférences et les performances. Il faudrait, en outre, étudier l'influence des capacités visuelles sur les préférences et les performances concernant les deux vues 3D et la vue matricielle classique, en expérimentant ces trois vues avec des sujets présentant des déficiences visuelles variées ; jusqu'à présent, tous les participants impliqués avaient une vue normale monoculaire et binoculaire. En revanche,

nous pouvons envisager, dès maintenant, de proposer aux concepteurs de logiciels de gestion de documents visuels à l'intention du grand public, l'interface utilisateur que nous avons conçue et mise au point pour les trois vues de façon itérative et avec la participation d'une cinquantaine d'utilisateurs futurs.

Annexe A

Etude préliminaire

L'annexe A contient, pour la vue interne et externe du cylindre, l'ordre de présentation des cibles dans les tâches 1 (T1) et 2 (T2) ainsi que les descriptions utilisées dans T1. Sont également présentés les questionnaires initial et final.



(a) Vue intérieure - tâche 1



(b) Vue intérieure - tâche 2



(c) Vue extérieure - tâche 1



(d) Vue extérieure - tâche 2

Fig. A.1 – Ordre de présentation des cibles

Cible 1

Une mini Cooper rouge à toit blanc effectue un virage sur une route en pente qui tourne vers la gauche.
La voiture avance face à l'objectif. Deux personnes sont visibles à l'avant.
Les bâtiments en arrière-plan "penchent" légèrement vers la droite en raison de l'angle de prise de vue.

En haut à droite un peu de ciel d'un bleu très pâle.

Cible 2

Un vestiaire avec, à gauche, partant du premier plan, quatre rangées de casiers bleu ciel superposés fixés au mur.

Les murs sont beige clair, le sol carrelé d'un beige plus foncé. Sur le mur du fond face à l'objectif, une porte d'ascenseur gris aluminium.

A droite, au second et à l'arrière plans, trois personnes de dos, deux femmes et un homme.
La femme au second plan à droite a de longs cheveux blonds. Elle porte un haut bleu turquoise et un jean bleu.

Cible 3

Vue aérienne d'un îlot de sable clair avec, au centre, un monticule couvert d'une végétation vert foncé. Au sommet du monticule, un palmier isolé se détache sur le sable en arrière-plan.

La mer transparente laisse voir les fonds marins. Elle est d'un bleu outremer derrière l'île, zébrée à gauche d'une traînée d'écume blanche. Elle est turquoise au premier plan.

Cible 4

Paon vu de face en gros plan faisant la roue.

Le paon est sur une pelouse.

Cible 5

Quatre enfants, trois filles et un garçon déguisés pour Halloween sont dans une pièce debout côte à côte, face à l'objectif.
Trois des enfants portent des chapeaux pointus noirs à large bord et des vestes ou des manteaux noirs.
Seul, le deuxième enfant à partir de la gauche, une fille aux longs cheveux châtains, est tête nue. Elle porte une jupe rouge et un pull-over noir sous un manteau noir ouvert.

Cible 6

A l'intérieur d'un cratère volcanique rocheux, un petit lac circulaire bleu turquoise ensoleillé. Les pentes extérieures du cratère sont sombres. La paroi interne du cratère est sombre à gauche où elle est à l'ombre, et de couleur claire, jaunâtre, à droite où le soleil l'éclaire.

A gauche, le cratère se détache sur des sommets enneigés et, à droite, sur quelques cumulus blancs. Au-dessus, une large bande de ciel limpide, bleu pâle à gauche, d'un bleu plus intense à droite.

FIG. A.2 – Descriptions des cibles utilisées dans la vue interne (VI)

Cible 1

A gauche, la tête d'un poteau télégraphique vertical se détache en noir sur un ciel nocturne bleuté, illuminé à droite par trois minces éclairs.

Du poteau partent des câbles fins : trois d'entre eux barrent la photo horizontalement, les deux autres, perpendiculaires aux précédents, la traversent en oblique, de l'angle supérieur gauche vers le bas.

Les éclairs partent de la droite, l'un en dessous des câbles horizontaux, parallèle à ceux-ci, les deux autres au-dessus, obliques, en direction de la tête du poteau.

Cible 2

Gros plan sur la tête, vue de face, d'un bébé phoque couché sur la glace.

Sa fourrure beige s'éclaircit sur les joues, autour du museau et au-dessus de ses yeux noirs que des zones sombres sur leurs bords font paraître énormes.
Les ombres sur la glace sont d'un bleu-vert pâle.

Cible 3

Six planètes se détachent sur un ciel noir. Celle au centre est bleue.

A gauche, au premier plan, la plus grosse des planètes, de couleur brunâtre, avec quelques tâches blanchâtres. La surface de l'hémisphère supérieur, éclairée, présente un léger relief irrégulier. L'hémisphère inférieur disparaît dans l'ombre.

Au second plan, au centre, une planète bleu, plus petite, de surface lisse.
A droite de cette planète, une autre beaucoup plus petite semble collée à elle au niveau de l'équateur. Elle est brune avec une surface crevassée.

Au-dessus de ces trois planètes, trois autres, minuscules, de taille croissante de gauche à droite.

Cible 4

Un large chemin de planches brun clair s'enfonce dans un sous-bois. Le feuillage peu fourni laisse voir, les troncs sombres et élancés des arbres à contre-jour , et entrevoir le ciel blanchâtre.

De chaque côté du chemin, une bordure en bois.
Sur ses bords et sur le sol, quelques feuilles mortes rousses.

Cible 5

Deux touaregs, suivis par un touriste, marchent l'un derrière l'autre sur l'arête d'une dune.
Les touaregs sont vêtus et enturbannés de bleu. Le touriste, habillé de jaune et coiffé de bleu, porte une planche de surf bleue. La dune est brun clair.

Les marcheurs, de profil, avancent de la droite vers la gauche. L'arête de la dune monte de la droite vers la gauche.
En arrière-plan, le ciel est bleu clair, limpide.

Cible 6

Portrait d'un homme jeune, coiffé d'un chapeau de paille clair à larges bords et portant une barbe noire qui lui masque le bas du visage.
L'homme, la tête légèrement tournée vers la droite, se détache sur une tenture à larges rayures obliques jaunes et blanches.

Il porte une chemise violette et une veste à revers bleu turquoise.

FIG. A.3 – Descriptions des cibles utilisées dans la vue externe (VE)

Nous vous invitons à participer à l'évaluation (robustesse, précision et facilité d'utilisation) de deux types de représentations 3D de grands ensembles d'informations visuelles (collections de photographies au cas particulier) : une vue extérieure et une vue intérieure d'un cylindre.

Cette expérimentation préservera votre anonymat. Les informations sollicitées dans ce questionnaire nous serviront simplement à préciser votre profil en tant qu'utilisateur de logiciels grand public (connaissances générales et expérience informatique). Elles resteront confidentielles.

QUESTIONNAIRE INITIAL

Nom :	Numéro :
Prénom :	Tél :
Age :	Email :

CONNAISSANCES GENERALES

Dernier diplôme obtenu (spécialité et année d'obtention) ou diplôme en cours de préparation (année) :

.....

LATERALISATION

Etes-vous :

Gaucher (gauchère) ◯
Droitier (droitière) ◯

COMPETENCES INFOMATIQUES

Quels types d'équipements informatiques avez-vous déjà utilisés ?

PC ou MAC ☐
Portable ☐
PDA ☐
TabletPC ☐
Autres - Précisez : .....

Quels systèmes d'exploitation vous sont familiers ?

Microsoft Windows ☐
MacOS ☐
Linux, Unix ☐

Quelles activités pratiquez vous ?

	Jamais	Parfois	Souvent
Internet	◯	◯	◯
Utilisation de logiciels grand public :			
Tableur	◯	◯	◯
Traitement de texte	◯	◯	◯
Utilisation de logiciels graphiques :			
Dessin	◯	◯	◯
PAO	◯	◯	◯
CAO/DAO	◯	◯	◯
Programmation, conception et développement d'applications	◯	◯	◯
Jeux	◯	◯	◯

Autres - Précisez : .....

FIG. A.4 – Questionnaire intial - page 1

Quels dispositifs et techniques d'interaction avez-vous déjà utilisés ?

	Jamais	Parfois	Souvent
Clavier	○	○	○
Souris	○	○	○
Ecran tactile	○	○	○
Commande vocale	○	○	○
Tablette graphique	○	○	○
Joystick	○	○	○
Joypad	○	○	○

Si vous avez une expérience de la programmation, donnez la liste des langages de programmation et/ou des environnements de développement que vous utilisez :

.....

Etes-vous utlisateur d'outils de conception graphique ?

Oui ○
Non ○

Si oui, quels sont les outils graphiques que vous utilisez ?

Photoshop, GIMP, ... ☐
Paintshop Pro ☐
Toolkit graphique (PJA, Zinc, Swing, ...) ☐
Autres - Précisez : .....

Avez-vous déjà réalisé des pages web ?

Oui ○
Non ○

Si oui, quels types d'outils ou environnements de développement avez-vous utilisés ?

	Jamais	Parfois	Souvent
Langage HTML	○	○	○
Editeur de pages web (Namo, Dreamweaver, ...)	○	○	○
Langages d'écriture de scripts (PHP, javascript, ...)	○	○	○
Outils de création d'animations (Flash, ...)	○	○	○

Lorsque vous abordez un nouveau logiciel, vous utilisez :

	Jamais	Parfois	Souvent
Une approche par essai-erreur	○	○	○
L'aide en ligne associée au logiciel	○	○	○
Un manuel/notice d'utilisation papier	○	○	○
Un ouvrage de référence	○	○	○
Des exemples	○	○	○
L'assistance d'un utilisateur expérimenté	○	○	○

Envoyer par messagerie

FIG. A.5 – Questionnaire intial - page 2

QUESTIONNAIRE FINAL

Pour simplifier, on désignera, dans le reste du document, la représentation extérieure (resp. intérieure) du cylindre par l'expression "vue externe" (resp. "interne").

- Quelle vue, externe ou interne, vous a semblé la plus facile à utiliser ?

Pourquoi ?

- Quelle vue, externe ou interne, vous a semblé la plus confortable, la moins fatigante ?

Pourquoi ?

- Quelle vue, externe ou interne, vous a semblé la plus efficace, concernant :
  - La facilité et la vitesse de parcours ou d'analyse des photographies de la collection ?

Pourquoi ?

  - La facilité et la précision du repérage de la position courante de l'affichage dans la collection ?

Pourquoi ?

- Le dispositif permettant le zoom vous a-t-il paru :
  - Efficace

Pourquoi ?

  - Agréable et simple à utiliser

Pourquoi ?

FIG. A.6 – Questionnaire final - page 1

- Le dispositif d'interaction associé à chacune des vues vous a-t-il paru approprié ?

Pourquoi ?

- L'affichage mural vous a-t-il paru adapté :
  - Aux représentations des collections choisies (i.e., les vues) ?

    Pourquoi ?

  - Aux deux types de tâches proposés ?

    Pourquoi ?

- Améliorations souhaitées, concernant :
  - La vue externe

  - La vue interne

  - Les dispositifs d'interaction avec chacune des deux vues

  - Le dispositif de zoom

FIG. A.7 – Questionnaire final - page 2

- La qualité des images

- La fluidité et la qualité de l'affichage

- Difficulté des tâches :
  - Photos non familières visuellement [1]

    Très faciles faciles sans difficulté particulière difficiles très difficiles infaisables

  - Photos familières visuellement [1]

    Très faciles faciles sans difficulté particulière difficiles très difficiles infaisables

- Problèmes et incidents :
  Les résumer brièvement en précisant éventuellement la tâche concernée.

- Autres commentaires :

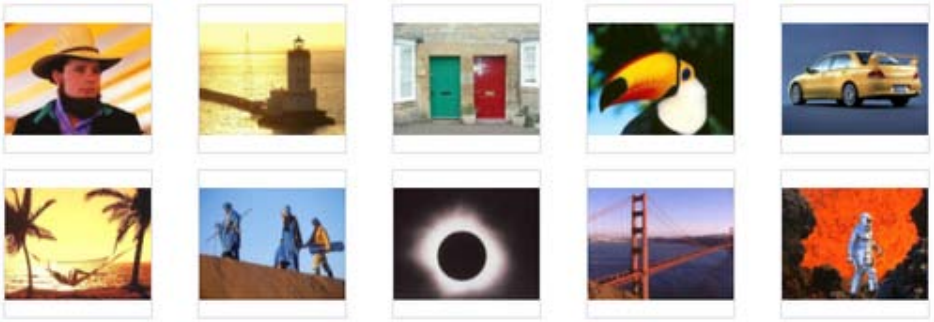
[1] Entourer le qualificatif qui exprime le mieux votre jugement.

FIG. A.8 – Questionnaire final - page 3

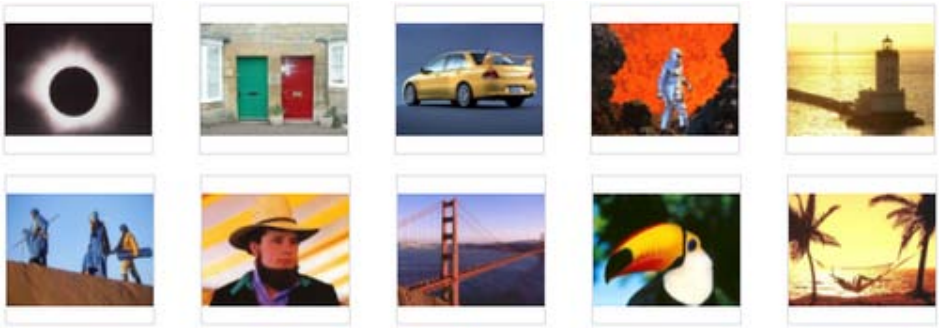
Annexe B

Premiere Etude

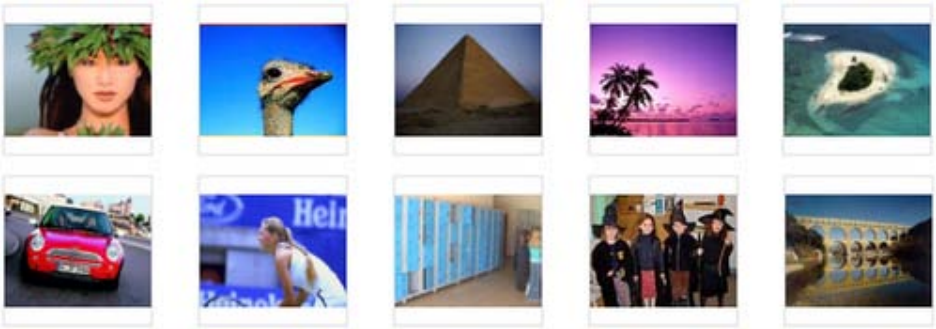
L'annexe B contient, pour la vue interne et externe du cylindre, l'ordre de présentation des cibles dans les tâches 1 (T1) et 2 (T2) ainsi que les descriptions utilisées dans T1. Sont également présentés les instructions et consignes données aux sujets.



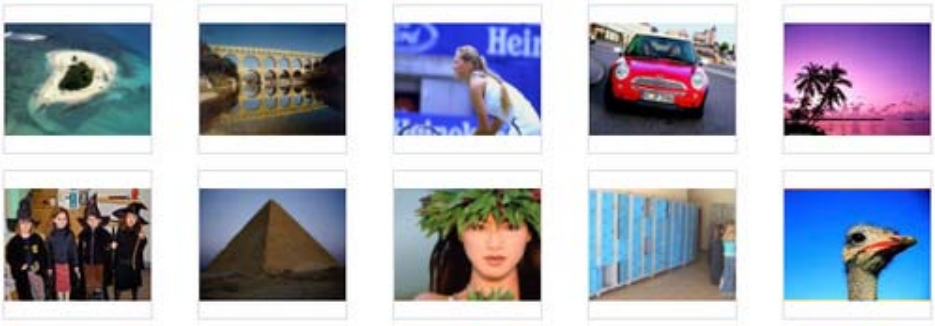
(a) Vue intérieure - tâche 1



(b) Vue intérieure - tâche 2



(c) Vue extérieure - tâche 1



(d) Vue extérieure - tâche 2

FIG. B.1 – Ordre de présentation des cibles

VUE INTERIEURE

Cible 1

Un homme avec une barbe noire, coiffé d'un chapeau de paille claire

La photographie est un portrait ; l'homme a la tête légèrement tournée vers la droite.

L'arrière-plan est constitué d'une tenture à larges rayures obliques jaunes et blanches.

L'homme porte une chemise violette et une veste noire à revers bleu turquoise.

Cible 2

Un phare, se détachant à contre-jour sur un coucher de soleil doré

Le phare est légèrement sur la droite de la photographie dont il en occupe presque toute la hauteur.

L'arrière plan est constitué d'une bande de ciel jaune doré dans le quart supérieur de la photographie. En dessous, la mer, qui prend la couleur du coucher de soleil est traversée par une fine bande de terre qui descend en oblique du bord gauche de l'image jusqu'au pied du phare.

Cible 3

Deux portes de maison, l'une verte et l'autre rouge

Les deux portes sont vues de faces, la porte verte à gauche et la porte rouge à droite. On voit partiellement les avancées de deux « bow windows » situées l'une à gauche, l'autre à droite des deux portes.

La façade des maisons est en pierres grises.

Cible 4

La tête d'un toucan avec un énorme bec jaune orangé

La photographie est un gros-plan sur la tête d'un toucan. Celui-ci a la tête tournée vers la gauche.

L'arrière plan est constitué de feuillages très sombres au-dessus et à gauche de la tête du toucan. En dessous du bec du toucan, on aperçoit le ciel, de couleur bleue.

Le cou du toucan est blanc, tandis que le haut de son corps est noir.

Cible 5

Une voiture de sport jaune

La voiture est vue de trois quarts arrière gauche.

L'arrière-plan est composé d'un fond gris-bleu, s'éclaircissant du bas vers le haut de la photographie.

La voiture est une Mitsubishi Lancer Evo VIII : c'est une berline tri-corps (4 portes), avec un imposant aileron sur le coffre.

FIG. B.2 – Descriptions des cibles utilisées dans la vue interne (VI) - page 1

Cible 6

Un hamac tendu entre deux palmiers, se détachant sur un coucher de soleil doré

Les palmiers sont de part et d'autre de la photographie, à contre-jour.

L'arrière-plan est composé d'un ciel jaune doré, qui occupe les deux tiers supérieurs de la photographie. En dessous, on peut voir la mer avec le reflet du coucher de soleil. En bas, au premier plan, une bande de terre, de couleur sombre.

Le hamac est occupé par une personne, qui a une jambe légèrement pliée.

Cible 7

Deux touaregs en bleu suivis par un homme en jaune sur l'arête d'une dune

Les marcheurs, de profil, avancent de la droite vers la gauche. L'arête monte de la droite vers la gauche.

L'arrière-plan est composé d'un ciel bleu, sans nuage, qui occupe les deux tiers supérieurs de la photographie.

L'homme en jaune porte une planche de surf bleue. Le sable de la dune est brun.

Cible 8

Une éclipse de soleil

La photographie est un gros-plan sur l'éclipse.

L'arrière plan est noir, et on peut voir un halo lumineux blanc autour de la planète qui masque le soleil.

La planète est noire.

Cible 9

Le "Golden Gate Bridge"

Au premier plan, à gauche, une des piles du pont. La seconde pile est visible au loin, très petite, contre le bord droit de la photographie.

L'arrière-plan est composé d'un ciel bleu sur la moitié supérieure de la photographie. Celui-ci devient blanc-rose sur la ligne d'horizon. Sur la moitié inférieure de la photographie, le bras de mer bleu sombre que le pont enjambe.

Le pont est de couleur rouge-rose.

Cible 10

Une personne en combinaison ignifugée devant de la lave en fusion

La personne est au centre de la photographie, elle est vue de profil, et se dirige vers la droite.

En arrière-plan, de la lave en fusion, de couleur rouge-orangée. De part et d'autre de la personne, ainsi que sur le sol, des rochers très sombres.

L'homme porte une combinaison intégrale de couleur argentée.

FIG. B.3 – Descriptions des cibles utilisées dans la vue interne (VI) - page 2

VUE EXTERIEURE

Cible 1

Une femme aux cheveux noirs, coiffée d'une couronne de feuillage vert

La photographie est un portrait ; la femme fait face à l'objectif, son visage occupe les deux tiers droits de la photographie.

L'arrière-plan, à gauche de la tête de la femme, est un ciel bleuté.

Cible 2

La tête d'une autruche

La photographie est un gros-plan sur la tête d'une autruche, vue de profil. L'autruche regarde vers la droite.

L'arrière plan est composé d'un ciel bleu uni, s'assombrissant légèrement vers le haut de la photographie.

La tête de l'autruche est de couleur gris-beige. Sous l'œil, une traînée rouge-orangée, en direction du bec.

Cible 3

Une pyramide d'Egypte

La photographie est un gros-plan sur une pyramide de couleur brunâtre. Deux faces de la pyramide sont visibles, l'une éclairé, l'autre dans l'ombre. L'arête les séparant est au centre de l'écran.

L'arrière-plan est composé d'un ciel d'un bleu sombre, occupant les deux tiers de la hauteur de la photographie. Le sol, irrégulier, est brun.

Cible 4

Un bouquet de palmiers se détachant sur un ciel rose violacé

Des palmiers, à contre-jour, sont visibles sur la gauche de la photographie, dans le coin inférieur. Ils penchent vers la droite.

L'arrière plan est composé d'un ciel violet devenant rose violacé au voisinage de la ligne d'horizon. Le ciel occupe les quatre cinquièmes supérieurs de la photographie. Dans la partie inférieure, une mer calme, sur laquelle se reflète le ciel.

Derrière les palmiers, le ciel s'éclaircit pour devenir presque blanc.

Cible 5

Vue aérienne d'un îlot

L'îlot est au centre de la photographie.

Il est couvert de végétation vert-foncé et bordé de plages de sable. La mer qui l'entoure est de couleur bleu outremer à l'arrière-plan, et bleu turquoise au premier plan.

Un palmier isolé, minuscule de part la distance de prise de vue, se détache au-dessus de la végétation.

FIG. B.4 – Descriptions des cibles utilisées dans la vue externe (VE) - page 1

Cible 6

Une Mini Cooper rouge

La photographie est un gros-plan sur la voiture, qui fait face à l'objectif. Celle-ci effectue un virage sur une route tournant vers la gauche.

L'arrière-plan est composé de bâtiments, qui penchent vers la droite, en raison de l'angle de prise de vue. En haut à droite, un peu de ciel bleu est visible.

On aperçoit deux personnes à l'avant de la voiture dont le toit et les rétroviseurs sont blancs.

Cible 7

Une joueuse de tennis, habillée en blanc

La joueuse regarde vers le côté gauche de la photographie. Elle est vue de profil et est penchée vers l'avant.

L'arrière-plan est bleu : des inscriptions sont partiellement visibles en haut à droite et en bas à gauche.

La joueuse a de longs cheveux blonds, formant une queue de cheval nattée.

Cible 8

Un vestiaire avec des casiers bleus

A gauche, partant du premier plan, quatre rangées de casiers bleus ciel superposés, fixés au mur. A droite, trois personnes de dos, deux femmes et un homme.

L'arrière-plan est composé des murs de couleur beige. Le sol est carrelé, également de couleur beige.

La femme à droite a de longs cheveux blonds, elle porte un haut bleu turquoise et un jean bleu.

Cible 9

Quatre enfants déguisés pour Halloween

Les quatre enfants sont dans une pièce, debout côte à côte, face à l'objectif.

Trois des enfants portent des chapeaux pointus noirs à large bord et des manteaux noirs. Le deuxième enfant en partant de la gauche est tête nue et porte une jupe rouge.

A l'arrière-plan, des étagères et une commode en bois sont visibles.

Cible 10

Un aqueduc à deux niveaux

La vue a été prise d'un emplacement plus proche de l'extrémité gauche de l'aqueduc que de la droite, et l'ancrage de ses extrémités dans deux massifs boisés sombres occupent toute la largeur de la photographie.

L'arrière-plan est composé d'un ciel bleu sur la moitié supérieure de la photographie, tandis que la rivière enjambée par l'aqueduc occupe la partie inférieure. Les arches des deux niveaux de l'aqueduc se reflètent dans l'eau qui s'assombrit progressivement vers l'angle inférieur droit de la photographie.

L'aqueduc est en pierres de couleur gris-beige.

FIG. B.5 – Descriptions des cibles utilisées dans la vue externe (VE) - page 2

LES DEUX INTERFACES : MODE D'EMPLOI

Instructions générales

1. Le grand nombre d'images dans les collections induit parfois un temps de chargement un peu long entre deux tâches, soyez patient.

2. Pour signifier au système que vous avez trouvé l'image, vous la « zoomez ». Une fois l'image agrandie au centre de l'écran, vous pouvez cliquer sur le bouton *Suivant*  pour valider votre choix et poursuivre l'expérience.

   Tant que vous n'avez pas cliqué sur le bouton *Suivant* , vous pouvez modifier votre choix. C'est le clic sur *Suivant* qui déclenche son enregistrement.

3. Lorsque vous cliquez sur le bouton *Suivant* , le passage à la tâche suivante est irréversible. Aucun message ne vous invite à confirmer votre intention. Assurez-vous, avant de cliquer sur ce bouton, que vous avez réellement terminé la tâche courante.

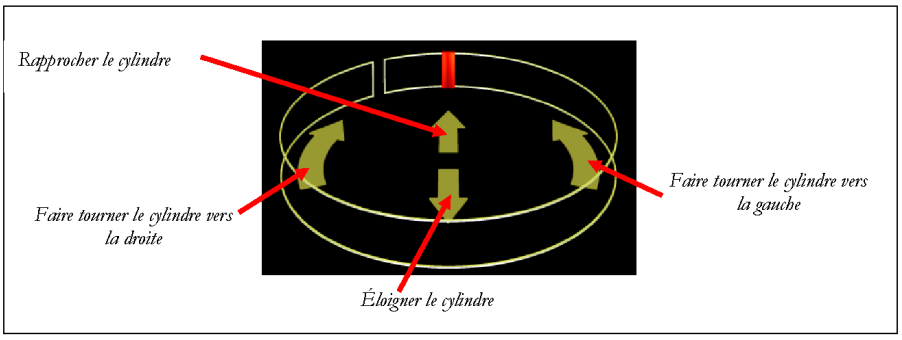
4. Dans la tâche de recherche des images décrite verbalement, une animation vous montre l'image solution. Pendant cette animation, on ne peut interagir avec les objets. A la fin de cette animation, l'image solution est agrandie au centre de l'écran, il suffit de cliquer sur celle-ci pour passer à la tâche suivante. Cette animation présente toujours le sens optimal de rotation optimal afin d'atteindre la solution en minimisant la distance.

FIG. B.6 – Instructions - page 1

Vue intérieure du cylindre

Déplacements

La commande permettant d'effectuer les déplacements se situe au centre de l'écran :



- Pour que le déplacement débute, il suffit de cliquer avec le bouton gauche sur une des flèches.
- Lorsque le cylindre tourne autour de vous, et seulement dans ce cas, vous avez la possibilité de moduler la vitesse de défilement, via la roulette de la souris.
- Pour arrêter le déplacement, cliquez simplement sur le bouton droit de la souris.

Remarque : pour passer d'un déplacement un autre, il faut d'abord arrêter le déplacement en cours.

Zoom

- Pour agrandir une image particulière, il suffit de cliquer avec le bouton gauche de la souris sur cette image.
- Pour faire disparaître l'image agrandie au centre de l'écran, il suffit de cliquer sur cette dernière.

Remarques : on ne peut agrandir plusieurs images en même temps ; il faut faire disparaître l'image agrandie courante avant de pouvoir « zoomer » une autre image. Si un déplacement est en cours, le fait de « zoomer » une image provoque l'arrêt du déplacement.

Repérage de la position

La barre rouge précise la position, sur le cylindre, de la colonne d'images au centre de l'écran ; elle est fixe. La bande transparente, représentant le cylindre, est mobile. Elle tourne sur elle-même en fonction des déplacements (rotations) que vous effectuez.

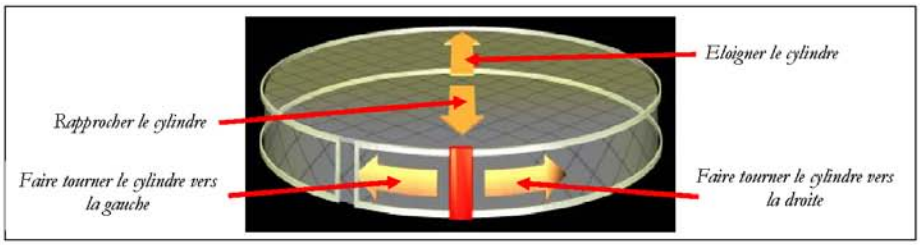
Remarque : cet outil de positionnement n'est pas déplaçable, et on ne peut interagir avec lui, il est donc inutile de le sélectionner ou de cliquer dessus.

FIG. B.7 – Instructions - page 2

Vue extérieure du cylindre

Déplacements

La commande permettant d'effectuer les déplacements se situe au centre de l'écran. Elle est couplée avec le dispositif permettant le repérage dans la collection :



- Lorsque qu'aucun déplacement, n'est en cours, le dispositif de repérage est actif :
  - le fait de passer la souris sur l'enveloppe du cylindre provoque l'apparition des flèches permettant la rotation.
  - le fait de passer la souris sur le haut du cylindre provoque l'apparition des flèches permettant de l'éloigner ou de le rapprocher.

  Pour que le déplacement débute, il suffit de cliquer avec le bouton gauche sur une des flèches.
- Lorsque le cylindre devant de vous, et seulement dans ce cas, vous avez la possibilité de moduler la vitesse de défilement, via la roulette de la souris.
- Pour arrêter le déplacement, cliquez simplement sur le bouton droit de la souris.

Remarques : pour passer d'un déplacement un autre, il faut d'abord arrêter le déplacement en cours. Lorsqu'un déplacement est en cours, l'enveloppe et le haut du cylindre disparaisse pour ne pas gêner la vision des photographies.

Zoom

- Pour agrandir une image particulière, il suffit de cliquer avec le bouton gauche de la souris sur cette image.
- Pour faire disparaître l'image agrandie au centre de l'écran, il suffit de cliquer sur cette dernière

Remarque : on ne peut agrandir plusieurs images en même temps ; il faut faire disparaître l'image agrandie courante avant de pouvoir « zoomer » une autre image.

Repérage de la position

La barre rouge précise la position, sur le cylindre, de la colonne d'images au centre de l'écran ; elle est fixe. La bande transparente, représentant le cylindre, est mobile. Elle tourne sur elle-même en fonction des déplacements (rotations) que vous effectuez.

FIG. B.8 – Instructions - page 3

EXPERIMENTATION

Evaluation d'une interface 3D de navigation dans des collections d'images

Objectif

Dans le cadre d'un projet de recherche soutenu par le Fonds National pour la Science, l'équipe MErLIn a conçu et implémenté deux modes de navigation et de recherche d'informations dans des ensembles de photographies non classées.

Un millier d'images environ, des photos couleur, sont disposées sur un cylindre vertical. Pour parcourir cet ensemble à la recherche d'une image particulière, vous utiliserez alternativement deux vues 3D du cylindre :

- l'une est une vue intérieure : vous êtes dans le cylindre et vous pouvez le faire tourner autour de vous, le rapprocher ou l'éloigner pour l'ajuster à votre vue ;
- l'autre est une vue extérieure : vous pouvez manipuler le cylindre, le faire tourner devant vous, le rapprocher ou l'éloigner pour l'ajuster à votre vue.

Les deux représentations permettent de "zoomer" une image particulière.

Après avoir testé successivement ces deux représentations interactives, vous évaluerez, en tant qu'utilisateurs potentiels, l'efficacité et la facilité d'utilisation de chacun d'entre eux pour naviguer dans une collection d'images. Les tâches proposées porteront :

- soit sur la recherche d'une image que l'on n'a jamais vue mais dont on connaît les caractéristiques, grâce à une description verbale,
- soit sur celle d'une image que l'on a déjà eu l'occasion de regarder, donc familière visuellement.

Par exemple, votre frère vous a offert un CD où il a gravé, sans prendre le temps de les classer, une sélection des photos qu'il a prises lors d'un séjour en Californie. Il vous a raconté qu'il avait réussi à photographier un pélican brun en train de nourrir ses petits.
Première situation : il n'a pas eu le temps de vous montrer cette photo. Vous voulez la voir.
Seconde situation : il vous l'a montrée, mais trop rapidement. Vous voulez la revoir pour l'examiner en détail.

FIG. B.9 – Consignes - page 1

Expérimentation

L'expérimentation comprend deux étapes :

- d'abord, la recherche de 20 photos dans des collections différentes à partir de leur description textuelle ;
- puis la recherche de ces 20 photos devenues familières ; chaque recherche est précédée de l'affichage, pendant 3 secondes, de la photo à chercher.

Les 10 premières tâches de chaque étape seront effectuées en utilisant l'une des deux représentations du cylindre, vue extérieure/intérieure, les 10 autres en utilisant l'autre.

Privilégiez la précision (i.e., trouver la "bonne" photo) plutôt que la rapidité. Cependant, efforcez-vous d'être méthodique car la durée de chaque tâche est limitée à 3 minutes.

N.B. Dès que vous avez sélectionné une photo ou que la limite de temps est dépassée, le système visualise la photo solution (uniquement dans le cas où les photos à chercher sont décrites verbalement).

Descriptions des photos

Les descriptions comprennent généralement quatre parties.

La première précise le thème général de la photographie et décrit son aspect visuel global.

Les trois parties suivantes détaillent et complètent la première et permettent :

- soit de se construire une représentation plus fidèle de la photo afin de faciliter son identification,
- soit, plus rarement, de la distinguer d'autres photos avec lesquelles elle pourrait être confondue sans de telles informations (thème, couleurs, formes).

N.B. Il existe peu de risques de confusion entre la photo à chercher et d'autres photos de la collection affichée. La première partie de la description suffit souvent pour identifier la photo sans ambiguïté.

FIG. B.10 – Consignes - page 2

Déroulement (G1)

Questionnaire initial d'identification

Tests de vision

Présentation générale de l'expérience

Etape 1

Présentation de l'interface d'interaction avec la vue intérieure du cylindre

Entraînement (2 tâches)

Réalisation de 10 tâches de recherche (à partir de descriptions des images)

Présentation de l'interface d'interaction avec la vue extérieure du cylindre

Entraînement (2 tâches)

Réalisation de 10 tâches de recherche (à partir de descriptions des images)

Etape 2

Au cours de cette étape, la recherche porte sur des images familières affichées préalablement pendant 3 secondes

Réalisation de 10 tâches de recherche en utilisant la vue intérieure du cylindre

Réalisation de 10 tâches de recherche en utilisant la vue extérieure du cylindre

Questionnaire d'évaluation des deux interfaces

Entretien post-expérience

FIG. B.11 – Consignes - page 3

Annexe C

Seconde Etude

L'annexe C contient, pour les vues interne et externe du cylindre et pour la vue 2D, les descriptions utilisées dans T1, les instructions données aux sujets et les questionnaires initiaux et finaux.

Cible 1

La tête d'un toucan avec un énorme bec jaune orangé

1. La photographie est un gros-plan sur la tête d'un toucan. Celui-ci a la tête tournée vers la gauche.
2. Au-dessus et à gauche de la tête du toucan, l'arrière plan est constitué de feuillages flous et très sombres. En dessous de son bec, on aperçoit un peu de ciel bleu.
3. Le cou du toucan est blanc, tandis que le haut de son corps est noir.

Cible 2

3 flûtes à champagne, remplies de liquide vert, rouge et bleu

1. La flûte verte est située le long du bord gauche de la photographie. A droite de cette flûte et plus éloignée de l'objectif, une flûte rouge puis une flûte bleue.
2. L'arrière-plan est bleu pâle. Le plan sur lequel sont disposées les flûtes, de la même couleur que l'arrière-plan, mais brillant, reflète partiellement les flûtes. Le reflet est coupé par le bord inférieur de la photographie, légèrement au dessus des pieds des flûtes.

Cible 3

Un feu d'artifice vert, sur un ciel noir

1. Le feu d'artifice occupe les 2/3 droits de la photographie. Il est composé de 5 « gerbes » qui se superposent légèrement : une grande gerbe au dessus de 4 plus petites en bas de la photographie.
2. L'arrière-plan est un ciel noir.

Cible 4

Un sentier s'enfonçant dans une forêt de feuillus

1. Face à l'objectif, le sentier, de couleur gris-beige, occupe au premier plan toute la largeur du bas de la photographie. Il s'enfonce dans la forêt avant de disparaître au loin.
2. De part et d'autre du sentier, des buissons verts, et des arbres.
3. Au premier plan, à gauche du sentier, on distingue le tronc d'un gros arbre.

Cible 5

Un homme avec un maillot orange à manches bleues et un casque de vélo blanc

1. La photographie est un portrait. L'homme a la tête légèrement tournée vers la droite et inclinée vers le bas.
2. L'arrière-plan, clair, est flou.
3. L'homme porte des lunettes de vélo non teintées, à montures noires. Un motif rouge orne le haut de sa manche droite (à gauche sur la photographie).

Cible 6

Une balle de paille, ronde, vue à contre-jour sur un champ de blé fauché

1. Le champ de paille brunâtre occupe les 2/3 inférieurs de la photographie. La balle de paille est de couleur très sombre. Elle est posée sur la tranche et on ne voit qu'une de ses faces circulaires qui se détache sur l'arête du champ.
2. Le ciel, à l'arrière-plan, occupe le 1/3 supérieur de la photographie. Il est composé dans sa partie inférieure de nuages gris et blancs. Il est bleu dans le haut de la photo.
3. L'arête du champ monte légèrement de la gauche vers la droite.

FIG. C.1 – Descriptions des cibles utilisées dans la vue interne (VI)

Cible 1

Un arbre isolé sur le sommet d'une colline d'herbe verte

1. L'arbre, de couleur sombre, se détache à contre-jour sur l'arête de la colline. La colline, vallonnée, occupe les 2/3 inférieurs de la photographie. Elle monte en pente douce de la gauche vers la droite.
2. L'arrière-plan, derrière l'arbre, est constitué d'un ciel bleu très clair.

Cible 2

Un lézard vert, vu de profil

1. La photographie est un gros-plan sur la tête et le torse d'un lézard vert vif. Il est vu de profil et regarde vers la droite.
2. L'arrière plan, dans la partie supérieure de la photographie, est très sombre. Le lézard est allongé sur l'écorce d'un tronc horizontal.
3. Le dessous de la gueule du lézard est blanc-crème.

Cible 3

Un homme avec une barbe noire, coiffé d'un chapeau de paille claire

1. La photographie est un portrait ; l'homme a la tête légèrement tournée vers la droite.
2. L'arrière-plan est constitué d'une tenture à larges rayures obliques jaunes et blanches.
3. L'homme porte une chemise violette et une veste noire à revers bleu turquoise.

Cible 4

La photographie d'un canal, prise d'une passerelle qui l'enjambe

1. Au premier plan, on distingue le garde-corps métallique de la passerelle d'où est prise la photographie. Face à l'objectif, le canal occupe toute la largeur du bas de la photographie, au premier plan.
2. De part et d'autre du canal, des berges obliques, en béton. Sur les berges, on distingue de l'herbe et des arbustes. Le canal disparaît à l'horizon. Dans la moitié supérieure de la photographie, le ciel est d'un bleu délavé.
3. Des arbres sont visibles au niveau de la ligne d'horizon.

Cible 5

Un arc-en-ciel

1. L'arc-en-ciel part du bas droit de la photographie et s'élève en diagonale vers la partie supérieure gauche de la photographie.
2. L'arrière-plan est divisé en deux : dans le coin inférieur droit de la photographie des arbres à contre-jour très sombres. Un ciel nuageux occupe le reste de la photographie. Il est sombre dans le coin inférieur gauche, et s'éclaircit au niveau du coin supérieur gauche de la photographie.
3. L'arc-en-ciel est très lumineux dans sa partie inférieure, et s'estompe dans la partie supérieure de la photographie.

Cible 6

Une fraise rouge, en gros plan

1. La photographie est un gros-plan sur une fraise, au centre de la photographie. Elle est posée sur une assiette blanc-beige que l'on distingue mal. La queue verte est tournée vers la gauche, la pointe vers la droite.
3. A droite de la fraise, dans la moitié supérieure de la photo, on distingue son ombre.

FIG. C.2 – Descriptions des cibles utilisées dans la vue externe (VE)

Cible 1

Une femme aux cheveux noirs, coiffée d'une couronne de feuillage vert

1. La photographie est un portrait. La femme fait face à l'objectif, son visage occupe les deux tiers droits de la photographie.
2. L'arrière-plan, à gauche de la tête de la femme, est un ciel bleuté.

Cible 2

Un phare vert pâle, de forme conique.

1. Le phare est au centre de la photographie, légèrement sur la gauche. Il est construit sur un massif de béton gris clair d'où part, sur la droite, un ponton en béton.
2. Au premier plan, la mer gris-bleu occupe toute la largeur de la photographie. Dans les 2/3 supérieurs, l'arrière-plan est composé d'un ciel pâle. Entre la mer et le ciel, derrière le phare, la côte est constituée d'une bande de broussailles, de couleur vert sombre.
3. La côte s'abaisse de la gauche vers la droite.

Cible 3

Une route bitumée avec une ligne centrale jaune

1. Face à l'objectif, la route, de couleur gris foncé, occupe, au premier plan, toute la largeur du bas de la photographie, avant de disparaître au loin. La chaussée est divisée en deux par une bande continue jaune, qui se dédouble. Les bords sont matérialisés par des bandes continues blanches ; celle de droite est plus visible que celle de gauche.
2. De part et d'autre de la chaussée, de la terre sablonneuse beige clair. En arrière-plan, un ciel bleuté occupe la moitié supérieure de la photographie. A l'horizon, une chaîne montagneuse sombre barre horizontalement la photographie en son centre.
3. Au premier plan, on peut distinguer deux fines balises blanches, de part et d'autre de la route.

Cible 4

La tête d'un perroquet au plumage gris et blanc

1. La photographie est un gros-plan sur la tête d'un perroquet. Le perroquet est vu de profil et regarde vers la gauche.
2. L'arrière plan, flou, est constitué de feuillages verts.
3. L'œil, petit et rond, a un iris d'un jaune très clair. Le bec du perroquet est gris foncé. Le plumage gris s'éclaircit autour de l'œil et au dessus du bec.

Cible 5

Un panneau de signalisation rond, bleu, avec un pictogramme blanc représentant un vélo

1. La photographie est un gros-plan sur le panneau de signalisation rond et bleu. Le pictogramme représente un vélo, dont l'avant est dirigé vers la gauche.
2. Le panneau se détache sur une pelouse, au pied d'un mur de briques de couleur rouge sombre. On distingue, à gauche, l'arrière d'une voiture gris aluminium.
3. Le panneau est bordé d'un double liseré blanc. On distingue, en dessous du panneau, le poteau sur lequel il est fixé, de couleur grise.

Cible 6

6 planètes se détachent sur un ciel noir

1. A gauche, au premier plan, la plus grosse des planètes, de couleur brunâtre. Au second plan, légèrement à droite du centre de la photographie, une planète bleue plus petite et sur son flanc droit, une autre, brune, beaucoup plus petite. Au dessus de ces 3 planètes, 3 autres planètes, minuscules, dont 2 très blanches, sur la droite.
2. L'arrière-plan est un ciel de couleur noire.
3. La taille des 3 planètes en haut de l'image est croissante de gauche à droite.

FIG. C.3 – Descriptions des cibles utilisées dans la vue 2D (V2D)

LES TROIS INTERFACES : MODE D'EMPLOI

Instructions générales

1. Le grand nombre d'images dans les collections induit parfois un temps de chargement un peu long entre deux tâches de recherche, soyez patient.

2. Lorsque vous pensez avoir identifié l'image, il faut faire un clic gauche sur celle-ci, puis valider votre choix par un clic gauche sur le bouton vert en bas à gauche de l'écran. Le choix est alors enregistré par le système.

   Tant que vous n'avez pas cliqué sur le bouton vous pouvez modifier votre choix. C'est le clic sur ce bouton qui déclenche la validation de l'enregistrement.

   Lorsque vous cliquez sur le bouton le passage à la tâche de recherche suivante est irréversible. Aucun message ne vous invite à confirmer votre intention.

   Assurez-vous, avant de cliquer sur ce bouton, que vous avez réellement terminé la tâche courante.

3. A la suite de chaque recherche d'une image décrite verbalement, une animation vous montre le déroulement, depuis la position de départ, de la recherche de l'image solution. A la fin de cette animation, l'image solution est zoomée au centre de l'écran, il suffit de cliquer sur celle-ci pour passer à la tâche suivante.
   Dans le cas des vues en 3D, cette animation présente toujours le sens optimal de rotation, c'est-à-dire celui qui permet d'atteindre la solution en minimisant la distance par rapport à la position de départ. Dans le cas de la vue 2D, cette animation part toujours du bord gauche de la représentation.

FIG. C.4 – Instructions - page 1

Interaction avec les représentations

DÉPLACEMENTS

Deux types de déplacements sont possibles :

1. Défilement latéral / rotation

- Pour que le déplacement débute, il suffit de survoler un des boutons latéraux avec la souris. Une fois le déplacement en cours, les boutons disparaissent.
- La vitesse de défilement latéral peut être modulée en utilisant la roulette de la souris, avec 12 niveaux différents.
- Pour arrêter le déplacement, deux solutions sont possibles : soit faire un clic droit sur n'importe quelle partie de l'écran, soit « zoomer » une image (action expliquée dans la partie suivante).

2. Rapprochement / éloignement du cylindre

Vous avez la possibilité d'ajuster la distance entre la représentation et vous. Cela se fait au moyen du bouton situé en bas à droite (2D et 3D extérieure) ou en bas au centre (vue 3D intérieure) :



Lorsque vous passez la souris sur cette zone, cela fait apparaître le panel suivant :



- Pour rapprocher la représentation, faire un clic gauche sur le bouton « + ».
- Pour éloigner la représentation, faire un clic gauche sur le bouton « – ».

Lorsque la souris ne survole plus le panel, celui-ci disparaît.

Remarque : pour passer d'un déplacement à un autre, il faut d'abord arrêter le déplacement en cours, en effectuant un « clic » droit sur n'importe quelle partie de l'écran.

FIG. C.5 – Instructions - page 2

ZOOM

- Pour agrandir une photographie particulière, faire un clic gauche sur cette photographie.
- Pour agrandir une autre photographie, faire simplement un clic gauche sur la photographie choisie.
- Pour mettre fin au mode « zoom », faire un clic droit sur n'importe quelle partie de l'écran.

Remarque : si un déplacement est en cours, le fait de « zoomer » une image provoque l'arrêt du déplacement.

REPÉRAGE DE LA POSITION

A tout moment, vous pouvez connaître précisément la position de la colonne affichée au centre de l'écran. Il suffit de maintenir appuyé le bouton central de la souris pour faire apparaître une « mini représentation » de la collection de photos. Celle-ci n'est pas déplaçable, et on ne peut interagir avec elle.

La barre rouge de cette visualisation de la collection de photos représente la colonne couramment affichée au centre de l'écran.

AFFECTATION DES BOUTONS DE DEFILEMENT LATERAL / ROTATION

Au début de l'expérimentation, pendant la phase de prise en main, il est vous est donné la possibilité de changer l'affectation des boutons de défilements latéral / rotation, en fonction de ce qui vous semble être le plus naturel.

Ce choix est ensuite fixé pour toute l'expérimentation avec la représentation.

FIG. C.6 – Instructions - page 3

PRESENTATION DE L'EXPERIMENTATION

Nous vous invitons à participer à l'évaluation de trois types de représentations de grands ensembles d'informations visuelles (collections de photographies au cas particulier) :

- deux représentations 3D : une vue extérieure et une vue intérieure d'un cylindre vertical,
- une représentation 2D "classique".

L'expérimentation consistera en la recherche de photographies particulières au sein de la collection.

Dans un objectif scientifique, la trace de vos interactions sera enregistrée, ainsi que votre activité visuelle (utilisation d'un *eye-tracker*).

Cette expérimentation préservera votre anonymat conformément à la loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Les informations sollicitées dans ce questionnaire nous serviront à préciser votre profil en tant qu'utilisateur de logiciels grand public. Elles resteront confidentielles.

Informations passation

Numéro Date

Informations personnelles

Nom Tél

Prénom E-mail

Age Latéralisation ○ gaucher ○ droitier

Numéro de bureau (si Loria)

CONNAISSANCES GENERALES

Dernier diplôme obtenu

Diplôme en cours de préparation (année 2006-2007)

FIG. C.7 – Questionnaire initial - page 1

COMPETENCES INFORMATIQUES

Quels types d'équipements informatiques avez-vous déjà utilisés ?

- [ ] PC ou Mac
- [ ] Portable
- [ ] PDA
- [ ] Tablet PC

Autres, précisez :

Quels types de systèmes d'exploitation avez-vous déjà utilisés ?

- [ ] Microsoft Windows
- [ ] Mac Os
- [ ] Linux, Unix

Autres, précisez :

Quels dispositifs et techniques d'interaction avez-vous déjà utilisés ?

	Jamais	Parfois	Souvent
Clavier	○	○	○
Souris	○	○	○
Ecran tactile	○	○	○
Commande vocale	○	○	○
Tablette graphique	○	○	○
Joystick	○	○	○
Manette de jeu	○	○	○

Estimez la quantité de documents visuels (photographies, dessin, images) que vous conservez sur votre ordinateur

Est-ce que vous organisez/ structurez ces documents visuels ? ○ oui ○ non

Si oui, selon quels critères ?

Si non, pourquoi ?

Utilisez-vous des logiciels pour explorer ces documents visuels ? ○ oui ○ non

Si oui, le(s)quel(s) ?

Quelles activités informatiques utilisant la 3D pratiquez-vous ?

	Jamais	Parfois	Souvent
Jeux vidéo à moteur 3D (Quake, Flight Simulator, ...)	○	○	○
Logiciels de CAO (Autocad Architectural Desktop, Sketchup, ...)	○	○	○
Logiciels de dessin, modeling (3dsmax, Maya, ...)	○	○	○

Autres, précisez :

FIG. C.8 – Questionnaire initial - page 2

QUESTIONNAIRE FINAL - G1

Informations passation

Numéro ☐ Date ☐

IMPRESSIONS GENERALES

➪ **L'expérimentation vous a semblé :**

longue	○	○	○	○	○	○	○	*courte*
difficile	○	○	○	○	○	○	○	*facile*
ennuyeuse	○	○	○	○	○	○	○	*attractive*

➪ **La première étape (descriptions textuelles des cibles) vous a semblé :**

longue	○	○	○	○	○	○	○	*courte*
difficile	○	○	○	○	○	○	○	*facile*
ennuyeuse	○	○	○	○	○	○	○	*attractive*

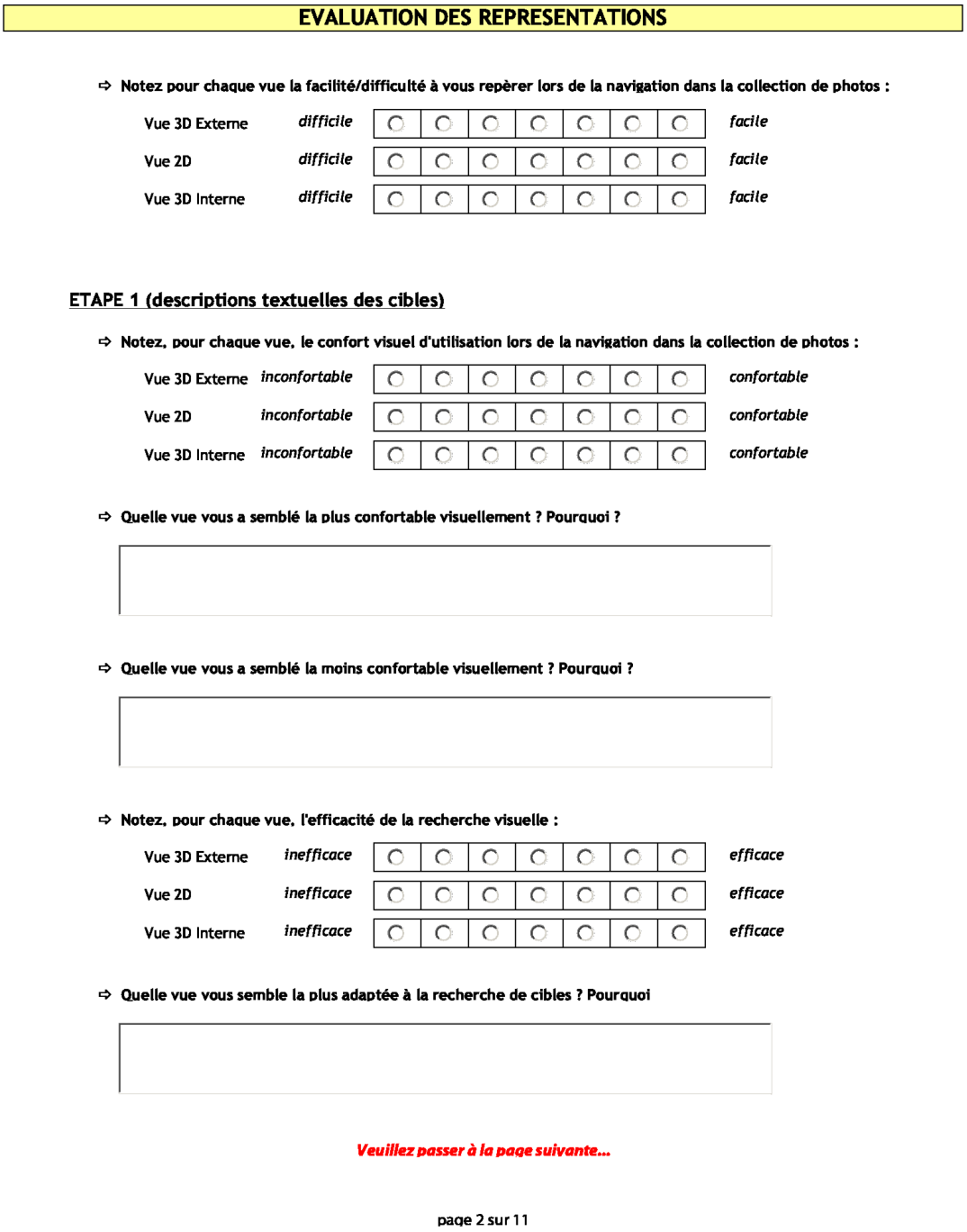
➪ **La seconde étape (présentations visuelles des cibles) vous a semblé :**

longue	○	○	○	○	○	○	○	*courte*
difficile	○	○	○	○	○	○	○	*facile*
ennuyeuse	○	○	○	○	○	○	○	*attractive*

Veuillez passer à la page suivante...

page 1 sur 11

FIG. C.9 – Questionnaire final - page 1



EVALUATION DES REPRESENTATIONS

⇨ **Notez pour chaque vue la facilité/difficulté à vous repèrer lors de la navigation dans la collection de photos :**

Vue 3D Externe	*difficile*	○	○	○	○	○	○	○	*facile*
Vue 2D	*difficile*	○	○	○	○	○	○	○	*facile*
Vue 3D Interne	*difficile*	○	○	○	○	○	○	○	*facile*

ETAPE 1 (descriptions textuelles des cibles)

⇨ **Notez, pour chaque vue, le confort visuel d'utilisation lors de la navigation dans la collection de photos :**

Vue 3D Externe	*inconfortable*	○	○	○	○	○	○	○	*confortable*
Vue 2D	*inconfortable*	○	○	○	○	○	○	○	*confortable*
Vue 3D Interne	*inconfortable*	○	○	○	○	○	○	○	*confortable*

⇨ **Quelle vue vous a semblé la plus confortable visuellement ? Pourquoi ?**

⇨ **Quelle vue vous a semblé la moins confortable visuellement ? Pourquoi ?**

⇨ **Notez, pour chaque vue, l'efficacité de la recherche visuelle :**

Vue 3D Externe	*inefficace*	○	○	○	○	○	○	○	*efficace*
Vue 2D	*inefficace*	○	○	○	○	○	○	○	*efficace*
Vue 3D Interne	*inefficace*	○	○	○	○	○	○	○	*efficace*

⇨ **Quelle vue vous semble la plus adaptée à la recherche de cibles ? Pourquoi**

Veuillez passer à la page suivante...

page 2 sur 11

FIG. C.10 – Questionnaire final - page 2

ETAPE 2 (présentations visuelles des cibles)

⇨ **Notez, pour chaque vue, le confort visuel d'utilisation lors de la navigation dans la collection de photos :**

Vue 3D Externe	*inconfortable*	○	○	○	○	○	○	○	*confortable*
Vue 2D	*inconfortable*	○	○	○	○	○	○	○	*confortable*
Vue 3D Interne	*inconfortable*	○	○	○	○	○	○	○	*confortable*

⇨ **Quelle vue vous a semblé la plus confortable visuellement ? Pourquoi ?**

⇨ **Quelle vue vous a semblé la moins confortable visuellement ? Pourquoi ?**

⇨ **Notez, pour chaque vue, l'efficacité de la recherche visuelle :**

Vue 3D Externe	*inefficace*	○	○	○	○	○	○	○	*efficace*
Vue 2D	*inefficace*	○	○	○	○	○	○	○	*efficace*
Vue 3D Interne	*inefficace*	○	○	○	○	○	○	○	*efficace*

⇨ **Quelle vue vous semble la plus adaptée à la recherche de cibles ? Pourquoi**

Veuillez passer à la page suivante...

page 3 sur 11

FIG. C.11 – Questionnaire final - page 3

APPRECIATION GLOBALE DES VUES

⇨ **Donnez une appréciation globale à chaque vue :**

Vue 3D Externe	*faible*	○	○	○	○	○	○	○	*élevée*
Vue 2D	*faible*	○	○	○	○	○	○	○	*élevée*
Vue 3D Interne	*faible*	○	○	○	○	○	○	○	*élevée*

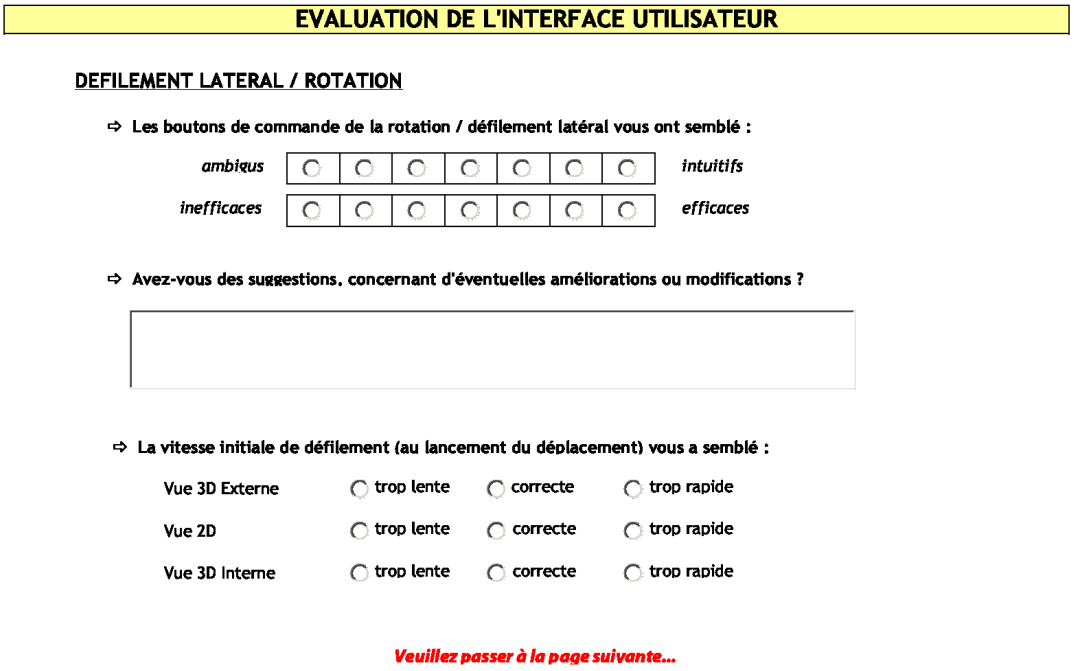
⇨ **Parmi les deux vues 3D (intérieure et extérieure), laquelle préferez-vous utiliser ? Pourquoi ?**

⇨ **Parmi les trois vues utilisées, laquelle seriez-vous prêt(e) à adopter pour votre usage courant ?**

⇨ **Certains aspects des représentations vous ont-ils posé problème ? Si oui, lesquels ?**

Veuillez passer à la page suivante...

FIG. C.12 – Questionnaire final - page 4



EVALUATION DE L'INTERFACE UTILISATEUR

DEFILEMENT LATERAL / ROTATION

⇨ **Les boutons de commande de la rotation / défilement latéral vous ont semblé :**

ambigus	○	○	○	○	○	○	○	*intuitifs*
inefficaces	○	○	○	○	○	○	○	*efficaces*

⇨ **Avez-vous des suggestions, concernant d'éventuelles améliorations ou modifications ?**

⇨ **La vitesse initiale de défilement (au lancement du déplacement) vous a semblé :**

Vue 3D Externe	○ trop lente	○ correcte	○ trop rapide
Vue 2D	○ trop lente	○ correcte	○ trop rapide
Vue 3D Interne	○ trop lente	○ correcte	○ trop rapide

Veuillez passer à la page suivante...

page 5 sur 11

FIG. C.13 – Questionnaire final - page 5

DISPOSITIF DE ZOOM DES PHOTOGRAPHIES

⇨ **Globalement, le dispositif de zoom des photographies vous a semblé :**

inutile	○	○	○	○	○	○	○	*utile*
difficile à utiliser	○	○	○	○	○	○	○	*facile à utiliser*
inefficace	○	○	○	○	○	○	○	*efficace*

⇨ **En dehors de la sélection des cibles, avez-vous utilisé le dispositif de zoom :**

- ○ *jamais*
- ○ *parfois*
- ○ *souvent*
- ○ *pour chaque recherche*

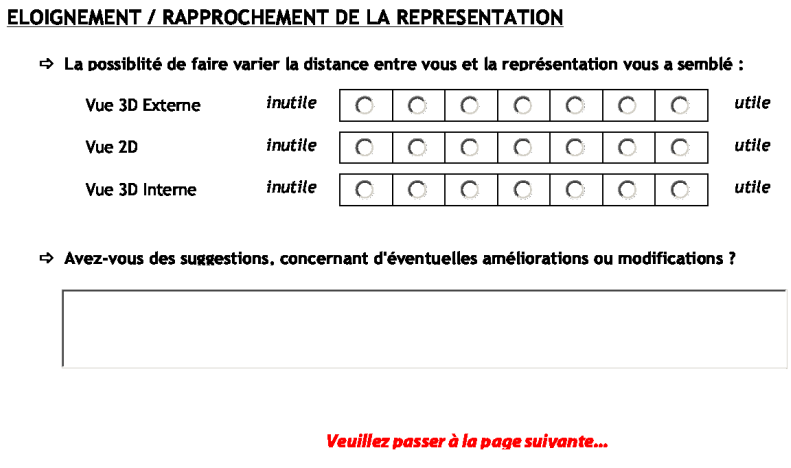
Pourquoi ?

⇨ **Avez-vous des suggestions, concernant d'éventuelles améliorations ou modifications ?**

Veuillez passer à la page suivante...

page 6 sur 11

FIG. C.14 – Questionnaire final - page 6



ELOIGNEMENT / RAPPROCHEMENT DE LA REPRESENTATION

⇨ **La possiblité de faire varier la distance entre vous et la représentation vous a semblé :**

Vue 3D Externe	*inutile*	○	○	○	○	○	○	○	*utile*
Vue 2D	*inutile*	○	○	○	○	○	○	○	*utile*
Vue 3D Interne	*inutile*	○	○	○	○	○	○	○	*utile*

⇨ **Avez-vous des suggestions, concernant d'éventuelles améliorations ou modifications ?**

Veuillez passer à la page suivante...

page 7 sur 11

FIG. C.15 – Questionnaire final - page 7

SYSTEME DE REPERAGE DE LA POSITION COURANTE

⇨ **Dans la première étape (descriptions textuelles), avez-vous utilisé le système de repérage pour contrôler votre position :**

Vue 3D Externe	◯ jamais	◯ parfois	◯ souvent
Vue 2D	◯ jamais	◯ parfois	◯ souvent
Vue 3D Interne	◯ jamais	◯ parfois	◯ souvent

⇨ **Dans la seconde étape (présentations visuelles), avez-vous utilisé le système de repérage pour retrouver plus facilement les cibles précédemment cherchées :**

Vue 3D Externe	◯ jamais	◯ parfois	◯ souvent
Vue 2D	◯ jamais	◯ parfois	◯ souvent
Vue 3D Interne	◯ jamais	◯ parfois	◯ souvent

⇨ **Avez-vous des suggestions, concernant d'éventuelles améliorations ou modifications ?**

Veuillez passer à la page suivante...

page 8 sur 11

FIG. C.16 – Questionnaire final - page 8

EVALUATION DES TACHES DE RECHERCHE

ETAPE 1 (descriptions textuelles des cibles)

⇨ **La recherche visuelle de cibles à partir de leur description textuelle vous a semblé une tâche :**

artificielle ○ ○ ○ ○ ○ ○ ○ *réaliste*

⇨ **Les descriptions des photographies étaient-elles suffisantes pour qu'il n'y ait pas d'ambiguïté ?**

○ oui
○ non

Si non, expliquez pourquoi :

⇨ **Quelles caractéristiques de la photographie cible avez-vous utilisées pour la retrouver ?**

ETAPE 2 (présentations visuelles des cibles)

⇨ **La recherche visuelle de cibles à partir de leur présentation visuelle vous a semblé une tâche :**

artificielle ○ ○ ○ ○ ○ ○ ○ *réaliste*

⇨ **Quelles caractéristiques de la photographie avez-vous utilisées pour la retrouver ?**

Veuillez passer à la page suivante...

FIG. C.17 – Questionnaire final - page 9

DISPOSITIF DE SUIVI DU REGARD

⇨ **Le port de l'eye tracker vous a-t-il gêné dans la réalisation des tâches de recherche ?**

pas du tout ○ ○ ○ ○ ○ ○ ○ *beaucoup*

⇨ **Avez-vous ressenti une certaine fatigue visuelle ?**

aucune ○ ○ ○ ○ ○ ○ ○ *élevée*

⇨ **Avez-vous ressenti une certaine fatigue au niveau de la nuque ?**

aucune ○ ○ ○ ○ ○ ○ ○ *élevée*

Veuillez passer à la page suivante...

page 10 sur 11

POUR CONCLURE...

⇨ **Comment jugez-vous l'apport de la 3D à l'activité de recherche visuelle ?**

⇨ **En tant que concepteur de logiciels grand public, quelles applications envisageriez-vous pour chacune des représentations 3D ?**

Merci beaucoup pour votre contribution

page 11 sur 11

FIG. C.18 – Questionnaire final - pages 10 et 11

Table des figures

2.1 Données intrinsèques et extrinsèques d'une photographie 7

2.2 Logiciel *XnView* : exemple d'affichage *overview + details* 12

2.3 Murs en perspective 16

2.4 Arbre conique : l'arborescence est présentée horizontalement et comprend ici 4 niveaux (plus la racine) ; les boutons situés à gauche permettent de changer le mode de représentation 16

2.5 *Source Viewer 3D* : chaque *poly cylinder* représente une ligne de code, la couleur représente le type de structure de contrôle et la hauteur, le niveau d'emboîtement 17

2.6 *Data Mountain* avec l'affichage de 100 favoris 18

3.1 Exemple de plan en perspective avec 16 photographies 27

3.2 Comparaison 2D (a, b) - plan en perspective (c, d) : la vue globale du document est représentée par les schémas b et d, les schémas a et c représentant la vue courante ; la partie réellement « exploitable » est en jaune 28

3.3 Vues intérieure et extérieure du cylindre vertical 29

3.4 Non équivalence des vues interne et externe 31

3.5 Evolution de la taille d'un document 36

3.6 Taille réelle sur un écran de 19 pouces (1280*1024) d'une image de 30 pixels (un veau couché sur de la paille) 36

3.7 Publicité pour le boîtier *AppleTV*, Apple, 2007 38

3.8 Publicité pour une offre de *VOD*, Club Internet, 2007 38

4.1 Photographies sur le thème des feux d'artifice 44

5.1 Interacteur pour la vue interne 52

5.2 Représentation du point de vue courant pour la locomotion (à gauche) et la manipulation (à droite) 54

5.3 Vue intérieure du cylindre vertical 55

5.4 Vue extérieure du cylindre vertical 56
5.5 Photographies servant de cibles 57
5.6 Cible (a) et description correspondante (b) 58
5.7 Nature des différentes informations présentes dans une description 58
5.8 Temps de recherche cumulés pour chaque sujet 62

6.1 Trois projets non retenus d'interacteurs 80
6.2 Interacteurs des vues interne et externe 81
6.3 Vue intérieure du cylindre vertical 83
6.4 Vue extérieure du cylindre vertical 84
6.5 Cibles 85
6.6 Exemple de description 85
6.7 Position horizontale des cibles sur le cylindre - l'arc en rouge indique la partie visible initialement 86

7.1 Trois projets non retenus d'interacteurs permettant l'activation du défilement . . 116
7.2 Dispositif de rapprochement-éloignement de la représentation : le survol du bouton sensitif à gauche fait apparaître la zone de droite comprenant deux boutons activables par un clic (« + » pour approcher, « - » pour éloigner) 117
7.3 Dispositif de repérage spatial non retenu : barre de progression graduée de 0 à 100 % 118
7.4 Dispositifs de repérage pour les 3 vues : la barre rouge indique la colonne affichée au centre de l'écran 119
7.5 Différence de qualité entre les 2 résolutions utilisées pour les photographies sous forme de vignettes (à gauche) et agrandies (à droite) 120
7.6 Vues interne et externe de la représentation cylindrique 121
7.7 Représentation 2D 121
7.8 Cibles utilisées pour chacune des trois vues pendant T1 et T2 122
7.9 Position des cibles : la partie de la vue visible initialement est représentée en jaune, les cibles en rouge 124
7.10 Indications données à la fin de chaque animation dans T1 : la cible agrandie, sa position dans la vue (en rouge) et le « chemin » le plus court pour l'atteindre (en vert) 126
7.11 ASL modèle 501 et head-tracker magnétique 127
7.12 Interface de calibrage et mire à 9 points 127
7.13 Exemple de question avec une échelle de Likert à 7 degrés 128

7.14 Cartes de température correspondant aux 2 stratégies d'exploration visuelle (1 sujet, 1 cible) : les zones colorées (du bleu au rouge) correspondent à la position des fixations ; les variations de couleur représentent la densité des fixations, croissante du bleu au rouge 137

7.15 Cartes de température correspondant aux 4 comportements visuels identifiés (4 sujets, 4 cibles) : les zones colorées (du bleu au rouge) correspondent à la position des fixations ; les variations de couleur représentent la densité des fixations, croissante du bleu au rouge 139

A.1 Ordre de présentation des cibles 169

A.2 Descriptions des cibles utilisées dans la vue interne (VI) 170

A.3 Descriptions des cibles utilisées dans la vue externe (VE) 171

A.4 Questionnaire intial - page 1 172

A.5 Questionnaire intial - page 2 173

A.6 Questionnaire final - page 1 174

A.7 Questionnaire final - page 2 175

A.8 Questionnaire final - page 3 176

B.1 Ordre de présentation des cibles 178

B.2 Descriptions des cibles utilisées dans la vue interne (VI) - page 1 179

B.3 Descriptions des cibles utilisées dans la vue interne (VI) - page 2 180

B.4 Descriptions des cibles utilisées dans la vue externe (VE) - page 1 181

B.5 Descriptions des cibles utilisées dans la vue externe (VE) - page 2 182

B.6 Instructions - page 1 183

B.7 Instructions - page 2 184

B.8 Instructions - page 3 185

B.9 Consignes - page 1 186

B.10 Consignes - page 2 187

B.11 Consignes - page 3 188

C.1 Descriptions des cibles utilisées dans la vue interne (VI) 190

C.2 Descriptions des cibles utilisées dans la vue externe (VE) 191

C.3 Descriptions des cibles utilisées dans la vue 2D (V2D) 192

C.4 Instructions - page 1 193

C.5 Instructions - page 2 194

C.6 Instructions - page 3 195

C.7 Questionnaire initial - page 1 196
C.8 Questionnaire initial - page 2 197
C.9 Questionnaire final - page 1 198
C.10 Questionnaire final - page 2 199
C.11 Questionnaire final - page 3 200
C.12 Questionnaire final - page 4 201
C.13 Questionnaire final - page 5 202
C.14 Questionnaire final - page 6 203
C.15 Questionnaire final - page 7 204
C.16 Questionnaire final - page 8 205
C.17 Questionnaire final - page 9 206
C.18 Questionnaire final - pages 10 et 11 207

Liste des tableaux

5.1 Identification des participants 59

5.2 Nombre de succès par vue (VE, VI) et par tâche (T1, T2) 63

5.3 Examens initiaux cumulés pour chaque sujet dans VE et VI (T2) 64

5.4 Sens de rotation pour chaque sujet dans VE et VI (T2) : un mauvais choix de sens de rotation est représenté en rouge, et un sens de rotation constant est représenté en bleu 64

5.5 Nombre d'erreurs cumulées (8 sujets) et nature de celles-ci, par vue (VI et VE) pour T1 65

5.6 Nombre cumulé d'utilisations des dispositifs de zoom, par vue et par tâche 67

6.1 Identification des participants 87

6.2 Performances des participants : comparaison entre les vues VI et VE (les pourcentages sont calculés sur la plus faible valeur du couple VE/VI) pour le nombre d'erreurs, le cumul des temps de recherche, la vitesse moyenne, et le cumul des longueurs de parcours 93

6.3 Performances brutes des participants : comparaison entre les vues VI et VE pour le nombre d'erreurs, le cumul des temps de recherche (secondes), la vitesse moyenne de défilement (colonnes/seconde) et le cumul des longueurs de parcours (colonnes) 94

6.4 Performances des participants : moyennes, écarts-types et tests statistiques pour le temps **T** (cumul pour les 10 cibles d'une vue, en secondes), le nombre de succès **S** (pour les 10 cibles d'une vue), la vitesse de parcours **V** (moyenne calculée pour 10 cibles d'une vue, en colonnes/seconde) et la longueur de parcours **L** (cumul pour les 10 cibles d'une vue, en colonnes) 95

6.5 Préférences des participants en faveur de VI ou VE exprimées dans le questionnaire final **Q** et l'entretien de débriefing **D**. **E** (resp. **I**) indique une préférence pour la vue externe (resp. interne) ; $<$ marque une préférence faible, et **H** indique une hésitation entre VI et VE 97

6.6 Stratégies d'exploration visuelle des sujets pour le groupe GI. **N** (resp. **F**) indique une bonne (resp. faible) acuité visuelle ; **E** (resp. **C**) indique une stratégie « colonnes entrantes » (resp. « colonnes centrales ») ; **?** marque l'absence d'identification de la stratégie 100

6.7 Stratégies d'exploration visuelle des sujets pour les groupes GE et GH (voir légende du tableau 6.6) 100

7.1 Identification des participants (20) : n° de passation, âge, sexe et formation universitaire 125

7.2 Evaluations des vues d'après les questionnaires finaux (20 participants) : note globale (**G**), facilité du repérage spatial (**RS**), confort visuel (**CV**) et efficacité de la recherche (**ER**). Tests statistiques de Friedman. En vert, les notes maximales pour chaque critère 132

7.3 Préférences des participants (20) d'après le questionnaire final (**Q**) et le débriefing (**D**) pour la vue interne (**I**), externe (**E**), **2D** ou pour les 2 vues 3D en général (**3D**). Le signe « > » indique une faible préférence. **2D-I** ou **2D-E** indiquent une hésitation entre ces 2 vues 132

7.4 Performances des participants pour T1 : comparaison entre les 3 vues **VI**, **VE** et **V2D** pour le nombre de succès **S** et la moyenne des temps de recherche **T** (secondes) 133

7.5 Performances des participants pour T2 : comparaison entre les 3 vues **VI**, **VE** et **V2D** pour le nombre de succès **S**, la moyenne des temps de recherche **T** (secondes), la longueur moyenne de parcours **L** (images) et la vitesse moyenne de rotation **V** (images/seconde) 134

7.6 Nombre de succès par vue dans T1 (20 participants) : moyennes, écarts-types et tests statistiques (Wilcoxon). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances 135

7.7 Vitesse de défilement/rotation dans T2 (20 participants) : moyennes, écarts-types et tests statistiques (Wilcoxon) En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances 135

7.8 Stratégies d'exploration visuelle des sujets des groupes GI et GE, et du sujet préférant indifféremment l'une ou l'autre vue 3D : balayage des colonnes entrantes (**E**), des colonnes centrales (**C**), de la largeur de l'écran (**L**) et stratégie indéterminée (**?**) 136

7.9 Stratégies d'exploration visuelle des sujets du groupe G2D : balayage des colonnes entrantes (**E**) et des colonnes centrales (**C**) 137

7.10 Cumul des comportements visuels sur l'ensemble des recherches effectuées avec chaque vue par les 20 sujets, pendant T2 140

7.11 Temps de recherche moyen (secondes), calculé sur l'ensemble des participants (20), pour les deux cibles les plus proches de l'affichage initial, trouvées dès leur première apparition à l'écran et, pour les vues 3D, en choisissant le sens de rotation optimal. « nb. rech » désigne le nombre de recherches visuelles satisfaisant ce(s) critère(s) . 141

7.12 Pour chaque vue, inverse de l'amplitude des saccades (**NF**/**SP**) et durée des fixations (**D**) pendant T2 : moyennes (20 participants), écarts-types et tests statistiques (Wilcoxon). NF désigne le nombre moyen de fixations et SP la longueur moyenne du parcours oculaire (en pixels), par recherche visuelle (pendant le défilement de la vue). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances 142

7.13 Pour chaque vue, inverse de l'amplitude des saccades (**NF**/**SP**) et durée des fixations (**D**) pendant les recherches visuelles de T2 pour lesquelles la position de la cible n'a pas été mémorisée : moyennes (20 participants), écarts-types et tests statistiques (Wilcoxon). NF désigne le nombre moyen de fixations et SP la longueur moyenne du parcours oculaire (en pixels), par recherche visuelle (pendant le défilement de la vue). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances 143

7.14 VI *vs* VE (seconde étude, 20 participants) : appréciations globales (questionnaire final) et tests statistiques (Wilcoxon). En vert la meilleure performance ; en orange les tendances 144

7.15 VI *vs* VE (seconde étude, 20 sujets) : performances pour T1 (nombre moyen de succès et temps moyen de recherche) et pour T2 (vitesse moyenne de rotation du cylindre et longueur moyenne de parcours), tests statistiques (Wilcoxon). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances 145

7.16 Jugements subjectifs sur la vue 2D et la vue 3D (préférée), d'après les questionnaires finaux : note globale (**G**), facilité du repérage spatial (**RS**), confort visuel (**CV**) et efficacité de la recherche (**ER**). Moyennes (20 participants), écarts-types et tests statistiques (Wilcoxon). En vert la meilleure performance ; en bleu les résultats significatifs des tests statistiques et en orange les tendances 150

Bibliographie

[Almeida *et al.*, 2006] R. Almeida, P. Cubaud, J. Dupire, S. Natkin et A. Topol. *Technologies for E-Learning and Digital Entertainment*, volume Volume 3942/2006 de *Lecture Notes in Computer Science*, chapitre Experiments Towards 3D Immersive Interaction for Digital Libraries, pages 1348–1357. Springer Berlin, 2006.

[Appert et Fekete, 2006] C. Appert et J.-D. Fekete. OrthoZoom scroller: 1D multi-scale navigation. Dans *CHI '06: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 21–30. ACM Press, 2006.

[Apple, 2005] Apple. *Spotlight, find anything on your Mac instantly.* Apple Computer, avril 2005.

[Attneave, 1955] F. Attneave. Symmetry, information, and memory for patterns. *American Journal of Psychology*, 68(2):209–222, 1955.

[Baudisch *et al.*, 2002] P. Baudisch, N. Good, V. Bellotti et P. Schraedley. Keeping things in context: a comparative evaluation of focus + context screens, overviews, and zooming. Dans *CHI '02: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 259–266. ACM Press, 2002.

[Beaudouin-Lafon, 2004] M. Beaudouin-Lafon. Designing interaction, not interfaces. Dans *AVI '04: Proceedings of the working conference on Advanced Visual Interfaces*, pages 15–22. ACM Press, 2004.

[Bederson *et al.*, 2002] B. B. Bederson, B. Shneiderman et M. Wattenberg. Ordered and quantum treemaps: making effective use of 2D space to display hierarchies. *ACM Transactions on Graphics*, 21(4):833–854, 2002.

[Bederson, 2001] B. B. Bederson. PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps. Dans *UIST '01: Proceedings of the 14th annual ACM symposium on User Interface Software and Technology*, pages 71–80. ACM Press, 2001.

[Bladh *et al.*, 2004] T. Bladh, D. Carr et J. Scholl. Extending treemaps to three dimensions: a comparative study. Dans *APCHI '04: Proceedings of the 6th Asia-Pacific conference on Computer-Human Interaction*, pages 50–59. Springer-Verlag, 2004.

[Borgman, 1989] C. L. Borgman. All users of information retrieval systems are not created equal: an exploration into individual differences. *Information Processing and Management*, 25(3):237–251, 1989.

[Card *et al.*, 1996] S. K. Card, G. G. Robertson et W. York. The WebBook and the Web Forager: an information workspace for the world-wide web. Dans *CHI '96: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 111–118. ACM Press, 1996.

[Card *et al.*, 2004] S. K. Card, L. Hong, J. D. Mackinlay et E. H. Chi. 3book: a scalable 3D virtual book. Dans *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1095–1098. ACM, 2004.

[Chen, 2000] C. Chen. Individual differences in a spatial-semantic virtual environment. *Journal of the American Society for Information Science*, 51(6):529–542, 2000.

[Chittaro et Scagnetto, 2001] L. Chittaro et I. Scagnetto. Is semitransparency useful for navigating virtual environments? Dans *VRST '01: Proceedings of the ACM symposium on Virtual Reality Software and Technology*, pages 159–166. ACM Press, 2001.

[Christmann et Carbonell, 2006] Olivier Christmann et Noëlle Carbonell. Browsing through 3D representations of unstructured picture collections: an empirical study. Dans *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 445–448. ACM, 2006.

[Christmann et Carbonell, 2008] Olivier Christmann et Noëlle Carbonell. Navigation and search in 3D visualizations of large unstructured photo collections: An empirical study. Dans *Visual '08: Proceedings of the10th International Conference on Visual Information Systems*, pages 31–43. Springer-Verlag, 2008.

[Chu *et al.*, 2004] Y.-C. Chu, D. Bainbridge, M. Jones et I. H. Witten. Realistic books: a bizarre homage to an obsolete medium? Dans *JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pages 78–86. ACM, 2004.

[Cockburn et McKenzie, 2000] A. Cockburn et B. McKenzie. An evaluation of cone trees. Dans *People and Computers XV: Proceedings of the 2000 british Computer society conference on Human-Computer Interaction*, pages 425–436. Springer-Verlag, 2000.

[Cockburn et McKenzie, 2001] A. Cockburn et B. McKenzie. 3D or not 3D?: evaluating the effect of the third dimension in a document management system. Dans *CHI '01: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 434–441. ACM Press, 2001.

[Cockburn et McKenzie, 2002] A. Cockburn et B. McKenzie. Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments. Dans *CHI '02: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 203–210. ACM Press, 2002.

[Cockburn, 2004] A. Cockburn. Revisiting 2D vs 3D implications on spatial memory. Dans *AUIC '04: Proceedings of the fifth conference on Australasian User Interface*, pages 25–31. Australian Computer Society, Inc., 2004.

[Combs et Bederson, 1999] T. T. A. Combs et B. B. Bederson. Does zooming improve image browsing? Dans *DL '99: Proceedings of the fourth ACM conference on Digital libraries*, pages 130–137. ACM Press, 1999.

[Cubaud *et al.*, 2002] P. Cubaud, P. Stokowski et A. Topol. Binding browsing and reading activities in a 3D digital library. Dans *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 281–282. ACM, 2002.

[Czerwinski *et al.*, 1999] M. Czerwinski, M. van Dantzich, G. G. Robertson et H. Hoffman. The contribution of thumbnail image, mouse-over text and spatial location memory to web page retrieval in 3D. Dans *INTERACT '99: Proceedings of the international conference on Human-Computer Interaction*, pages 163–170. IOS Press, 1999.

[Czerwinski *et al.*, 2002] M. Czerwinski, D. S. Tan et G. G. Robertson. Women take a wider view. Dans *CHI '02: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 195–202. ACM Press, 2002.

[Darken *et al.*, 1998] R. P. Darken, T. Allard et L. B. Achille. Spatial orientation and wayfinding in large-scale virtual spaces: an introduction. *Presence*, 7(2):101–107, 1998.

[Dollinger et Hoyer, 1996] S. Dollinger et W. Hoyer. Age and skill differences in the processing demands of visual inspection. *Applied Cognitive Psychology*, 10:225–239, 1996.

[Drucker *et al.*, 2004] S. M. Drucker, C. Wong, A. Roseway, S. Glenner et S. De Mar. MediaBrowser: reclaiming the Shoebox. Dans *AVI '04: Proceedings of the working conference on Advanced Visual Interfaces*, pages 433–436. ACM Press, 2004.

[Eakins et Graham, 1999] J. P. Eakins et M. E. Graham. Content-based image retrieval. *JISC Technology Applications Programme Report*, 39, octobre 1999.

[Egan, 1988] D. E. Egan. *Individual differences in human-computer interaction*, volume Handbook of Human-Computer Interaction, pages 543–568. M. Helander, 1988.

[Enser, 1995] P. G. B. Enser. Pictorial information retrieval. *Journal of Documentation*, 51(2):126–170, 1995.

[Fairchild *et al.*, 1999] K. M. Fairchild, S. E. Poltrock et G. W. Furnas. *SemNet: time-dimensional graphic representations of large knowledge bases*, volume Readings in information visualization: using vision to think, pages 190–206. Morgan Kaufmann Publishers Inc., 1999.

[Fekete et Plaisant, 2002] J.-D. Fekete et C. Plaisant. Interactive information visualization of a million items. Dans *INFOVIS '02: Proceedings of the IEEE Symposium on Information Visualization*, pages 117–124. IEEE Computer Society, 2002.

[Folk *et al.*, 1992] C. L. Folk, R. W. Remington et J. C. Johnston. Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology*, 18:1030–1044, 1992.

[Furnas, 1986] G. W. Furnas. Generalized fisheye views. Dans *CHI '86: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 16–23. ACM Press, 1986.

[Guiard *et al.*, 2006] Y. Guiard, O. Chapuis, Y. Du et M. Beaudouin-Lafon. Allowing camera tilts for document navigation in the standard gui: a discussion and an experiment. Dans *AVI '06: Proceedings of the working conference on Advanced Visual Interfaces*, pages 241–244. ACM Press, 2006.

[Harrouet *et al.*, 2006] F. Harrouet, T. Jourdan et E. Cazeaux. *ARéVi*, chapitre Outils et modèles informatiques des environnements virtuels. Le traité de la réalité virtuelle. Les Presses de l'Ecole des Mines de Paris, 2006.

[Harrouet, 2000] F. Harrouet. *oRis : s'immerger par le langage pour le prototypage d'univers virtuels à base d'entités autonomes*. PhD thesis, Université de Bretagne Occidentale, 2000.

[Heer et Card, 2003] J. Heer et S. K. Card. Efficient user interest estimation in fisheye views. Dans *CHI '03: CHI '03 extended abstracts on Human Factors in Computing Systems*, pages 836–837. ACM Press, 2003.

[Huynh *et al.*, 2005] D. F. Huynh, S. M. Drucker, P. Baudisch et C. Wong. Time Quilt: scaling up zoomable photo browsers for large, unstructured photo collections. Dans *CHI '05: extended abstracts on Human Factors in Computing Systems*, pages 1937–1940. ACM Press, 2005.

[Itti et Koch, 2000] L. Itti et C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10–12):1489–1506, 2000.

[Kang *et al.*, 2003] H. Kang, C. Plaisant et B. Shneiderman. New approaches to help users get started with visual interfaces: multi-layered interfaces and integrated initial guidance. Dans *DG '03: Proceedings of the 2003 annual national conference on Digital Government Research*, pages 1–6. Digital Government Research Center, 2003.

[Kang et Shneiderman, 2000] H. Kang et B. Shneiderman. Visualization methods for personal photo collections: Browsing and searching in the photofinder. Dans *IEEE International Conference on Multimedia and Expo (III)*, pages 1539–1542. IEEE, 2000.

[Kato, 1992] T. Kato. Database architecture for content-based image retrieval. Dans *Proceedings of the SPIE, Image Storage and Retrieval Systems*, volume 1662, pages 112–123, 1992.

[Kearsley, 1994] G. Kearsley. *Public Access Systems: Bringing Computer Power to the People*. Greenwood Publishing Group Inc., Westport, CT, USA, 1994.

[Keim *et al.*, 2007] D. A. Keim, J. Schneidewind et M. Sips. *Scalable pixel based visual data exploration*, volume Pixelization paradigm, pages 12–14. Springer, May 2007.

[Keim, 2001] D. A. Keim. Visual exploration of large data sets. *Communications of the ACM*, 44(8):38–44, 2001.

[Khella et Bederson, 2004] A. Khella et B. B. Bederson. Pocket photoMesa: a zoomable image browser for pdas. Dans *MUM '04: Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia*, pages 19–24. ACM Press, 2004.

[Kobsa, 2004] A. Kobsa. User experiments with tree visualization systems. Dans *INFOVIS '04: Proceedings of the IEEE symposium on Information Visualization*, pages 9–16. IEEE Computer Society, 2004.

[Kuchinsky *et al.*, 1999] A. Kuchinsky, C. Pering, M. L. Creech, D.Freeze, B. Serra et J. Gwizdka. FotoFile: a consumer multimedia organization and retrieval system. Dans *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 496–503. ACM Press, 1999.

[Kules *et al.*, 2004] B. Kules, H. Kang, C. Plaisant, A. Rose et B. Shneiderman. Immediate usability: a case study of public access design for a community photo library. *Interacting with Computers*, 16(3):1171–1193, 2004.

[Kustanowitz et Shneiderman, 2005] J. Kustanowitz et B. Shneiderman. Meaningful presentations of photo libraries: rationale and applications of bi-level radial quantum layouts. Dans *JCDL '05: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*, pages 188–196. ACM Press, 2005.

[Lamping *et al.*, 1995] J. Lamping, R. Rao et P. Pirolli. A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. Dans *CHI '95: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 401–408. ACM Press, 1995.

[Likert, 1932] R. Likert. *A technique for the measurement of attitudes*. s.n., 1932.

[Mackinlay *et al.*, 1991] J. D. Mackinlay, G. G. Robertson et S. K. Card. The Perspective Wall: detail and context smoothly integrated. Dans *CHI '91: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 173–176. ACM Press, 1991.

[Marcus *et al.*, 2003] A. Marcus, L. Feng et J. I. Maletic. 3D representations for software visualization. Dans *SoftVis '03: Proceedings of the 2003 ACM symposium on Software Visualization*, pages 27–36. ACM Press, 2003.

[Messick, 1976] S. Messick. *Individuality in Learning*. Jossey-Bass, 1976.

[Mills *et al.*, 2000] T. J. Mills, D. Pye, D. Sinclair et K. R. Wood. Shoebox: a digital photo management system. Rapport technique, AT&T Laboratories, 2000.

[Moghaddam *et al.*, 2004] B. Moghaddam, Q. Tian, N. Lesh, C. Shen et T. S. Huang. Visualization and user-modeling for browsing personal photo libraries. *International Journal of Computer Vision*, 56(1-2):109–130, 2004.

[Mou *et al.*, 2006] W. Mou, T. P. McNamara, B. Rump et C. Xiao. Roles of egocentric and allocentric spatial representations in locomotion and reorientation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 32(6):1274–1290, november 2006.

[Olston et Chi, 2003] C. Olston et E. H. Chi. ScentTrails: integrating browsing and searching on the web. *ACM Transactions on Computer-Human Interactions*, 10(3):177–197, 2003.

[Parker *et al.*, 1992] D. M. Parker, J. R. Lishman et J. Hugues. Temporal integration of spatially filtered visual images. *Perception*, 21(2):147–160, 1992.

[Pirolli *et al.*, 2000] P. Pirolli, S. K. Card et M. M. Van Der Wege. The effect of information scent on searching information: visualizations of large tree structures. Dans *AVI '00: Proceedings of the working conference on Advanced Visual Interfaces*, pages 161–172. ACM Press, 2000.

[Pirolli *et al.*, 2001] P. Pirolli, S. K. Card et M. M. Van Der Wege. Visual information foraging in a focus + context visualization. Dans *CHI '01: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 506–513. ACM Press, 2001.

[Plaisant *et al.*, 1995] C. Plaisant, D. Carr et B. Shneiderman. Image-browser taxonomy and guidelines for designers. *IEEE Software*, 12(2):21–32, 1995.

[Plaisant, 2004] C. Plaisant. The challenge of information visualization evaluation. Dans *AVI '04: Proceedings of the working conference on Advanced Visual Interfaces*, pages 109–116. ACM Press, 2004.

[Platt *et al.*, 2002] J. C. Platt, M. Czerwinski et B. A. Field. Phototoc: Automatic clustering for browsing personal photographs. Rapport technique, Microsoft Research Technical Report MSR-TR-2002-17, 2002.

[Rekimoto et Green, 1993] J. Rekimoto et M. Green. The information cube using transparency in 3D information visualization. Dans *WIST '93: Proceedings of the third annual Workshop Information Technologies & Systems*, pages 125–132, 1993.

[Robertson *et al.*, 1991] G. G. Robertson, J. D. Mackinlay et S. K. Card. Cone Trees: animated 3d visualizations of hierarchical information. Dans *CHI '91: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 189–194. ACM Press, 1991.

[Robertson *et al.*, 1998] G. G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel et M. van Dantzich. Data Mountain: using spatial memory for document management. Dans *UIST '98: Proceedings of the symposium on User Interface Software and Technology*, pages 153–162. ACM Press, 1998.

[Robertson *et al.*, 2000] G. G. Robertson, M. van Dantzich, D. C. Robbins, M. Czerwinski, K. Hinckley, K. Risden, D. Thiel et V. Gorokhovsky. The Task Gallery: a 3D window manager. Dans *CHI '00: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 494–501. ACM Press, 2000.

[Robertson et Mackinlay, 1993] G. G. Robertson et J. D. Mackinlay. The document lens. Dans *UIST '93: Proceedings of the 6th annual ACM symposium on User interface software and technology*, pages 101–108. ACM Press, 1993.

[Rodden et Wood, 2003] K. Rodden et K. R. Wood. How do people manage their digital photographs? Dans *CHI '03: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 409–416. ACM Press, 2003.

[Rodden, 2002] K. Rodden. Evaluating similarity-based visualisations as interfaces for image browsing. Rapport Technique UCAM-CL-TR-543, University of Cambridge, Computer Laboratory, septembre 2002.

[Rogers *et al.*, 1994] B. Rogers, S. J.Cunningham et G.Holmes. Navigating the virtual library: a 3D browsing interface for information retrieval. Dans *Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems*, pages 467–471, 1994.

[Schyns et Oliva, 1994] P. G. Schyns et A. Oliva. From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5(4):195–200, 1994.

[Sebrechts *et al.*, 1999] M. M. Sebrechts, J. V. Cugini, S. J. Laskowski, J. Vasilakis et M. S. Miller. Visualization of search results: a comparative evaluation of text, 2D, and 3D interfaces. Dans *SIGIR '99: Proceedings of the 22nd annual international ACM SIGIR conference on Research and Development in Information Retrieval*, pages 3–10. ACM Press, 1999.

[Shen *et al.*, 2003] C. Shen, N. Lesh et F. Vernier. Personal digital historian: story sharing around the table. *interactions*, 10(2):15–22, 2003.

[Shneiderman *et al.*, 2000] B. Shneiderman, D. Feldman, A. Rose et X. Ferré-Grau. Visualizing digital library search results with categorical and hierarchical axes. Dans *DL '00: Proceedings of the fifth ACM conference on Digital Libraries*, pages 57–66. ACM Press, 2000.

[Shneiderman *et al.*, 2002] B. Shneiderman, H. Kang, B. Kules, C. Plaisant, A. Rose et R. Rucheir. A photo history of SIGCHI: evolution of design from personal to public. *Interactions*, 9(3):17–23, 2002.

[Shneiderman, 1992] B. Shneiderman. Tree visualization with tree-maps: 2D space-filling approach. *ACM Transactions on Graphics*, 11(1):92–99, 1992.

[Shneiderman, 1996] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. Dans *VL '96: Proceedings of the 1996 IEEE symposium on Visual Languages*, pages 336–343. IEEE Computer Society, 1996.

[Shneiderman, 2003] B. Shneiderman. Why not make interfaces better than 3D reality? *IEEE Computer Graphics and Applications*, 23(6):12–15, 2003.

[Simonin *et al.*, 2005] J. Simonin, S. Kieffer et N. Carbonell. Effects of display layout on gaze activity during visual search. Dans *INTERACT '05: Proceedings of the international conference IFIP TC13 on Human-Computer Interaction*, pages 1054–1058. Springer-Verlag, 2005.

[Snavely *et al.*, 2006] N. Snavely, S. M. Seitz et R. Szeliski. Photo Tourism: exploring photo collections in 3D. Dans *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, pages 835–846. ACM Press, 2006.

[Stasko et Zhang, 2000] J. Stasko et E. Zhang. Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. Dans *INFOVIS '00: Proceedings*

of the IEEE Symposium on Information Visualization, pages 57–65. IEEE Computer Society, 2000.

[Suh *et al.*, 2003] B. Suh, H. Ling, B. B. Bederson et D. W. Jacobs. Automatic thumbnail cropping and its effectiveness. Dans *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 95–104. ACM Press, 2003.

[Tan *et al.*, 2006] D. S. Tan, D. Gergle, P. Scupelli et R. Pausch. Physically large displays improve performance on spatial tasks. *ACM Transaction on Computer-Human Interaction*, 13(1):71–99, 2006.

[Tanaka *et al.*, 2004] Y. Tanaka, Y. Okada et K. Niijima. Interactive interfaces of treecube for browsing 3D multimedia data. Dans *AVI '04: Proceedings of the working conference on Advanced Visual Interfaces*, pages 298–302. ACM Press, 2004.

[Tavanti et Lind, 2001] M. Tavanti et M. Lind. 2D vs 3D, implications on spatial memory. Dans *INFOVIS '01: Proceedings of the IEEE Symposium on Information Visualization 2001*, pages 139–145. IEEE Computer Society, 2001.

[Treisman, 1985] A. Treisman. Preattentive processing in vision. *Computer Vision, Graphics and Image Processing*, 31(2):156–177, 1985.

[Trumbo, 1998] J. Trumbo. Spatial memory and design: a conceptual approach to the creation of navigable space in multimedia design. *Interactions*, 5(4):26–34, 1998.

[Vernier et Nigay, 1997] F. Vernier et L. Nigay. Représentations multiples d'une grande quantité d'informations. Dans *IHM 1997 : Actes de la conférence de l'Association Francophone d'Interaction Homme-Machine*, pages 183–190. Cépaduès Publications, 1997.

[Winer, 1971] B. J. Winer. *Statistical principles in experimental design (2nd edition)*. McGraw-Hill, 1971.

[Wiss et Carr, 1999] U. Wiss et D. A. Carr. An empirical study of task support in 3d information visualizations. Dans *INFOVIS '99: Proceedings of the IEEE symposium on Information Visualization*, pages 392–399. IEEE Computer Society, 1999.

[Wiza *et al.*, 2004] W. Wiza, K. Walczak et W. Cellary. Periscope: a system for adaptive 3D visualization of search results. Dans *Web3D '04: Proceedings of the ninth international conference on 3D Web Technology*, pages 29–40. ACM Press, 2004.

[Yang *et al.*, 2002] M.-H. Yang, D. J. Kriegman et N. Ahuja. Detecting faces in images: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(1):34–58, 2002.

[Zhao *et al.*, 2003] W. Zhao, R. Chellappa, P. J. Phillips et A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4):399–458, 2003.

Résumé

Le but de cette thèse est de proposer et d'évaluer une représentation 3D de collections non structurées de documents visuels. L'objectif est d'accroître le confort visuel, de réduire le temps de recherche et de favoriser la mémorisation des positions des documents, par rapport aux logiciels actuels de gestion de collections d'images. Avec une représentation sous forme d'un cylindre vertical, nous avons conçu 2 visualisations : les photographies sont plaquées sur la face interne ou externe du cylindre, pour implémenter 2 métaphores courantes d'interaction, respectivement locomotion et manipulation.

3 études expérimentales ont été réalisées, avec la même méthodologie : la réalisation par les participants de 2 tâches courantes de recherche visuelle. La première étude est exploratoire, la seconde compare les 2 vues 3D plus finement et la dernière compare les vues 3D à une vue 2D classique. Dans cette dernière étude, l'usage d'un oculomètre permet de comprendre les relations entre les caractéristiques géométriques des vues et les stratégies visuelles des participants.

Les participants préfèrent majoritairement les vues 3D. Les performances n'influent pas sur les préférences ; les caractéristiques géométriques statiques ont principalement été mentionnées pour justifier les préférences. Les stratégies de parcours visuel sont davantage influencées par la caractéristique dynamique des vues (i.e., défilement horizontal), ce qui avantage la vue interne. Enfin, les vues 3D garantissent un meilleur confort visuel, en offrant un meilleur guidage du regard (principalement la vue interne) et en minimisant les saccades désordonnées observées dans la vue 2D.

Abstract

The goal of this thesis is to propose and to assess a 3D representation of large unstructured collections of images. The purpose is to improve comfort of visual search, to reduce search times and to better support memorization of documents' locations, compared with standard photo browsers. With a representation as a vertical cylinder, we designed two visualizations: photos are plastered on the inner or the outer side of the cylinder, to implement 2 common interaction metaphors, respectively locomotion and manipulation.

3 experimental studies were performed, with the same methodology: subjects performed 2 common types of visual search tasks. The first study is exploratory, the second one compares both 3D views in-depth and the last one compares 3D views with a standard 2D view. In this study, an eye-tracker helps us to understand relations between geometrical features and visual exploration strategies.

3D views are preferred by a majority of participants. Preferences are not influenced by performances; static visual features were mainly used to justified preferences. Visual exploration strategies are mainly influenced by the common dynamic visual feature (i.e., horizontal scrolling), which is an advantage for the inner view. At last, 3D views ensure a better comfort of visual search, with a better gaze guidance (especially the inner view) and by minimizing uncoordinated saccades which were observed in the 2D view.