

Thèse

présentée pour l'obtention du titre de

Docteur de l'Université Paris XII - Val de Marne

Spécialité : Génie Informatique, Automatique et Traitement du Signal

par **Matthieu VOIRY**

Étude et Conception d'un Système Automatisé de Contrôle d'Aspect des Pièces Optiques Basé sur des Techniques Connexionnistes

Commission d'examen composée de

| | | |
|----------------------|-----------------------|----------------------------|
| <i>Rapporteurs :</i> | M. Pierre CHAVEL | Directeur de Recherches |
| | M. Patrick GARDA | Professeur des Universités |
| <i>Examineurs :</i> | Mme Véronique AMARGER | Maître de Conférences |
| | M. Raymond CHEVALLIER | Maître de Conférences |
| | M. Hichem MAAREF | Professeur des Universités |
| | M. Kurosh MADANI | Professeur des Universités |
| <i>Invité :</i> | M. Joël BERNIER | Ingénieur |

LiSSI - Laboratoire Images, Signaux et Systèmes Intelligents

Université Paris12 Val-de-Marne (Bâtiment P2 - Marie Curie 2ème étage)

Remerciements

Je tiens tout d'abord à remercier Monsieur Pierre Chavel, Directeur de Recherches à IOTA et Monsieur Patrick Garda, Professeur à l'Université Pierre et Marie Curie, qui m'ont fait l'honneur d'accepter de rapporter ce travail de thèse.

Merci également à Monsieur Raymond Chevallier, Maître de Conférences à l'ENST Bretagne et à Monsieur Hichen Maaref, Professeur à l'Université d'Evry Val d'Essone, pour l'intérêt qu'ils ont porté à mon travail en acceptant de faire partie du jury.

Mes très sincères remerciements à :

- mon directeur de thèse, le Professeur Kurosh MADANI, qui a initié le projet de système automatique de contrôle d'aspect puis a bien voulu m'accueillir dans son Laboratoire. Il a su canaliser mon travail en douceur, ce dont je lui suis très reconnaissant,*
- mon co-encadreur de thèse, Madame Véronique Amarger, pour sa disponibilité et le soutien exemplaire qu'elle m'a apporté,*
- Abdennasser Chebira, Amine Chohra, et Christophe Sabourin, et mes collègues doctorant et docteurs : Saliou, Sofianne, Nadia, Moustapha, Ivan, Arash, Weiwei, Lamine, pour leur aide et l'ambiance très chaleureuse qu'ils ont toujours su faire régner au Laboratoire,*
- Monsieur François Houbre sans qui ce projet n'aurait jamais vu le jour,*
- Monsieur Joël Bernier pour sa confiance et la liberté qu'il a su me laisser,*
- Rémi, Nicolas, Caroline, Anne-Laure, Emmanuel, Benjamin, Predrag, Guillaume, Slimane, Jean-Michel, Alexandre ainsi que tous mes collègues du service Instrumentation et les personnes que j'ai cotoyées à REOSC, pour l'ambiance de travail exceptionnelle du lieu.*
- Nicolas Poirot pour son aide précieuse dans le développement de DCD, et Florient Schmitz pour les mesures sur défauts-étalons.*

Enfin, merci de tout coeur à toute ma famille ainsi qu'à tous mes amis pour leur soutien inestimable.

Mention spéciale à ma Maman Marie-Reine qui m'a toujours soutenu et fait confiance pendant ces longues années d'études, même quand mes choix n'étaient pas toujours limpides.

A mes parents.

Table des matières

| | |
|--|-----------|
| Table des matières | v |
| Liste des tableaux | ix |
| Table des figures | xi |
| Introduction générale | 1 |
| I Présentation et analyse de la problématique | 7 |
| I.1 Optiques hautes performances | 8 |
| I.1.1 Fabrication et application | 8 |
| I.1.2 Spécifications | 12 |
| I.1.3 Contrôle qualité et métrologies | 12 |
| I.1.4 Le contrôle d' aspect et les normes associées | 13 |
| I.2 Mise en oeuvre du contrôle d'aspect des optiques hautes performances . . . | 16 |
| I.2.1 Le contrôle d'aspect conventionnel et ses limites | 17 |
| I.2.2 Étude du besoin et système proposé | 22 |
| I.3 Conclusion du chapitre | 25 |
| II Système de détection des défauts d'aspect | 27 |
| II.1 Capteurs pour le contrôle d'aspect des pièces optiques | 29 |
| II.1.1 Synthèse des méthodes optiques existantes | 30 |
| II.1.2 Microscopie Nomarski | 36 |
| II.1.2.1 Principe de la mesure | 36 |

| | | |
|------------|--|-----------|
| II.1.2.2 | Théorie et réalisation technologique | 36 |
| II.1.2.3 | Avantages de la technique Nomarski | 38 |
| II.2 | Mise en œuvre de la microscopie Nomarski pour l'étude des pièces optiques | 40 |
| II.2.1 | Techniques d'exploitation d'images Nomarski existantes | 40 |
| II.2.2 | Méthode de détection des défauts d'aspect proposée | 42 |
| II.2.2.1 | Présentation | 42 |
| II.2.2.2 | Robustesse et répétabilité de la méthode | 49 |
| II.2.3 | Mesure de la rugosité | 52 |
| II.2.3.1 | Définition de la rugosité | 53 |
| II.2.3.2 | Méthodes de mesure existantes pour évaluer la rugosité | 53 |
| II.2.3.3 | Méthode de mesure proposée | 54 |
| II.2.3.4 | Répétabilité de la méthode | 57 |
| II.2.3.5 | Etalonnage de la méthode | 59 |
| II.2.4 | Implémentation et validation des méthodes | 61 |
| II.2.4.1 | Présentation du prototype | 61 |
| II.2.4.2 | Le logiciel de post-traitement | 64 |
| II.2.4.3 | Complément au prototype : recalage de cartes | 67 |
| II.2.4.4 | Validation du prototype | 70 |
| II.3 | Conclusion du chapitre | 75 |
| III | Extraction de données et classification des défauts | 77 |
| III.1 | Représentation des défauts pour le système de classification | 80 |
| III.1.1 | Paramétrisation des données : état de l'art | 81 |
| III.1.1.1 | Descripteurs de contour | 82 |
| III.1.1.2 | Descripteurs d'image | 82 |
| III.1.1.3 | Descripteurs relevant d'une heuristique | 84 |
| III.1.2 | Implantation et validation du codage par les invariants de Fourier Mellin | 84 |
| III.1.2.1 | Implantation | 85 |
| III.1.2.2 | Validation | 90 |
| III.2 | Réduction de dimension et classification neuronale | 102 |
| III.2.1 | Dimension intrinsèque d'une distribution | 102 |

| | | |
|--|--|------------|
| III.2.2 | Le « fléau de la dimension » | 106 |
| III.2.3 | Les techniques de réduction de dimension | 111 |
| III.2.3.1 | Techniques par sélection d'un sous-ensemble de paramètre | 111 |
| III.2.3.2 | Techniques projectives | 113 |
| III.2.4 | Réduction de dimension et classification neuronale des défauts . . . | 121 |
| III.2.4.1 | Etude des techniques SOM, CCA et CDA | 122 |
| III.2.4.2 | Classification neuronale : discrimination des défauts de type « poussière » | 129 |
| III.3 | Classification neuronale modulaire | 141 |
| III.3.1 | Classification modulaire : définition et rapide état de l'art | 141 |
| III.3.2 | La technique Treelike Divide To Simplify (T-DTS) | 143 |
| III.3.3 | Technique de décomposition proposée | 146 |
| III.3.3.1 | Fondement principal | 146 |
| III.3.3.2 | Estimation locale de la dimension intrinsèque | 149 |
| III.3.3.3 | Formation de groupes de dimensions intrinsèques homogènes | 151 |
| III.3.4 | Validation de la méthode sur un problème « jouet » | 153 |
| III.3.4.1 | Présentation du problème | 153 |
| III.3.4.2 | Protocole expérimental | 156 |
| III.3.4.3 | Résultats et discussion | 156 |
| III.3.5 | Classification neuronale modulaire des défauts d'aspect | 158 |
| III.4 | Conclusions du chapitre | 161 |
| Conclusion générale | | 165 |
| Annexes | | 173 |
| A Normes concernant le contrôle d'aspect | | 175 |
| B Spécification du système à concevoir | | 193 |
| C Capteurs pour la détection et la caractérisation des défauts d'aspect | | 207 |
| Bibliographie | | 251 |

Bibliographie de l'auteur

265

Liste des tableaux

| | | |
|-------|---|-----|
| I.1 | Ordres de grandeur relatifs aux dimensions de pièces optiques et à la taille minimale des défauts devant être détectés sur ces composants. | 17 |
| II.1 | Principales caractéristiques des techniques étudiées dans l'Annexe C. . . | 32 |
| II.2 | Mesures dimensionnelles d'un défaut étalon de 16 μm de diamètre imagé avec différents temps d'exposition. | 51 |
| II.3 | Mesures dimensionnelles d'un défaut étalon de 63 μm de diamètre imagé avec différents temps d'exposition. | 51 |
| II.4 | Mesures dimensionnelles d'un défaut étalon de 100 μm de diamètre imagé avec différents temps d'exposition. | 52 |
| II.5 | Répétabilité par rapport au temps de pose. | 58 |
| II.6 | Répétabilité par rapport à la position de la pupille. | 58 |
| II.7 | Mesures de rugosité (en \AA) et valeur calculée par la méthode proposée. . | 59 |
| III.1 | Paramètres des trois expériences pendant lesquelles ont été acquises les imagerie. | 92 |
| III.2 | Paramètres des trois expériences pendant lesquelles ont été acquises les imagerie. | 122 |
| III.3 | Différentes mesures caractérisant la distribution des éléments de la base de données C (défauts permanents) projetés dans les SOMs obtenues précédemment. | 127 |
| III.4 | Description des bases de données utilisées pour l'expérience. | 132 |

| | | |
|--------|---|-----|
| III.5 | Performances du classifieur pour les différentes techniques de réduction de dimensionnalité, pour une dimension finale de 2. | 138 |
| III.6 | Performances « remarquables » atteintes par le classifieur pour les techniques de réduction de dimensionnalité CCA et CDA. | 138 |
| III.7 | Description des bases de données utilisées pour le problème jouet. | 155 |
| III.8 | Description des clusters générés par notre méthode de décomposition appliquée aux données du problème jouet. | 157 |
| III.9 | Nombre de neurones dans la couche cachée et performances de généralisation des classifieurs modulaires et globaux pour le problème jouet. | 158 |
| III.10 | Description des bases de données utilisées pour le problème réel. | 159 |
| III.11 | Description des clusters générés par notre méthode de décomposition appliquée aux données du problème réel. | 159 |
| III.12 | Nombre de neurones dans la couche cachée et performances de généralisation des classifieurs modulaires et globaux pour le problème réel. | 160 |
| III.13 | Temps caractéristiques associés aux classifieurs modulaires et globaux pour les phases d'apprentissage et de généralisation. | 160 |
| C.1 | Comparaison entre capteurs à acquisition séquentielle et capteurs à acquisition parallèle. | 211 |

Table des figures

| | | |
|------|---|----|
| | Défauts d'aspect sur un miroir (mis en évidence par strioscopie). | 2 |
| | Schéma de principe du système à concevoir. | 3 |
| I.1 | Miroir destiné à un satellite d'observation. | 10 |
| I.2 | Miroir primaire de 8,2m de diamètre, destiné au VLT. | 10 |
| I.3 | Objectif, lentille et miroir, destinés à l'industrie. | 11 |
| I.4 | Plaque amplificatrice de laser. | 11 |
| I.5 | Représentation schématique des différents types de défauts d'aspect pouvant être présents sur une pièce optique. | 15 |
| I.6 | Principe du contrôle d'aspect en réflexion. | 19 |
| I.7 | Principe du contrôle d'aspect en transmission. | 19 |
| I.8 | Poste de contrôle et plaquette de défauts étalons. | 20 |
| I.9 | Schéma de principe de la solution proposée. | 24 |
| II.1 | Image d'un champ obtenue par microscopie Nomarski. | 37 |
| II.2 | Schéma optique d'un microscope DIC (extrait de (Chatterjee 2003)). | 37 |
| II.3 | Le même champ imagé à l'aide de la microscopie (a) Nomarski ; (b) Dark-Field ; (c) Bright-Field. | 39 |
| II.4 | Le même champ imagé à l'aide de la microscopie Dark-Field avec différents temps d'exposition : a) 200 ms ; b) 500 ms ; c) 1s. | 39 |
| II.5 | Image originale issue du mode DIC. | 43 |
| II.6 | Image de la Figure II.5 après prétraitement. | 44 |
| II.7 | Image pré-traitée seuillée avec $s=78$, $s=92$, $s=112$ | 45 |

| | | |
|-------|--|----|
| II.8 | Nombre de carrés 8x8 dont le pixel de niveau de gris minimum est du niveau de gris reporté en abscisse. | 45 |
| II.9 | Découpage de l'image en 12 parties de taille équivalentes. | 46 |
| II.10 | Image de la Figure II.5 après seuillage. | 47 |
| II.11 | Image de la Figure II.5 avec les contours détectés. | 48 |
| II.12 | Extraction d'images de défauts à partir d'une image Nomarski brute. | 49 |
| II.13 | Image Nomarski originale (à gauche) et représentation des défauts détectés (à droite). | 49 |
| II.14 | Graphique correspondant aux données du Tableau II.7, avec la régression linéaire correspondante. | 60 |
| II.15 | Le microscope Olympus BX-52. | 61 |
| II.16 | Flux de données lors de la mesure d'une pièce optique avec le prototype. | 62 |
| II.17 | Rectangle englobant orienté d'un défaut. | 66 |
| II.18 | Cartographie des défauts d'aspect d'une pièce optique de 100mm x 100mm réalisée avec le prototype présenté. | 70 |
| II.19 | Cartographie de rugosité d'une pièce optique de 100mm x 100mm réalisée avec le prototype présenté. | 71 |
| II.20 | Cartographie d'une pièce optique de diamètre 42mm. | 72 |
| II.21 | Cartographie de la même pièce (Figure II.20) après nettoyage. | 73 |
| II.22 | a) Cartographie de la Figure II.21 recalée par rapport à la cartographie de la Figure II.20; (b) Cartographie de la Figure II.20; (c) Mise en correspondance des cartographies représentées en a) et b). | 74 |
| III.1 | Reconstruction d'une image de papillon à partir de sa TFM pour des tailles variables de spectre. | 88 |
| III.2 | SOM obtenue avec les données de l'expérience A. | 93 |
| III.3 | SOM obtenue avec les données de l'expérience B. | 94 |
| III.4 | SOM obtenue avec les données de l'expérience C. | 94 |
| III.5 | Différentes images de défauts projetées dans le nœud n°28 de la carte représentée dans la Figure III.2. | 95 |
| III.6 | Différentes images de défauts projetées dans le nœud n°93 de la carte représentée dans la Figure III.2. | 95 |

| | | |
|--------|--|-----|
| III.7 | Répartition des différents labels définis par un expert dans la grille de la Figure III.2. | 96 |
| III.8 | Construction de la représentation dy-dx : on trace la distribution jointe des distances entre les poids (axe vertical) et des distances de grille (axe horizontal). | 98 |
| III.9 | Graphique dy-dx correspondant à la SOM obtenue avec les données de l'expérience A. (la moyenne \square et l'écart-type \diamond des dx sont également représentés pour chaque dy) | 98 |
| III.10 | Construction de la représentation curviligne à gauche la grille de la SOM ; à droite, sa représentation curviligne | 99 |
| III.11 | Représentation curviligne correspondant à la SOM de la Figure III.2 (les numéros des neurones sont identiques) | 100 |
| III.12 | A gauche, la distribution « en fer à cheval » et la surface curviligne qui la sous-tend (en gris) ; à droite, projection de la distribution sur cette surface.) | 103 |
| III.13 | Pourcentage des échantillons issus d'une distribution gaussienne qui appartiennent à la sphère de rayon 1,65 en fonction de la dimension de l'espace (tiré de (Verleysen 2001)). | 108 |
| III.14 | Ratio entre les volumes d'une hypersphère de rayon 1 et d'un hypercube de côté 1, en fonction de la dimension de l'espace (tiré de (Verleysen 2001)). | 108 |
| III.15 | Ratio entre les volumes d'hypersphères de rayons 0,9 et 1, en fonction de la dimension de l'espace (tiré de (Verleysen 2001)). | 109 |
| III.16 | Différences entre distances Euclidienne et curviligne (tiré de (Lee et al. 2000)). | 119 |
| III.17 | Représentation schématique des trois expériences menées. | 123 |
| III.18 | Representation dy-dx des trois SOMs obtenues (la moyenne \square et l'écart-type \diamond des dx sont également représentés). Gauche : SOM ; milieu : CCA ; droite : CDA. | 124 |
| III.19 | Répartition des éléments de la base de données C projetés dans la carte SOM. | 126 |

| | | |
|--------|--|-----|
| III.20 | Répartition des éléments de la base de données C projetés dans la carte SOM, après réduction de dimension par CCA. | 126 |
| III.21 | Répartition des éléments de la base de données C projetés dans la carte SOM, après réduction de dimension par CDA. | 127 |
| III.22 | Estimation de la dimension intrinsèque de la distribution de données de la base B : $DI = 5,50$ | 130 |
| III.23 | Estimation de la dimension intrinsèque de la Figure III.22 corrigée grâce à une calibration. | 131 |
| III.24 | Exemples de défauts à classer : a) défaut de type « poussière » ; b) autre type de défauts | 132 |
| III.25 | Principe des expériences menées. En haut, phase d'apprentissage, en bas, phase de test. | 133 |
| III.26 | Performances globales de classification lors des différentes expériences. . . | 136 |
| III.27 | Performances de reconnaissance de la classe NP (« autres défauts ») ; l'écart-type relatif aux 5 expériences est également représenté. | 136 |
| III.28 | Performances de reconnaissance de la classe P (« défauts de type poussière ») ; l'écart-type relatif aux 5 expériences est également représenté. . | 137 |
| III.29 | Structure générale d'un classifieur de type T-DTS (tiré de (Bouyoucef 2007)). | 143 |
| III.30 | Les courbes de Grassberger-Proccacia obtenues pour a) les données réelles ; b) les données synthétiques. | 147 |
| III.31 | Performances de reconnaissance de la classe P ; l'écart-type relatif aux 5 expériences est également représenté. | 147 |
| III.32 | Construction du voisinage du point y_i , avec $k=3$ | 149 |
| III.33 | La distribution utilisée dans le problème jouet. | 154 |
| III.34 | Répartition des classes 1 et -1 dans la partie plane de la distribution de points du problème jouet. | 155 |
| | Chaine de traitement conduisant d'une image Nomarski à la nature d'un défaut. | 168 |
| | Schéma de principe du prototype complet. | 171 |
| C.1 | Principe de la microscopie confocale, extrait de (Charron 2004b) | 212 |

| | | |
|------|---|-----|
| C.2 | Principe du codage chromatique, extrait de (STIL 2005) | 214 |
| C.3 | Schéma optique de l'éclairage de Köhler, inspiré de (Roblin 1999a) | 217 |
| C.4 | Schéma optique d'un microscope DIC, extrait de (Chatterjee 2003) | 220 |
| C.5 | Principe de la photogrammétrie. inspiré de (Charron 2004b) | 223 |
| C.6 | Interféromètre de Mirau, extrait de (Charron 2004b) | 224 |
| C.7 | Interféromètre de Michelson, extrait de (Charron 2004b) | 224 |
| C.8 | Montage d'enregistrement (deux poses successives), extrait de (Smigielski 2001a) | 228 |
| C.9 | Montage de restitution (les deux images interfèrent), extrait de (Smi- gielski 2001a) | 229 |
| C.10 | Montage optique de strioscopie, extrait de (Zanier 2002) | 231 |
| C.11 | Observation de deux plans par déplacement du capteur CCD, extrait de (Bertin 2005) | 232 |
| C.12 | Système complet de contrôle de pièces optiques (un microscope monté sur une table XY pour caractériser les défauts détectés est également représenté sur le schéma), extrait de (Rainer 1997) | 234 |
| C.13 | Schéma de la tête de mesure d'un microscope acoustique, extrait de (Attal 1998) | 240 |
| C.14 | Schéma de principe de la mesure par réflexion, extrait de (Charron 2004a) | 242 |
| C.15 | Courbes des intensités lumineuses recueillies par les capteurs n°1 et n°2 en fonction de la distance à la pièce mesurée, extrait de (Charron 2004a) | 242 |
| C.16 | Principe de l'imagerie IIM, extrait de (Schwarz et al. 2003) | 244 |
| C.17 | Montage optique associé à la technique d'imagerie IIM, extrait (Schwarz et al. 2003) | 245 |

Introduction générale

Ce travail de recherche a été mené dans le cadre d'un contrat CIFRE liant le Laboratoire Images, Signaux et Systèmes Intelligents (LISSI / EA 3956 de l'Université PARIS XII – Val de Marne) à la Société REOSC. De ce fait, au travers des différentes études réalisées, il tente d'amener une réponse à une problématique industrielle sensible pour la société partenaire, et vise une solution pouvant être transposable à des situations similaires.

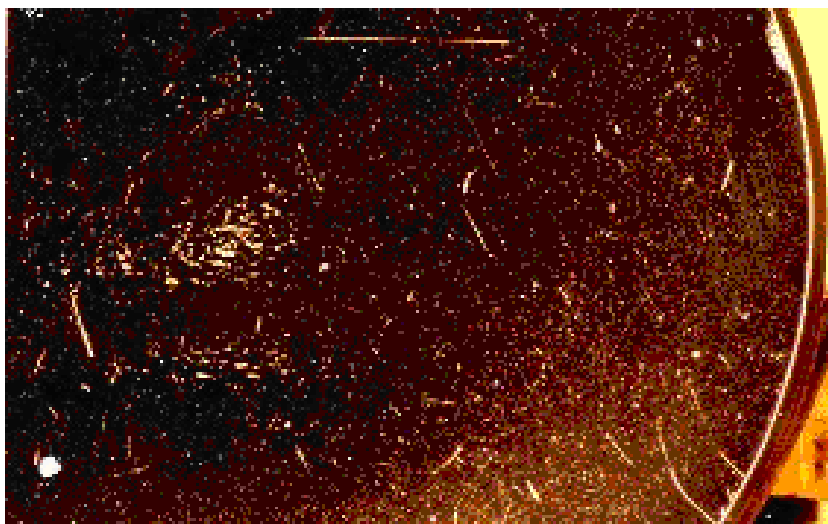
La société REOSC (Recherche et Études en Optique et Sciences Connexes) fut fondée en 1937 par un groupe de scientifiques de l'Institut d'Optique de Paris incluant notamment Henri Chrétien (télescope Ritchey-Chrétien) et Charles Fabry (interféromètre Fabry-Pérot). Rachetée en 1999, REOSC fait maintenant partie de la division Optronique et Systèmes Aéroterrestres de la Société SAGEM Défense Sécurité, elle-même filiale du groupe SAFRAN, groupe international de haute technologie.

Le site de Saint Pierre du Perray, qui accueille aujourd'hui REOSC, est un centre de R&D ainsi que de fabrication de prototypes et de petites séries. Des optiques de très hautes performances y sont réalisées pour des programmes de recherche français et internationaux. Avec la gamme d'optiques hautes performances représentées par REOSC, SAFRAN est un des leaders mondiaux en conception, étude et réalisation d'équipements opto-mécaniques de pointe pour tous types d'applications :

- Optiques spatiales : télescopes, objectifs, filtres. . .
- Optiques pour l'astronomie : grands miroirs et miroirs segmentés. . .
- Défense : optiques de haute résolution pour la reconnaissance aérienne. . .
- Composants optiques : laser, UV, X Ray. . .
- Couches minces optiques : photolithographie, traitements métalliques, gravure. . .

Le symbole le plus significatif de cette excellence reste sans doute le jeu de miroirs primaires du célèbre VLT (Very Large Telescope), réalisé à Saint Pierre à la fin des années 90.

Dans un cadre industriel de production de grande série, on mise généralement sur la maîtrise de la chaîne de fabrication afin de réduire les contrôles qualité et les rebuts associés. Les activités de REOSC, quant à elles, impliquent la fabrication de petites séries à forte valeur ajoutée et nécessitent donc un contrôle unitaire des composants fabriqués, garantissant la qualité du produit fini. Lors de cette phase de contrôle, un des paramètres inspectés est l'aspect des pièces optiques : on examine la gravité des « rayures » et des « piqûres » éventuellement présents sur les miroirs.



Défauts d'aspect sur un miroir (mis en évidence par strioscopie).

Il est en effet pertinent de noter que les imperfections de surface, générées pendant les différentes phases de fabrication d'un composant, engendrent des lumières parasites qui peuvent dégrader ses performances optiques. Dans le but de limiter leur impact à un niveau acceptable (compatible avec l'application visée), le nombre et l'étendue de ces défauts sont soumis à une spécification, exprimée par le client par le biais d'une norme.

Malheureusement, il existe pour l'heure peu d'équipements automatisés de contrôle permettant le diagnostic des défauts d'aspect, et les systèmes existants ne sont pas, pour diverses raisons, adaptés aux conditions de production réelles. Les inspections restent donc essentiellement visuelles, ce qui n'est pas sans poser de sérieux problèmes. Le contrôle vi-

suel s'avère en effet insuffisant en termes d'exhaustivité et de sensibilité; il représente d'autre part une tâche pénible demandant une grande expertise et beaucoup de temps. Le caractère critique de cette métrologie est aggravé par l'apparition de besoins nouveaux liés à des marchés émergents (typiquement celui des optiques laser ou dédiées à la micro-lithographie). Ces applications mettent en jeu des optiques de plus en plus grandes et/ou exigent la détection de défauts de plus en plus petits. De plus, dans le cas des optiques laser, les défauts d'aspect sont des critères fonctionnels prépondérants puisqu'ils peuvent mettre en péril l'intégrité physique de la pièce. A ceci, s'ajoutent des contraintes de coûts et de cadences incompatibles avec le contrôle d'aspect traditionnel ainsi que des exigences de maîtrise du procédé (détection des dérives) nécessitant une meilleure caractérisation des défauts d'aspect.

Ces différentes considérations ont permis d'identifier un besoin spécifique qui conduit naturellement à la problématique industrielle au cœur de cette thèse : l'étude et la conception d'un système automatisé de contrôle d'aspect des pièces optiques.

Afin de doter le système d'une autonomie suffisante pour permettre une automatisation des contrôles, nos recherches ont porté sur le choix des capteurs utilisés et sur l'« intelligence » associée aux traitements des données qu'ils fournissent. Schématiquement, le système que nous nous proposons de mettre en place, se présente de la manière suivante :

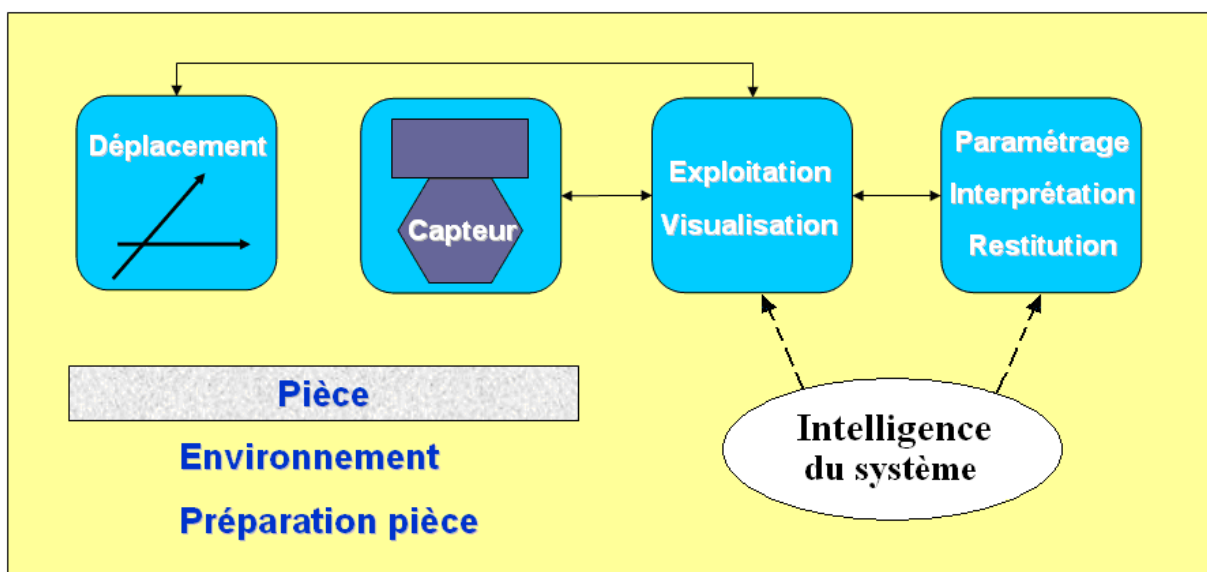


Schéma de principe du système à concevoir.

On peut distinguer dans ce schéma plusieurs entités distinctes :

- La pièce optique à contrôler, dont les caractéristiques sont influencées par l’environnement et une préparation donnée.
- Le bloc capteur, qui permet au système de mesurer les grandeurs physiques servant à caractériser la pièce.
- Le mécanisme de déplacement, qui impose des mouvements au bloc capteur et permet de la sorte des mesures sur toute la surface de la pièce.
- Le module d’exploitation et de visualisation, qui exploite ces mesures afin d’identifier les défauts d’aspect présents sur la pièce et de déterminer leurs caractéristiques, pour construire finalement une cartographie de celle-ci.
- Le module de paramétrage, d’interprétation et de restitution, qui analyse les données fournies par le module précédent pour dégager un diagnostic de conformité de la pièce par rapport à une norme donnée en paramètre.

La première phase d’étude du système à concevoir correspond à une analyse de ses différents composants et des performances qu’il doit atteindre. Elle permet d’identifier deux aspects particulièrement complexes du problème qui constituent de véritables défis. Le premier consiste à détecter les défauts présents sur la pièce optique contrôlée et à en extraire des informations de bas niveau (position, taille et représentation image), ce dans un temps raisonnable, c’est-à-dire compatible avec les cadences de fabrication. La deuxième phase réside dans l’utilisation de la représentation obtenue pour assigner le défaut à une classe prédéfinie et ainsi déterminer sa nature. Une telle étape de classification est nécessaire, entre autres, à cause de l’environnement et de la préparation de la pièce, qui introduisent un certain nombre d’artefacts ne devant pas être confondus avec des défauts réels pour garantir la fiabilité du diagnostic.

Le premier des défis à relever concerne en fait le compromis entre la vitesse du contrôle et la précision qu’il doit atteindre. Ce problème délicat nécessite un soin particulier dans le choix des capteurs employés par le système. Les différentes recherches entreprises dans ce sens ont permis de dégager une solution satisfaisante qui met en œuvre deux capteurs différents travaillant à deux échelles distinctes. Dans ce travail, nous nous sommes cependant presque exclusivement intéressé au capteur de « précision » qui permet de caractériser les défauts d’aspect. Après une étude approfondie, notre choix s’est porté sur

un capteur utilisant la microscopie Nomarski. Mais aucune des méthodes existantes d'exploitation d'images obtenues par ce biais n'était satisfaisante pour l'application visée. Des outils permettant la détection des défauts d'aspect et l'étude de la rugosité ont donc été développés et cet ensemble de techniques a été intégré dans un prototype permettant le contrôle automatique de petits composants optiques. D'autre part, la méthode de détection proposée permet d'isoler une « imagerie » d'un défaut donné, qui constitue la donnée de départ pour la phase de classification.

Le deuxième défi consiste à reproduire un niveau d'expertise suffisant pour discriminer les différents types d'imperfections présents sur les pièces optiques et permettre ainsi une mise en oeuvre pertinente des normes. Afin de remplir cette tâche complexe, nous avons choisi d'utiliser les réseaux de neurones artificiels. Ils sont en effet communément utilisés pour des classifications difficiles car ils possèdent des caractéristiques attrayantes pour ce type d'applications (Zhang 2000) (Egmont-Petersen, de Ridder, and Handels 2002).

Cependant, pour mettre en oeuvre des techniques neuronales, un codage préalable des données est nécessaire. Les données brutes à traiter (des images) sont en effet de tailles importantes et variables, ce qui interdit de les exploiter directement. Pour notre système, nous avons retenu une méthode de codage basée sur la transformée de Fourier-Mellin.

Comme dans une majorité de cas d'applications réelles, la représentation des défauts obtenue est d'une taille assez importante. Ceci peut entraîner d'importants problèmes, connus sous le nom générique de « fléau de la dimension » (Yeh et al. 2005), qui peuvent être limités par des méthodes de réduction de la dimensionnalité. Nous avons mis en oeuvre certaines d'entre elles, Self Organizing Maps (SOM), Curvilinear Components Analysis (CCA) et Curvilinear Distance Analysis (CDA) sur des données réelles afin d'évaluer leur impact probable sur une tâche de classification. Dans un deuxième temps, ces techniques ont été intégrées dans un système neuronal de discrimination des défauts de type « poussière ».

Afin d'obtenir une classification plus performante en termes de taux de généralisation et de temps d'exécution, un traitement par réseaux de neurones modulaires a été envisagé. Nous nous sommes intéressés plus précisément à des modèles utilisant une décomposition en sous-problèmes locaux. Ce type de structures tire du concept de modularité un certain nombre d'avantages précieux, notamment en ce qui concerne les temps d'apprentissage

et de classification. Malheureusement le mécanisme de décomposition qui les sous-tend reste flou et parfois même peu pertinent. Nous proposons dans cette thèse une nouvelle méthode de décomposition basée sur les dimensions intrinsèques locales, permettant de modéliser la structure topologique du problème étudié. Le problème de discrimination des poussières a été également traité par un classifieur neuronal modulaire utilisant cette approche.

Ce manuscrit se structure de la manière suivante :

Le chapitre I expose plus en détails la problématique globale de la thèse, les différentes difficultés que sous-tend cette problématique et présente les solutions que nous leur avons apportées.

Le chapitre II décrit le système de détection des défauts d'aspect dans son ensemble : après une brève description des différents capteurs envisageables pour le contrôle d'aspect des pièces optiques, nous étudions en détail le capteur retenu, basé sur la microscopie Nomarski. Nous présentons alors les techniques développées permettant l'exploitation des images fournies par ce capteur ainsi que le prototype de contrôle automatique qui les met en œuvre.

Le chapitre III traite de l'extraction, du traitement et de la classification des données issues des « imageries » de défauts : nous y abordons tout d'abord la problématique de la paramétrisation des données avant de présenter la solution choisie qui utilise la transformée de Fourier Mellin. Nous discutons ensuite du concept de réduction de dimension, en introduisant la notion de dimension intrinsèque d'une distribution et de « fléau de la dimension ». Après avoir donné un état de l'art des techniques de réduction de dimension, nous étudions l'intérêt de trois d'entre elles (SOM, CCA et CDA) dans le cadre d'une classification neuronale mettant en œuvre nos données réelles. Enfin, nous abordons le domaine de la classification neuronale modulaire par le biais de la technique Treelike Divide To Simplify (T-DTS). Nous exposons ensuite notre technique de décomposition en sous-problèmes de dimension intrinsèque homogène ainsi que sa validation sur un problème jouet ; une classification neuronale modulaire de nos « imageries » utilisant cette approche est également discutée.

Finalement, nous concluons ce travail et en présentons les différentes perspectives.

Présentation et analyse de la problématique

Sommaire

| | | |
|------------|---|-----------|
| I.1 | Optiques hautes performances | 8 |
| I.1.1 | Fabrication et application | 8 |
| I.1.2 | Spécifications | 12 |
| I.1.3 | Contrôle qualité et métrologies | 12 |
| I.1.4 | Le contrôle d' aspect et les normes associées | 13 |
| I.2 | Mise en oeuvre du contrôle d'aspect des optiques hautes performances | 16 |
| I.2.1 | Le contrôle d'aspect conventionnel et ses limites | 17 |
| I.2.2 | Étude du besoin et système proposé | 22 |
| I.3 | Conclusion du chapitre | 25 |

Pour développer un système automatique de contrôle d'aspect des pièces optiques hautes performances, il est nécessaire d'avoir une bonne connaissance de ces composants spécifiques. La manière dont ils sont fabriqués et les applications auxquelles ils sont destinés sont particulièrement intéressantes. En effet, les défauts recherchés peuvent être générés à divers stades du processus de fabrication, qui constitue par ailleurs le contexte opérationnel du système à concevoir. D'autre part, la nature des applications visées influe directement sur la qualité des optiques fabriquées à travers des spécifications exprimées selon le formalisme de normes. Par ce biais, elles ont un impact sur la métrologie des pièces optiques et permettent donc de mettre en évidence une partie des contraintes à respecter dans la mise en œuvre du système de contrôle d'aspect.

Ce chapitre présente les optiques hautes performances et leurs spécifications, ainsi que le contrôle d'aspect et ses limites actuelles. Avec cette étude comme support, une analyse des caractéristiques du système à concevoir a été menée. Celle-ci permet d'isoler deux problèmes majeurs traités dans cette thèse. Le principe de la solution proposée est exposé à la fin de ce chapitre.

I.1 Optiques hautes performances

De manière générale, une pièce optique est un composant en verre, métal ou céramique qui assure une fonction optique en modifiant un flux de lumière incidente par réflexion ou déflexion sur un ou plusieurs dioptries. On parle d'optiques hautes performances lorsque leurs caractéristiques sont très pointues (en termes de forme et de rugosité principalement) et/ou que leur taille est inhabituellement grande.

I.1.1 Fabrication et application

Un tel composant est obtenu après plusieurs étapes de fabrication :

- Le cubage : on part d'un bloc de matière que l'on met en forme grossièrement, afin d'obtenir une pièce aux dimensions extérieures adéquates.
- La mise en courbe : on s'intéresse alors aux surfaces constituant les dioptries de la pièce, dont on approche la forme finale en une première approximation.

Ces deux premières étapes d'ébauche utilisent les mêmes outils : meules diamantées et machines d'usinage variées.

- Le doucissage : cette phase a pour but de corriger la forme et de supprimer les défauts d'ébauche. A cette fin, on utilise des polissoirs traditionnels associés à des liquides de polissage (slurry) dont les grains sont de plus en plus fins. A ce moment, le miroir n'est pas encore réfléchissant ou transparent. . .
- Le polissage : on continue à diminuer la taille des grains utilisés; on obtient, de proche en proche une surface plus conforme au dioptre recherché. D'autre part, la rugosité baisse et le miroir devient réfléchissant. Cette étape peut mettre en œuvre d'autres moyens de polissage tels que la P.A.O. (Polissage Assisté par Ordinateur).
- La finition : on élimine les derniers défauts de formes et on atteint la rugosité désirée. Pour ce faire, sont utilisés la P.A.O., l'usinage ionique et d'autres procédés dédiés.

Ces deux dernières étapes sont essentiellement itératives, ce qui explique en grande partie la difficulté de fabrication des optiques hautes performances dont la réussite dépend complètement de la convergence du process. On verra plus loin que cela donne également un rôle prépondérant à la métrologie.

- Le traitement : cette dernière étape consiste à déposer sur la pièce optique, par évaporation, une couche mince dont le rôle est de donner à la surface concernée des propriétés optiques spéciales (polarisation de la lumière, filtrage passe-bande du spectre lumineux, très grand coefficient de réflexion. . .)

Les pièces optiques ainsi fabriquées couvrent un large éventail d'applications dans des domaines divers :

- Le spatial (Figure I.1) : objectifs, télescopes et filtres embarqués dans les satellites, destinés essentiellement à l’astronomie et à la défense.



Figure I.1 : Miroir destiné à un satellite d’observation.

- L’astronomie (Figure I.2) : grands miroirs et miroirs segmentés pour les observatoires.

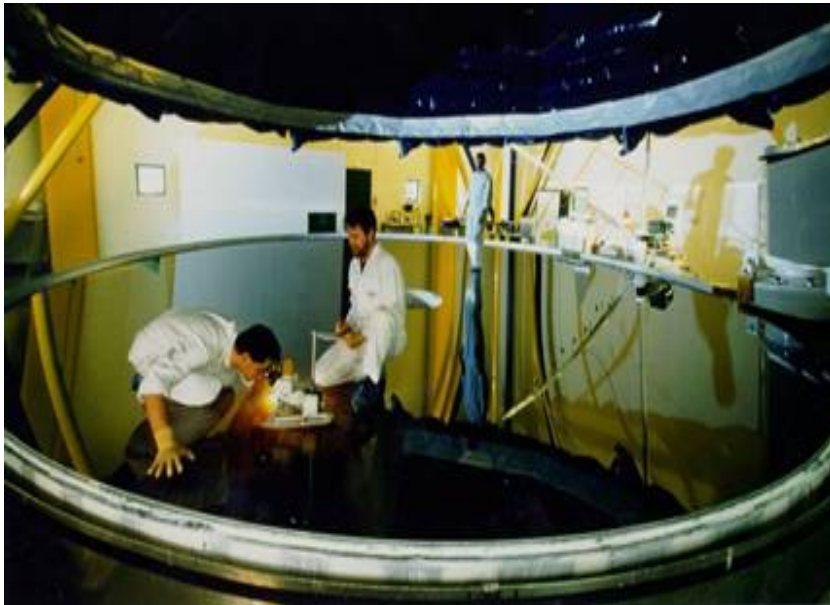


Figure I.2 : Miroir primaire de 8,2m de diamètre, destiné au VLT.

- L'industrie (Figure I.3) : miroirs, objectifs et collecteurs, en particulier pour des applications de microlithographie (qui requièrent des optiques nombreuses et de qualité très élevée).



Figure I.3 : Objectif, lentille et miroir, destinés à l'industrie.

- L'énergie (Figure I.4) : optiques de transport et d'amplification, pour les lasers, les rayonnements UV et les rayons X.



Figure I.4 : Plaque amplificatrice de laser.

Les illustrations des figures I.1, I.2, I.3 et I.4 montrent des optiques fabriquées au sein de la Société REOSC, présente sur tous ces marchés.

I.1.2 Spécifications

Compte tenu du large spectre d'applications couvert par les optiques hautes performances, les pièces fabriquées sont très variées. Afin d'assurer l'adéquation des caractéristiques d'une pièce donnée avec l'usage auquel elle est destinée, un certain nombre de contraintes sont fixées. Ces spécifications, déterminées par le client en fonction de ses besoins, portent sur quatre points principaux :

- Le dimensionnement : il fixe les différentes cotes que doit respecter la pièce, essentiellement pour permettre son intégration dans le système opto-mécanique qui doit la recevoir.
- La forme : cette caractéristique est fondamentale puisqu'elle détermine les propriétés optiques de la pièce et influe donc directement sur ses performances opérationnelles.
- La rugosité : cette propriété physique (présentée en détail dans le paragraphe II.2.3.1, p.53) revêt une grande importance puisqu'elle entraîne une perte de flux et dégrade le rapport signal sur bruit observé en sortie de la pièce.
- L'aspect : les piqûres et les rayures, générés lors du processus de polissage, engendrent de la lumière parasite, et d'autres problèmes plus spécifiques (risque de détérioration des pièces optiques laser) ; leur nombre doit donc être limité et en accord avec l'application visée.

Sont souvent ajoutées à ces spécifications de base des précisions quant à la matière utilisée pour fabriquer le miroir ainsi que les traitements de surface qui doivent lui être appliqués. Il faut également noter qu'elles s'accompagnent de tolérances associées avec chacun des paramètres spécifiés.

I.1.3 Contrôle qualité et métrologies

Le contrôle qualité a pour but de garantir que la pièce fabriquée répond aux spécifications qui lui sont associées. Puisqu'il valide les performances finales de la pièce et donc conditionne directement son acceptation par le client, ce contrôle est entrepris lorsque la pièce est achevée. Cependant, certains paramètres sont régulièrement contrôlés du-

rant la fabrication, afin de guider le process de polissage et d'éviter des dérives qui ne pourraient être corrigées.

Différentes métrologies sont mises en œuvre pour vérifier la conformité des pièces par rapport à leurs spécifications et réaliser ces contrôles périodiques :

- Les paramètres dimensionnels sont contrôlés à l'aide d'outils classiques (règles graduées, pieds à coulisse...), de colonnes à palpeur ou de machines de mesure 3D.
- La rugosité est mesurée grâce à des techniques de microscopie interférométrique, Nomarski, à force atomique, ou à des profilomètres mécaniques.
- Diverses méthodes d'interférométrie sont mises en œuvre pour accéder à la forme de la pièce (le type d'interféromètre et sa configuration dépendent des caractéristiques de celle-ci). On utilise également des machines de mesure 3D et des capteurs de type Shack-Hartmann.
- Pour ce qui est du contrôle d'aspect, les moyens de contrôle disponibles sont beaucoup plus limités. Pour comprendre pourquoi, nous allons étudier par le détail cette problématique centrale pour la thèse dans les prochains paragraphes.

L'amélioration de ces métrologies est un challenge constant pour l'industrie optique : elles doivent en effet répondre à des exigences croissantes en termes de précision et de temps de contrôle, ce afin de garantir aux clients la qualité de pièces toujours plus complexes. D'autre part, dans le cas des optiques hautes performances, elles jouent un rôle critique puisqu'elles conditionnent largement la convergence du processus de fabrication en guidant les corrections à effectuer entre chaque itération par la connaissance de divers paramètres concernant le composant.

Au final, le coût des différents contrôles représente environ 20% du prix total d'une pièce optique fabriquée par REOSC (ce taux varie de 5 à 50% selon les pièces) : la métrologie est donc un enjeu économique majeur.

I.1.4 Le contrôle d' aspect et les normes associées

Le contrôle d'aspect consiste à détecter les différents défauts d'aspect présents sur une pièce optique et à vérifier qu'ils ne compromettent pas les performances qu'elle doit atteindre.

Un défaut d'aspect est une imperfection de surface ; la norme ISO 9211-1 de 1994 donne une typologie des différents défauts que l'on peut trouver sur les pièces optiques avant traitement (voir Figure I.5) :

1. Ecailles : Eclats de verre, généralement sur la périphérie d'une pièce.
2. Egrénures : Petits éclats qui se suivent sur les bords de la pièce.
3. Rayures : Sillons continus ou discontinus à flancs dépolis ou éclatés, dont la largeur et la profondeur peuvent être irrégulières.
4. Frayures : Fines rayures continues.
5. Filandres : Défauts linéaires superficiels laissés par le polissoir, visibles seulement sous certaines incidences.
6. Filoches : Défauts linéaires superficiels dus à un frottement, généralement plus fins que les filandres.
7. Piqûres : Défauts ponctuels de surface.
8. Gris : Concentration de piqûres qui trahit le fait que la surface est insuffisamment polie après la phase de douci.
9. Oxyde : Altération chimique de la surface du verre en cours de polissage.

A ces types de défauts se rajoutent des défauts après traitement :

- Taches : Décoloration inégale localisée sur la surface, provoquée par des changements produits par une action chimique.
- Comètes : Défauts ressemblant aux corps célestes du même nom (noyau suivi d'une queue voilée).
- Projections : Défauts résultant du dépôt sur la surface du substrat de particules de matière provenant du creuset chaud, lors de traitements par évaporation.
- Traces de nettoyage : Traces pouvant ressembler à des taches ou à des raies, dues à un nettoyage insuffisant (défauts se retrouvant également sur des pièces non-traitées).

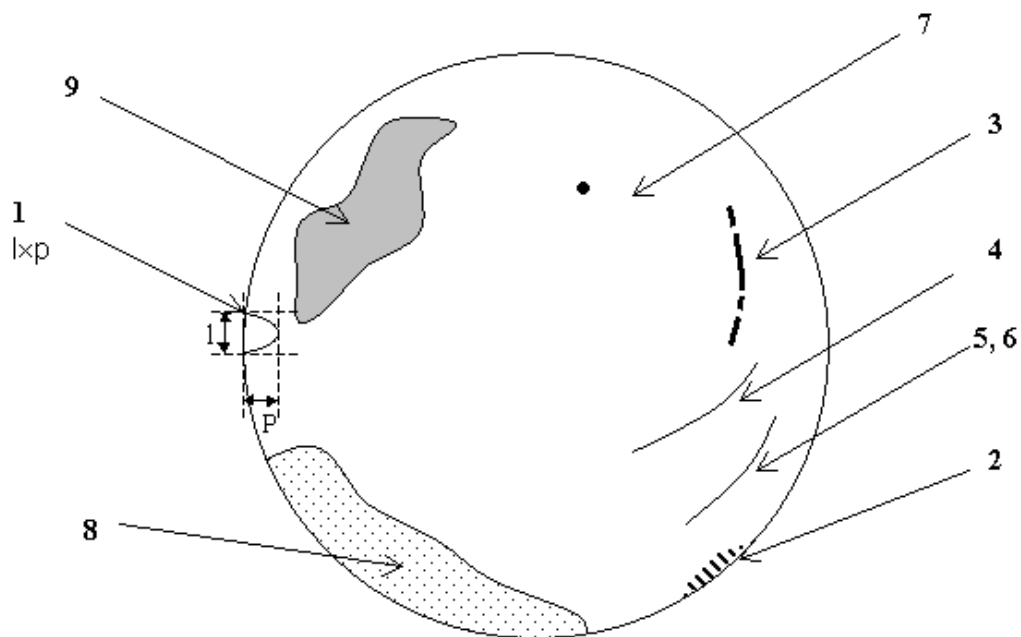


Figure I.5 : Représentation schématique des différents types de défauts d'aspect pouvant être présents sur une pièce optique.

On ne rapporte ici que les défauts d'aspect après traitement les plus communs. Pour une description plus complète, se reporter à la norme ISO 9211-1 :1994.

Rappelons ici que ces différentes imperfections de surface produisent des lumières parasites indésirables à des degrés divers selon le type de pièce fabriquée : par exemple, seuls des défauts importants peuvent être gênants pour les optiques des grands télescopes, alors que les pièces utilisées par la microlithographie doivent être quasiment exemptes de défauts d'aspect. D'autre part, dans certains cas, comme celui des optiques laser, les défauts d'aspect revêtent une importance particulière car ils ont un impact fonctionnel critique. En effet, ces défauts, lorsqu'ils sont soumis à un important flux laser, créent des « points chauds » qui peuvent entraîner la destruction pure et simple du composant...

Pour limiter ces effets néfastes ou tout du moins pour qu'ils soient compatibles avec l'application à laquelle est destinée la pièce optique considérée, le nombre et la taille des défauts d'aspect doivent être maîtrisés.

Afin de formaliser les spécifications en aspect, un certain nombre de normes ont été mises en place. Ces différentes normes sont présentées en détails et analysées dans l'Annexe A. A l'heure actuelle, l'industrie optique utilise presque exclusivement deux d'entre elles, la

norme MIL et la norme ISO. Même si l'utilisation de telles normes n'est pas toujours aisée, elles constituent le seul outil de dialogue entre client et fabricant, qui permet de garantir la qualité de la pièce et son adéquation avec l'application visée.

L'analyse de ces deux normes montre que la confrontation d'une pièce optique à une spécification d'aspect exprimée à l'aide de l'une d'entre elle, nécessite de détecter toutes les imperfections de surface présentes sur la pièce étudiée et de déterminer différents paramètres les caractérisant :

- Leur localisation : la connaissance de la position des défauts sur la pièce permet d'une part d'appliquer le critère de « concentration » imposé par les normes et d'autre part de discriminer les défauts de type « écaille » ou « égrénure » qui doivent éventuellement être pris en compte séparément.
- Leurs dimensions : les paramètres morphologiques sont fondamentaux pour l'application des normes ISO et MIL puisque dans ce cadre, la largeur, la longueur et l'aire d'un défauts sont les seuls critères qui permettent d'évaluer sa gravité (et donc son acceptabilité). Il permet par ailleurs de faire la distinction entre défaut ponctuel et longiligne.
- Leur nature : on doit être capable de faire la distinction entre défauts avant et après traitement puisque ces deux types de défauts sont soumis à deux spécifications différentes. D'autre part, comme nous allons le voir plus loin, il est primordial de distinguer les défauts permanents des défauts qui peuvent être corrigés.

I.2 Mise en oeuvre du contrôle d'aspect des optiques hautes performances

Si le nombre de paramètres dont la connaissance est nécessaire pour appliquer les normes qui sous-tendent le contrôle d'aspect est plutôt limité, ce contrôle pose dans les faits de sérieux problèmes.

I.2.1 Le contrôle d'aspect conventionnel et ses limites

Premièrement, le fait que la taille des défauts recherchés (de l'ordre de la dizaine de microns) soit très petite par rapport aux dimensions du substrat qui les porte (de l'ordre du mètre) est problématique. En effet, la précision nécessaire pour étudier de tels défauts semble incompatible avec l'étude de surfaces importantes en un temps raisonnable. Ceci explique que les quelques systèmes existants (proposés dans le commerce ou présentés dans des brevets) ne sont pas capables de concilier ces deux aspects : soit ils ne prennent en charge que des pièces de petite taille (Tomie 2003) (Knollenberg, Hoxie, and Clinton 1996) (Imaino et al. 1999) , soit ils se concentrent sur des défauts bien plus grands (Kieckhäfer 1996) (Tanaka et al. 1995) (Kihira 1997).

Ce rapport désavantageux est illustré par le Tableau I.1 ; y sont présentés, pour différents domaines d'applications, des ordres de grandeur relatifs aux dimensions des pièces optiques utilisées et à la taille minimale des défauts devant être détectés sur ces composants.

| Application | Very Extreme Ultraviolet (VEUV) | Extreme Ultraviolet (EUV) | Laser | Optique Conventionnelle | Astronomie |
|--|---------------------------------|---------------------------|------------------|-------------------------|-------------------|
| Taille des pièces (ordre de grandeur) | 100 cm ² | 100 cm ² | 1 m ² | 0,1 m ² | 10 m ² |
| Taille minimale des défauts à détecter | 50 nm | 2 μm | 5-30 μm | 50 μm | 100 μm |

Table I.1 : Ordres de grandeur relatifs aux dimensions de pièces optiques et à la taille minimale des défauts devant être détectés sur ces composants.

La deuxième difficulté majeure tient à l'expertise nécessaire à la discrimination des différents types de défauts qui peuvent être détectés sur les pièces optiques.

Outre la discrimination des défauts avant et après traitement imposée par les normes, on doit également distinguer les « vrais » défauts des « faux » défauts. Le contrôle d'aspect est en effet mené la plupart du temps dans l'atelier, ou tout du moins dans un environnement à la propreté pas forcément maîtrisée. Par conséquent, un nombre important d'imperfections de surface pouvant être corrigées (tels que poussières et traces de net-

toyage, entre autres) sont présents sur les pièces à contrôler. Evidemment, ces « faux » défauts doivent être écartés lors de l'analyse présidant à la confrontation de la pièce avec ses spécifications, sous peine de refuser une pièce en réalité conforme. Enfin, il est possible d'isoler certaines classes de défauts d'aspect constituant de véritables signatures révélant des problèmes précis dans le process de polissage. Ceci est bien évidemment souhaitable pour améliorer le contrôle du procédé de fabrication et ainsi augmenter la qualité et le rendement de la production.

Si certains systèmes permettent de discriminer les poussières et certaines traces de nettoyage (Chen, Zeng, and Akbulut 2004) (Knollenberg 1990) (Imaino and Latta 1999) (Sekine et al. 2000), c'est au prix d'une complexification considérable des capteurs. Ces solutions mettent en effet à profit le comportement particulier des poussières vis-à-vis de la polarisation et de la diffusion ; cela implique par ailleurs que le diagnostic des défauts d'aspect proposé n'est pas très détaillé. La capacité d'expertise des défauts, pourtant fondamentale, est donc pour l'instant inaccessible aux systèmes automatiques de contrôle.

Au final, les deux difficultés majeures que nous venons d'évoquer expliquant que le contrôle d'aspect ne soit, dans l'état actuel des choses, pas du tout automatisé, on se base exclusivement sur l'expertise humaine.

Pour effectuer un contrôle d'aspect, l'opérateur dispose de différentes sources de lumière : spot halogène, lampe fibrée, et lampe Nachet, dont l'utilisation est fonction de la taille et du type de pièce considérée. Un premier examen permet de déceler les écailles de la pièce. L'homogénéité de la matière est ensuite contrôlée en éclairant la pièce par ses tranches, ce qui révèle d'éventuelles « bulles ». Vient enfin la partie la plus délicate du contrôle : trouver les défauts d'aspect sur la surface du miroir. Pour mettre en évidence ces défauts, l'opérateur regarde la pièce sous un angle proche de celui de la réflexion spéculaire (Figure I.6) - ou de la transmission directe (Figure I.7) - du spot de lumière. Le défaut diffusant de la lumière, il apparaît alors brillant sur fond noir.

En général, le contrôle est effectué dans l'obscurité et on dispose un fond noir derrière la pièce pour obtenir un meilleur contraste des défauts. On peut au besoin déplacer légèrement la direction du faisceau de la lampe : cela permet de mettre en évidence certains défauts de façon dynamique et rend la méthode encore plus sensible.

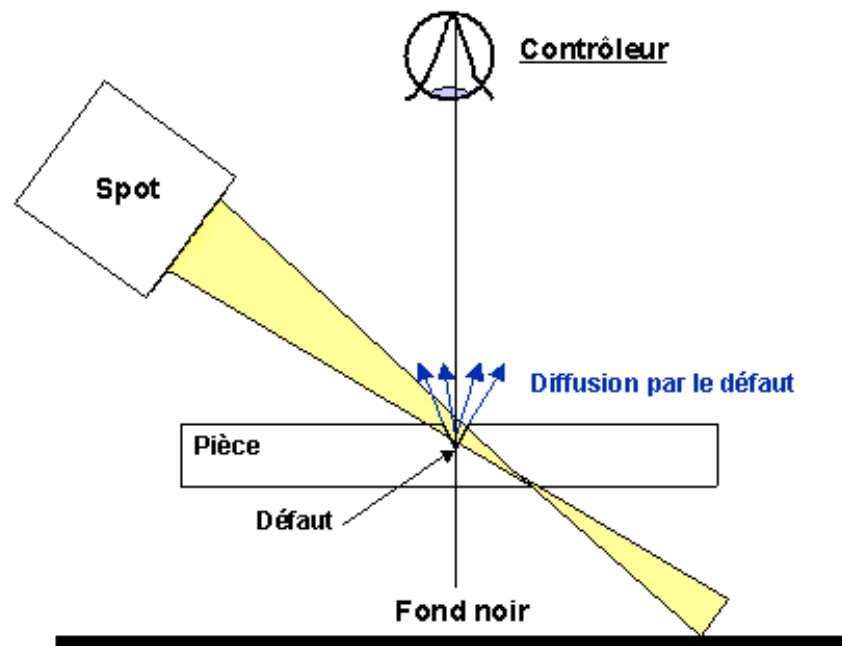


Figure I.6 : Principe du contrôle d'aspect en réflexion.

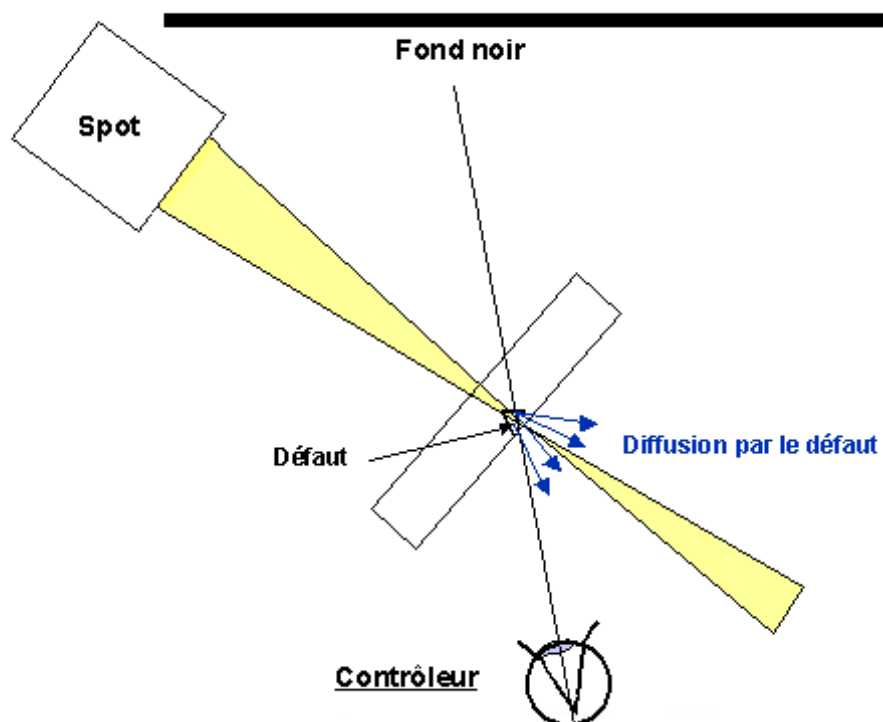


Figure I.7 : Principe du contrôle d'aspect en transmission.

Le contrôle s'effectue autant que faire se peut sous une hotte ou un plafond soufflant pour éviter que la pièce ne soit contaminée par des poussières ambiantes qui rendraient le contrôle fastidieux et nuiraient à sa fiabilité (dilemme poussière/défaut). Dans le même esprit, on dispose d'une source d'air sous pression, de compresseurs, ainsi que de différents solvants : éthanol, isopropanol, acétone, pour nettoyer la pièce à contrôler et pour ponctuellement trancher le dilemme évoqué précédemment (si un défaut disparaît au nettoyage, ce n'en était pas un !). De plus, on opère généralement en portant des gants.

Enfin, les contrôleurs utilisent des règles graduées pour positionner les défauts ainsi qu'une binoculaire x100 avec un réticule calibré ou une plaquette de défauts étalons pour les dimensionner (voir Figure I.8).

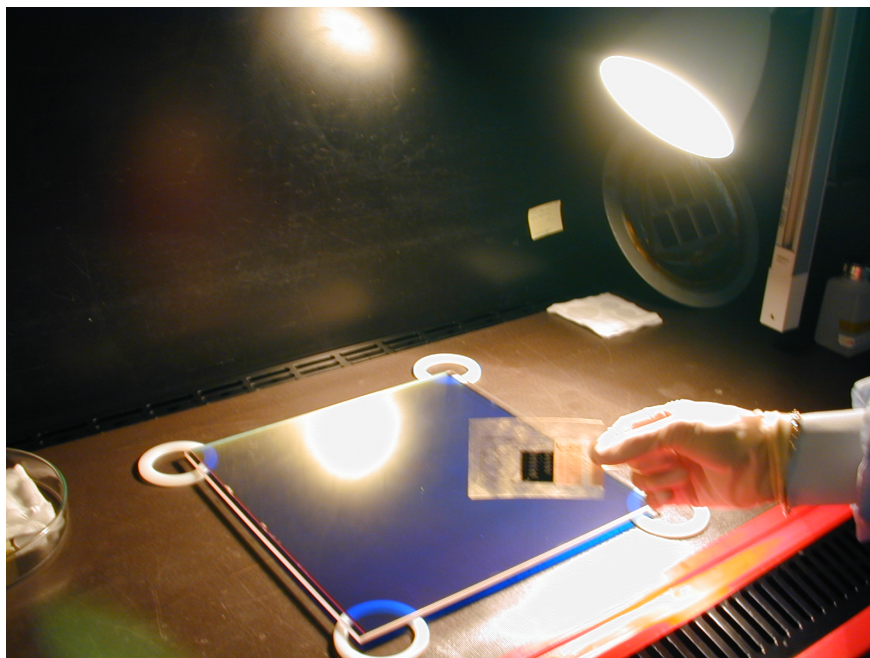


Figure I.8 : Poste de contrôle et plaquette de défauts étalons.

Il convient ici d'insister sur le niveau d'expertise nécessaire à ce type de contrôle : un contrôleur doit être formé plusieurs mois avant d'être capable de détecter des défauts d'aspect de l'ordre du micron avec une fiabilité satisfaisante. Pour ce qui est de la discrimination efficace des différents types de défaut, l'apprentissage est encore plus long. . . D'autre part, ces contrôles sont fastidieux : il faut environ 3 heures pour effectuer le contrôle finale d'une pièce d'un demi mètre carré.

Cette solution, basée sur une inspection visuelle par un opérateur, est utilisée de façon exclusive par l'industrie optique mais s'avère malheureusement loin d'être pleinement satisfaisante.

Premièrement, elle est entièrement fondée sur l'expertise humaine et souffre donc des limitations classiques inhérentes à ce type de techniques. Tout d'abord, le contrôle d'aspect est une tâche répétitive qui nécessite une grande minutie ; il entraîne donc une importante fatigue de l'opérateur. Cette lassitude entraîne quelques diminutions d'attention pendant la durée (assez importante) du contrôle ; il lui est donc difficile d'être exhaustif. Ce constat est aggravé par le fait que, les défauts recherchés étant de plus en plus petits, on s'approche du seuil critique de détection par l'être humain. Enfin, un expert est par nature subjectif : les appréciations portées par deux personnes différentes sur un même défaut d'aspect ne sont donc pas forcément concordantes.

Ces différentes limitations : subjectivité, manque d'exhaustivité et donc de répétabilité, font du contrôle d'aspect classique une métrologie d'une fiabilité douteuse ou tout du moins très perfectible.

D'autre part, si cette technique est adaptée à la fabrication de pièce optique à l'unité, elle apparaît clairement incompatible avec une production industrielle. Considérons à ce propos l'exemple du projet français Laser Megajoule (LMJ) mis en oeuvre par le CEA (informations disponibles à l'adresse <http://www-lmj.cea.fr/html/cea.htm>). Dans le cadre de ce projet, de nombreuses pièces optiques de taille respectable (de l'ordre du m²) doivent être fabriquées. En prenant en compte le fait que ces optiques sont des consommables du montage, on peut évaluer le nombre de composants à produire à environ 600 par année. Etant donné que ces pièces doivent être contrôlées 5 à 6 fois pendant leur fabrication, cela représente finalement une surface totale d'environ 1500 m² (l'équivalent du quart d'un terrain de football!) à examiner par an. Atteindre un tel objectif dans des temps raisonnables nécessiterait un nombre très important de contrôleurs ce qui est impossible pour des raisons évidentes de coût et de faisabilité (comment former autant d'experts?). Ces contrôles sont pourtant indispensables pour assurer la pérennité des pièces qui, rappelons le, doivent soutenir un flux laser intense.

I.2.2 Étude du besoin et système proposé

Nous venons de voir que le contrôle d'aspect est une métrologie critique, qui doit garantir des exigences fonctionnelles de haut niveau et de plus en plus pointues. Dans ce contexte, le contrôle d'aspect classique, effectué par un expert humain devient insuffisant en termes de fiabilité, coût et cadence. Malheureusement, il n'existe sur le marché aucun système capable d'assurer automatiquement cette tâche.

Partant de ce double constat, on arrive naturellement à la problématique à laquelle cette thèse va tenter d'apporter une réponse : comment concevoir un système automatique de contrôle d'aspect des pièces optiques ? Afin d'affiner cette problématique industrielle, nous avons mené une étude de besoin en collaboration avec les équipes de SAGEM Saint Pierre du Perray. Cette réflexion a ensuite présidé à une spécification complète du système à concevoir. Les résultats de ce travail de conception préliminaire, décrivant par le détail les différentes fonctions de service à satisfaire sont reportés en Annexe B.

Globalement, on peut retenir de cette spécification que le système à concevoir doit garantir la qualité des contrôles effectués et ce en termes de :

- *Répétabilité* : cette propriété, fondamentale pour un outil de contrôle, associée au caractère automatique du traitement, assure l'impartialité du contrôle.
- *Exhaustivité* : le système doit permettre par ce biais une utilisation pertinente du formalisme des normes et un suivi des pièces via un archivage des données.
- *Autonomie* : on vise ainsi une réduction maximale de l'intervention des experts.
- *Respect des cadences* : le système doit s'intégrer à la chaîne de production avec un coût minimum en termes de temps et de disponibilité.
- *Robustesse vis-à-vis de l'environnement de travail* : le système doit être capable de fonctionner de manière correcte dans les conditions correspondant à son lieu d'intégration.
- *Respect de l'intégrité des pièces* : le contrôle d'aspect des pièces optiques ne doit pas avoir d'impact sur celles-ci.

Cette spécification, permet par ailleurs d'envisager le système de contrôle comme un maillon de la chaîne de production et fixe les performances qu'il devra atteindre.

Rappelons ici que les paramètres nécessaires au diagnostic de conformité d'un composant optique suivant le formalisme d'une norme sont la position, la taille et la nature de tous les défauts présents sur la pièce. Une première analyse du système à concevoir et des contraintes auxquelles il doit satisfaire permet d'identifier deux étapes fondamentales du diagnostic : la détection des défauts d'aspect d'une part, et leur catégorisation d'autre part.

La détection des défauts d'aspect :

Le système à concevoir doit pouvoir, en utilisant les informations fournies par le bloc capteur, détecter les défauts d'aspect présents sur le composant optique contrôlé et en extraire des informations de bas niveau (positions, tailles, représentations image). D'autre part, même si cela n'a pas de rapport direct avec le contrôle d'aspect, il est souhaitable d'obtenir également une information sur la rugosité de la pièce mesurée ; ce paramètre revêt en effet une grande importance pour les surfaces optiques, puisqu'il est lié à la quantité de lumière qu'elles diffusent.

En outre, la qualité de cette phase est prépondérante au regard des contraintes de répétabilité, d'exhaustivité, de robustesse et de cadences introduites précédemment. En clair, tous les défauts d'aspect présents sur la pièce optique contrôlée doivent être détectés de manière fiable, ce dans un temps compatible avec le temps total de contrôle recherché.

Pour atteindre cet objectif, nous avons travaillé d'une part sur le type de capteur utilisé par le système, et d'autre part sur les techniques de traitement du signal permettant de traiter les informations qu'il fournit.

La catégorisation des défauts :

Cette étape doit permettre, à partir de la représentation extraite dans la phase précédente de déterminer la nature d'un défaut donné. Elle a ainsi pour but final une mise en oeuvre pertinente des normes, qui nécessite au moins de pouvoir discriminer les vrais défauts des « faux » (traces de nettoyage, poussières...). Le système assurant cette tâche doit être capable de proposer une classification d'un défaut d'aspect à partir d'une représentation image, de manière suffisamment fiable et en un temps très réduit (il est en effet très probable que le nombre de défauts détectés, donc à classer, soit très important lors d'un contrôle standard). Finalement, cette étape de traitement est essentielle au respect des

contraintes d'autonomie (le diagnostic des défauts se fait sans intervention humaine) et de robustesse à l'environnement de travail (le système de contrôle est capable de gérer des environnements pas forcément propres).

Afin de satisfaire à ces différentes contraintes, nous avons choisi de traiter cette problématique à l'aide de techniques connexionnistes.

Ces phases principales s'articulent conformément au schéma de la Figure 11.

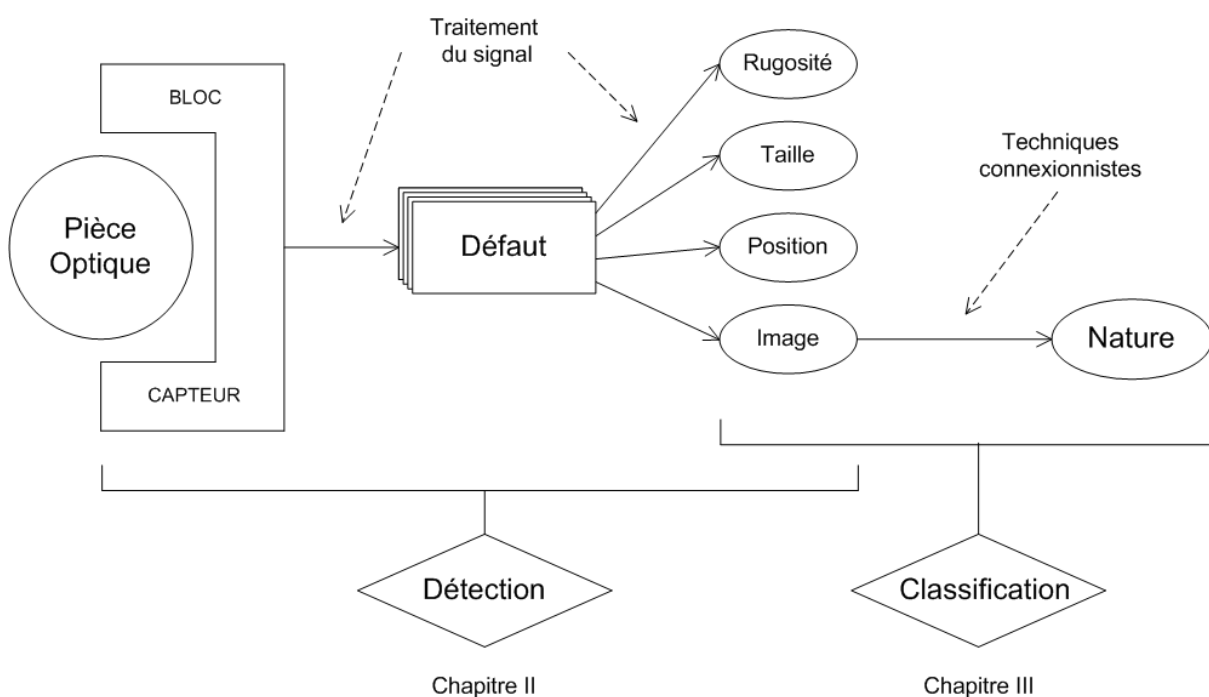


Figure I.9 : Schéma de principe de la solution proposée.

Une dernière phase n'est pas représentée dans le schéma : elle consiste à former un diagnostic de conformité de la pièce contrôlée par rapport à des spécifications exprimées dans le formalisme d'une norme. Pour ce faire, il faut analyser les renseignements disponibles (c'est-à-dire tailles, positions et natures des défauts) en suivant les recommandations de la norme utilisée. Un système expert sommaire peut aisément remplir cette tâche : les règles qu'il doit coder proviennent directement de l'interprétation des normes présentée dans l'Annexe A.

I.3 Conclusion du chapitre

Les imperfections de surfaces, telles que les rayures et les piqûres, générées pendant le processus de polissage des pièces optiques hautes performances, peuvent engendrer des lumières parasites incompatibles avec les performances recherchées. Malgré l'importance de ces défauts, leur métrologie, basée sur l'expertise humaine, est critique. Les applications nouvelles exigent en effet de détecter des défauts de plus en plus petits sur des surfaces de plus en plus grandes, ce qui aggrave les limitations inhérentes à ce type de contrôle. Malheureusement, il n'existe pour l'heure aucune solution automatique satisfaisante.

La présente thèse tente de répondre à cette problématique. Après l'analyse fonctionnelle d'une telle solution et les spécifications liées, nous proposons un système automatique de contrôle d'aspect des pièces optiques construit autour de deux phases principales. La première consiste à détecter les défauts présents sur la pièce optique contrôlée et à en extraire des informations de bas niveau (position, taille et représentation) ; la deuxième utilise la représentation obtenue pour assigner le défaut à une classe prédéfinie et ainsi déterminer sa nature.

Ces deux phases principales correspondent à deux problèmes ardues traités respectivement dans les chapitres II et III de ce manuscrit.

Systeme de detection des defauts d'aspect

Sommaire

| | | |
|-------------|--|-----------|
| II.1 | Capteurs pour le contr4ole d'aspect des pi4eces optiques | 29 |
| II.1.1 | Synth4ese des m4ethodes optiques existantes | 30 |
| II.1.2 | Microscopie Nomarski | 36 |
| II.1.2.1 | Principe de la mesure | 36 |
| II.1.2.2 | Th4eorie et r4ealisation technologique | 36 |
| II.1.2.3 | Avantages de la technique Nomarski | 38 |
| II.2 | Mise en 4oeuvre de la microscopie Nomarski pour l'4etude des pi4eces optiques | 40 |
| II.2.1 | Techniques d'exploitation d'images Nomarski existantes | 40 |
| II.2.2 | M4ethode de d4etection des d4efauts d'aspect propos4ee | 42 |
| II.2.2.1 | Pr4esentation | 42 |
| II.2.2.2 | Robustesse et r4ep4etabilit4e de la m4ethode | 49 |
| II.2.3 | Mesure de la rugosit4e | 52 |
| II.2.3.1 | D4efinition de la rugosit4e | 53 |
| II.2.3.2 | M4ethodes de mesure existantes pour 4evaluer la rugosit4e | 53 |
| II.2.3.3 | M4ethode de mesure propos4ee | 54 |
| II.2.3.4 | R4ep4etabilit4e de la m4ethode | 57 |
| II.2.3.5 | Etalonnage de la m4ethode | 59 |
| II.2.4 | Impl4ementation et validation des m4ethodes | 61 |
| II.2.4.1 | Pr4esentation du prototype | 61 |

| | | |
|-------------|--|-----------|
| II.2.4.2 | Le logiciel de post-traitement | 64 |
| II.2.4.3 | Complément au prototype : recalage de cartes | 67 |
| II.2.4.4 | Validation du prototype | 70 |
| II.3 | Conclusion du chapitre | 75 |

Le premier défi que nous devons relever consiste à détecter des défauts d'aspect de l'ordre du micromètre sur des pièces optiques de grande taille dans un temps compatible avec les contraintes liées à la production, et à en extraire des informations de bas niveau. Pour cela, une étude des différents capteurs pouvant avoir un intérêt pour le système visé a été menée. Nous avons finalement opté pour une solution utilisant deux capteurs travaillant à des échelles différentes. Les raisons de ce choix sont discutées dans ce chapitre. Le capteur choisi pour la détection à petite échelle est basé sur la microscopie Nomarski. Nous concentrons donc ensuite notre travail sur la mise au point d'un système de détection utilisant ce type de capteur. Dans ce cadre, différentes méthodes de traitement sont développées afin d'exploiter les images obtenues par le biais de cette technologie. Enfin, ces techniques sont validées par la réalisation d'un prototype les mettant en œuvre. Ce prototype, capable de détecter la présence de défauts d'aspect sur des composants optiques de taille limitée, et de mesurer la rugosité de leurs surfaces, est présenté à la fin de ce chapitre.

II.1 Capteurs pour le contrôle d'aspect des pièces optiques

Le choix de ces capteurs est primordial car il influence directement la nature et la qualité des informations qui seront traitées par le système pour détecter et caractériser les défauts d'aspect.

L'analyse des exigences présentées dans le chapitre précédent (et exposées en détail dans l'Annexe B) permet de dégager les principales contraintes auxquelles doivent satisfaire les capteurs recherchés :

- *Sensibilité* : ils doivent permettre de détecter des défauts dont une des dimensions latérales dépassent $2\mu\text{m}$ et d'une profondeur d'une dizaine de nm de manière exhaustive.
- *Rapidité* : ils doivent permettre de contrôler une pièce dans des délais raisonnables (typiquement 10 minutes pour 0.5 m^2).
- *Quantité d'informations* : ils doivent permettre de caractériser les défauts en termes

de taille, morphologie et nature (défauts matière, traces de nettoyage, poussière. . .).

- *Robustesse* : ils doivent être capable d'avoir un comportement stable et pertinent dans un environnement dont des facteurs physiques sont changeants (propreté, luminosité, température, vibrations) et sur des pièces de nature, de géométrie et de niveau de finition différents.
- *Sécurité* : ils doivent respecter l'intégrité physique de la pièce contrôlée.

Ces différentes contraintes vont nous guider dans la recherche des capteurs qui seront à la base du système de contrôle d'aspect automatique.

II.1.1 Synthèse des méthodes optiques existantes

Afin de recenser les techniques qui paraissent les mieux adaptées à la détection et à la caractérisation des défauts d'aspect sur les pièces optiques, une importante recherche documentaire a été effectuée. Cette recherche s'est appuyée sur la littérature d'optique instrumentale, les brevets industriels et les informations commerciales fournies par les entreprises spécialisées.

Pour une vue d'ensemble des moyens de mesure qui sont utilisés conventionnellement, le lecteur se reportera avec intérêt aux articles de revue (Chen, Brown, and Song 2000), (Charron 2004), (Gomez et al. 1998), (Whitehouse 1985), et (Whitehouse 1997). Il convenait cependant de fixer quelques critères permettant de circonscrire le champ d'investigation. La recherche a premièrement été limitée aux métrologies sans contact ; la nature et la géométrie des pièces à inspecter interdit en effet le recours à des méthodes nécessitant un contact physique, qui pourrait mettre en péril l'intégrité physique de la pièce optique contrôlée. Le deuxième facteur limitant est la sensibilité : les techniques dont la sensibilité est très en deçà de la sensibilité requise pour détecter les défauts d'aspect doivent être écartées.

Les méthodes répondant à ces critères sont présentées dans l'Annexe C, où sont exposés pour chaque technique abordée :

- le principe de la mesure
- la réalisation technique de celle-ci
- les spécifications d'un produit commercial dérivé de cette technique

- les contraintes de mesure
- ses avantages et ses inconvénients
- une application à un contrôle qualité

Le Tableau II.1 présente de manière synthétique les principales caractéristiques des techniques étudiées dans l'Annexe C. A noter, que pour les différentes microscopies, les diverses estimations (définition, taille des champs...) sont données pour un grossissement x200. Dans ce tableau, la colonne « C » désigne la capacité à caractériser les défauts d'aspect ; la colonne « D » désigne quant à elle la capacité à détecter ces mêmes défauts. Ces propriétés, fondamentales pour notre application, sont liées directement aux résolutions en X, Y et Z (colonne 5 et 6 du Tableau II.1) accessibles grâce au capteur considéré. Les paramètres de durée d'une acquisition (colonne 8 du Tableau II.1) et de champ mesuré (colonne 9 du Tableau II.1) permettent quant à eux d'évaluer le temps total de mesure d'une pièce optique (d'une taille typique de l'ordre du m^2) ; on peut ainsi évaluer la compatibilité d'un capteur avec la contrainte forte du temps de contrôle. Le coût (colonne 7 du Tableau II.1) est évidemment à prendre en compte dans le cadre industriel qui est le nôtre. Enfin certaines spécificités des capteurs envisagés (colonne 4 du Tableau II.1) influent de manières diverses sur leur intérêt potentiel pour le système visé.

On peut classer les différents capteurs étudiés en quatre familles distinctes :

Capteurs à acquisition séquentielle :

Ces capteurs sont capables de fournir les coordonnées 3D d'un point de l'objet étudié. En répétant la mesure selon un tramage donné, on reconstruit un profil en trois dimensions de la surface de cet objet.

Leur principal avantage est le fait qu'ils fournissent une information 3D, particulièrement intéressante du point de vue de l'interprétation des données. D'autre part, ces capteurs sont en général très précis. Cependant, la définition en X et Y est directement dépendante de la mécanique qui donne le maillage des mesures sur la pièce. Cette technique de maillage limite également le traitement de petits défauts (par rapport à la taille caractéristique du maillage) présentant de grandes pentes. D'autre part, il est impossible de mener de telles mesures sur de grandes surfaces car les temps de mesure sont prohibitifs.

Finalement, l'utilisation de tels capteurs est envisageable pour l'application que nous

| Méthode | D | C | Spécificités | Précision en Z | Définition en X,Y | Coût | Durée d'une acquisition | Champ mesuré |
|---|---|---|--|------------------|----------------------------|------|-------------------------|---------------------------------|
| Focalisation dynamique conventionnelle | X | X | Sur pièce relativement propre | ≈ 1 nm | ≈ 1 μ m | +++ | $\approx 10e-3$ s | - |
| Focalisation dynamique par codage chromatique | X | X | Sur pièce relativement propre | ≈ 10 nm | ≈ 0.5 μ m | +++ | $\approx 10e-6$ s | - |
| Microscopie Bright Field | X | X | | - | ≈ 0.5 μ m | + | $\approx 10e-4$ s | $\approx 500\mu$ m X 500μ m |
| Microscopie Dark Field | X | X | Pour des matériaux peu diffusant | - | ≈ 0.5 μ m | + | ≈ 1 s | $\approx 500\mu$ m X 500μ m |
| Microscopie Nomarski | X | X | Sensibilité en Z en dessous de 10 nm Possibilité d'avoir une mesure quantitative en Z | - | ≈ 0.5 μ m | + | $\approx 10e-3$ s | $\approx 500\mu$ m X 500μ m |
| Photogrammétrie microscopique | X | X | Reconstruction lourde algorithmiquement | ≈ 100 nm | ≈ 0.5 μ m | + | < 1s | $\approx 500\mu$ m X 500μ m |
| Microscopie interférométrique | X | X | Sensibilité aux vibrations | ≈ 1 A | ≈ 0.5 μ m | ++ | ≈ 5 s | $\approx 500\mu$ m X 500μ m |
| Interférométrie holographique | X | | la pièce étudiée doit être suffisamment diffusante | < 1 nm | dépend de la configuration | + | < 5 min | $\approx 1m^2$ |
| Strioscopie | X | | sensibilité inconnue | - | dépend de la configuration | + | < 1 s | $\approx 1m^2$ |
| Illumination avec pose prolongée | X | | sensibilité inconnue | - | dépend de la configuration | + | 1 min | $\approx 1m^2$ |
| Diffusométrie | X | | Possibilité d'avoir des informations utiles à la caractérisation du défaut | - | ≈ 10 μ m | ++ | $\approx 10e-6$ s | $\approx 10\mu$ m X 10μ m |

Table II.1 : Principales caractéristiques des techniques étudiées dans l'Annexe C.

cherchons à développer à condition qu'ils soient utilisés sur des champs de mesure très limités. Ceci permettrait en effet de jouir de la précision et de la richesse de l'information fournie, sans pâtir de la lenteur de la mesure.

Capteurs à acquisition parallèle :

Ces capteurs acquièrent en une mesure, une image ou une pseudo-image correspondant à un champ équivalent à un grand nombre de mesures ponctuelles (typiquement de l'ordre de 10^6).

Leur principal avantage est le fait qu'ils fournissent en une mesure rapide, une information précise sur une surface relativement étendue (typiquement de l'ordre du mm^2 pour des grossissements moyens). Cependant, la définition en X et Y de la mesure est directement dépendante du grossissement employé qui fixe également la taille du champ mesuré. Ces méthodes nécessitent d'autre part des réglages qui peuvent être délicats à mettre en œuvre automatiquement. Leur inconvénient majeur est qu'il semble difficile de les utiliser pour le contrôle de grandes surfaces, qui nécessiterait des temps de mesure conséquents.

Ces capteurs sont potentiellement intéressants pour l'application que nous cherchons à développer à deux titres : premièrement, ils sont tout à fait adaptés à l'étude de champs de petite taille puisqu'ils peuvent en donner rapidement une image précise ; d'autre part, on pourrait mener ce genre de mesure sur des pièces entières en augmentant significativement la vitesse du scan (le temps le plus limitant n'étant pas le temps de mesure mais le temps de déplacement du capteur par rapport à la pièce).

Enfin, il est à noter que ces différents capteurs présentent des avantages qui leur sont propres, le plus intéressant étant sans doute la possibilité, pour certains, d'avoir accès à une information tridimensionnelle.

Capteurs sensibles à la dynamique :

Ces capteurs n'effectuent pas de mesures caractérisant l'état statique de l'objet à étudier mais sont capables de détecter des variations très fines de cet état entre deux instants.

Leur principal avantage est le fait qu'ils fournissent rapidement une information sur un champ de mesure très étendu. D'autre part, ces informations sont très précises pour ce qui est du déplacement en Z. Cependant, la définition en X et Y est directement liée au mode d'acquisition de la mesure et, compte tenu des larges surfaces à analyser, limitée. Un autre

inconvenient est la difficulté d'automatiser certaines de ces techniques. Mais le principal problème est de trouver une contrainte adéquate qui mette en relief dynamiquement les défauts, de manière à ce qu'ils deviennent détectables malgré une faible définition en X et Y.

Il n'existe pas, à notre connaissance, de telles contraintes, compatibles avec le respect de l'intégrité des pièces optiques mesurées, ce qui exclut l'utilisation de ce type de capteurs.

Capteurs utilisant des moyens d'amplification des défauts :

Le principe de base de ces capteurs est d'amplifier les défauts pour qu'il soient détectables macroscopiquement.

Leur principal avantage est le fait qu'ils permettent l'étude d'une surface très étendue en un temps relativement court. Cependant, la définition de la mesure en X et Y est directement liée à son mode d'acquisition, et, au vu des larges surfaces à analyser, limitée. Ceci n'est en fait pas gênant si l'on ne cherche pas à imager les défauts mais juste à les détecter. Le point sensible de ces méthodes est l'amplification des défauts, qui doit être adaptée à l'objet recherché et assurer une sensibilité correcte.

Ces capteurs sont donc potentiellement intéressants pour l'application que nous cherchons à développer. En effet, ils permettraient, sous réserve qu'ils aient une sensibilité suffisante, de détecter les défauts sur de grandes surfaces en un temps restreint. Il faut noter qu'ils peuvent également fournir une information, certes limitée, utile à la caractérisation du défaut.

Au final, ces différentes techniques se répartissent en deux catégories distinctes. D'une part les techniques permettant la détection rapide des défauts mais n'amenant pas suffisamment d'information pour rendre possible leur caractérisation (elles correspondent aux deux dernières familles décrites). D'autre part, des techniques plus précises permettant cette caractérisation et a fortiori la détection des défauts recherchés (les deux premières familles décrites). Si cette deuxième catégorie est la plus séduisante au premier abord, elle souffre d'un inconvénient majeur : elle nécessite un temps de contrôle prohibitif pour des pièces de grande taille (contrôle d'une durée de l'ordre de la semaine!).

Donc, s'il existe un nombre raisonnable de capteurs satisfaisant aux contraintes de sécurité, de sensibilité et de robustesse, il semble que les contraintes de rapidité et de

quantité d'information soient incompatibles. Une solution mettant en œuvre deux capteurs différents appartenant aux deux catégories précitées doit donc être envisagée. Cette approche utilisée dans (Henzler et al. 2000), (Kayser, Bothe, and Osten 2004) et (Weiss and Saranlı 2004) conduit à des mesures comportant deux étapes. Tout d'abord, un premier capteur relève les emplacements de défauts potentiels. Ensuite, ces emplacements sont examinés par un deuxième capteur, capable de détecter plus finement les éventuels défauts et de les caractériser. On profite ainsi de la rapidité du premier capteur qui délimite seulement des zones de présence de défauts et de la précision du deuxième capteur qui permet leur détection plus précise et leur caractérisation (en termes de taille et de nature).

Le système de contrôle d'aspect utilisera donc *deux capteurs distincts* :

- un capteur de **détection « grossière »**
- un capteur de **détection « fine » et de caractérisation**

Cette recherche documentaire a par ailleurs permis d'identifier les techniques les plus prometteuses sur lesquelles pourraient être basés ces deux capteurs. Ce choix a été basé sur un compromis coût / qualité des capteurs (en termes de robustesse, définition et rapidité), qui peut être apprécié à partir des données reportées dans le Tableau II.1.

En ce qui concerne le choix du capteur de détection « grossière » des défauts, notre choix s'est porté sur deux techniques différentes :

- d'une part, sur la strioscopie, technique à la base de la norme NF ISO 10110-7-méthode 2 concernant le contrôle d'aspect ;
- d'autre part, sur la méthode d'éclairage par les tranches, technique employée pour les contrôles d'aspect des composants du NIF, équivalent américain du Mégajoule.

Ces techniques, dont la sensibilité exacte était inconnue au départ, ont été testées en interne au sein de REOSC et ont donné des résultats concluants, qui valident leur utilisation par le système à concevoir.

Pour ce qui est du capteur de détection « fine », qui permet la caractérisation des défauts, notre choix s'est porté sur la microscopie optique dans un premier temps, puis plus précisément sur la technique de visualisation Nomarski.

Dans la suite de ce travail, on s'intéressera exclusivement à la phase de caractérisation puisqu'elle met en œuvre la véritable intelligence du système. En effet, à partir des données fournies par ce deuxième capteur, on devra déterminer la taille, les paramètres morphologiques et la nature des défauts détectés lors de la première phase (détection « grossière ») du contrôle. Par conséquent, toutes les données étudiées dans les chapitres suivants proviennent de la technique de microscopie Nomarski, qui est donc présentée en détails dans le paragraphe suivant.

II.1.2 Microscopie Nomarski

II.1.2.1 Principe de la mesure

La microscopie Nomarski (ou Differential Interference Contrast) est une technique résultant de l'utilisation des concepts de la microscopie associés avec ceux de l'interférométrie Nomarski (visualisation directe des pentes de la surface étudiée).

Nous ne reviendrons pas ici sur le principe bien connu du microscope (Flewitt and Wild 1994). L'imagerie en mode Nomarski est quant à elle une technique d'interférométrie différentielle (Bouchareine 1999) : après dédoublement latéral de l'onde incidente en deux ondes subissant l'influence de l'objet, leur recombinaison fournit un interférogramme mettant en évidence la différence de chemin optique (on visualise en fait un gradient de longueur de chemin optique).

En d'autres termes une image obtenue en utilisant la microscopie Nomarski représente les pentes de l'objet étudié. Ceci explique l'impression de 3D très marquée qui caractérise ces images. Ainsi, dans l'image de la Figure II.1 représentant un champ de 1.7 mm x 1.3 mm obtenu grâce à cette technique, on distingue bien la rugosité de la pièce étudiée.

II.1.2.2 Théorie et réalisation technologique

On produit deux images juxtaposées d'une source lumineuse ponctuelle en « jouant » avec deux directions orthogonales de polarisation de l'onde électromagnétique (en utilisant un prisme de Wollaston par exemple). Les interférences sont possibles en repolarisant les deux ondes dans le même plan grâce à un analyseur. Le schéma d'un tel montage optique est présenté dans la Figure II.2.

Finalement, on obtient une image en pseudo 3D, une différence de teinte ou de niveau

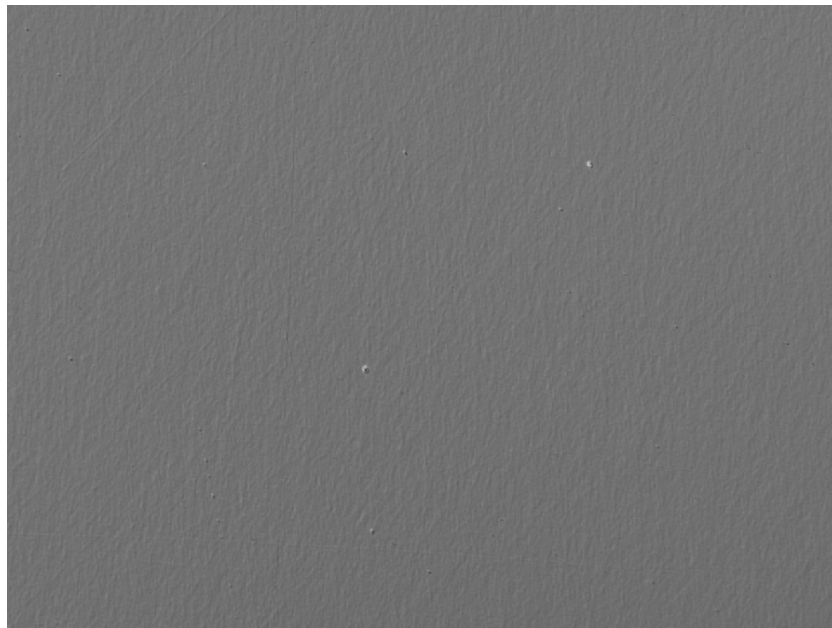
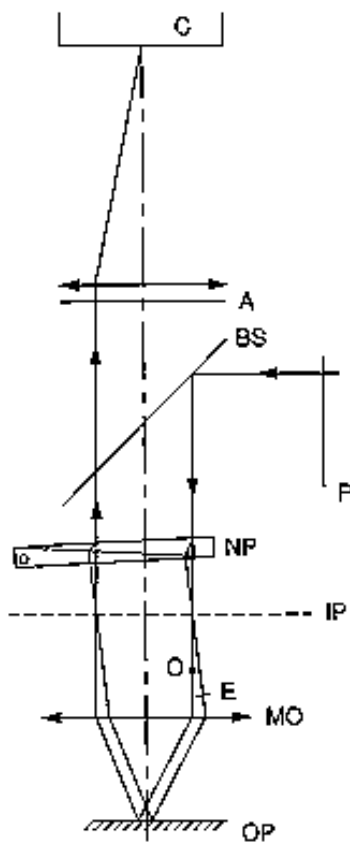


Figure II.1 : Image d'un champ obtenue par microscopie Nomarski.



Où
 OP désigne le plan objet
 MO, l'objectif du microscope
 IP, le plan d'intersection des rayons
 NP, le prisme de Nomarski
 BS, le cube séparateur
 P, le polariseur
 A, l'analyseur,
 et C, le plan image

Figure II.2 : Schéma optique d'un microscope DIC (extrait de (Chatterjee 2003)).

de gris correspondant à une différence de pente. En effet, si le dédoublement d est effectué dans la direction de x , en tout point introduisant une variation de chemin optique $\delta(x)$, la différence de marche $\Delta(x)$ entre les deux ondes est (Roblin 1999) :

$$\Delta(x) = \delta(x + d/2) - \delta(x - d/2) \approx d\delta'(x) \quad (\text{II.1})$$

avec $\delta'(x)$, le gradient en x de $\delta(x)$.

De plus, on peut montrer que (Holmes and Levy 1987) que :

$$I(x, y) = 2a^2[1 + \Delta(x)] \quad (\text{II.2})$$

où $I(x, y)$ désigne l'intensité lumineuse observée à travers le microscope DIC , et a est une constante dépendant du microscope. Ainsi, l'intensité lumineuse détectée dépend linéairement du gradient en x du chemin optique au point considéré, c'est à dire de la pente en x du relief de la pièce en ce point. Pour des indications théoriques plus complètes, le lecteur pourra se rapporter à (Lessor, Hartman, and Gordon 1979) et pour une étude détaillée sur la conception et la fabrication d'un tel microscope à (Chatterjee 2003).

II.1.2.3 Avantages de la technique Nomarski

Cette technique présente trois avantages déterminants qui ont motivé son choix pour le système à concevoir.

Premièrement, elle présente, de part sa nature différentielle, une très bonne sensibilité par rapport au relief de la pièce mesurée. Cette qualité que l'on peut apprécier sur la Figure II.1 – la rugosité de la pièce imagée est faible : 11 Å RMS (Root Mean Square, voir le paragraphe II.2.3.1, p.53) – nous assure que tous les défauts à détecter seront bien imagés lors du contrôle. En fait, cette technique permet d'imager des défauts dont la profondeur est de l'ordre du nm (Lessor, Hartman, and Gordon 1979), ce qui est un ordre de grandeur plus petit que les défauts recherchés. Il est même possible de détecter des défauts dont la taille (longueur et largeur) est inférieure à la limite de résolution de l'objectif microscopique employé (Holmes and Levy 1987). Sur ce point, elle surpasse les techniques de la microscopie conventionnelle (présentées en Annexe C) Bright-Field (fond blanc) et Dark-Field (fond noir) pour des éclairages moyens. Ceci est illustré par la Figure II.3 représentant le même champ microscopique mesuré avec le même microscope utilisant

les trois différents modes précités. Dans cet exemple, les deux raies au centre et à droite du champ sont beaucoup plus visibles dans l'image Nomarski que dans l'image Dark-Field alors que pratiquement invisibles dans l'image Bright-Field.

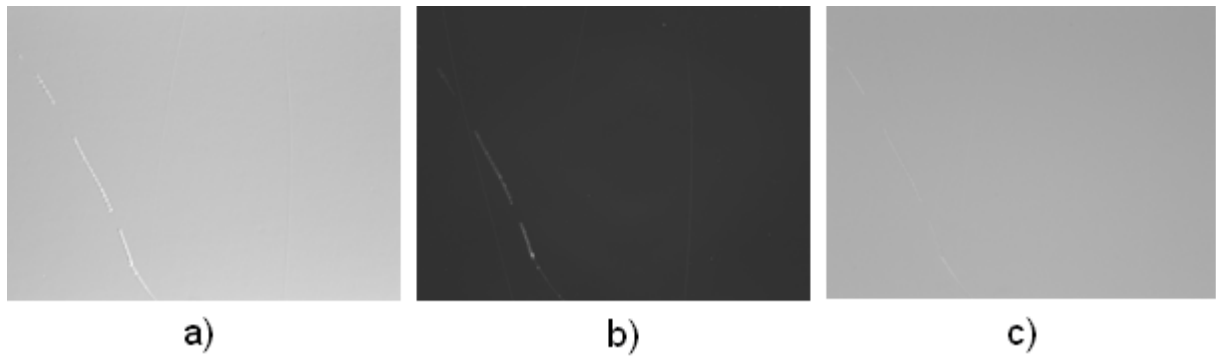


Figure II.3 : Le même champ imagé à l'aide de la microscopie (a) Nomarski ; (b) Dark-Field ; (c) Bright-Field.

D'autre part, cette technique présente une grande robustesse par rapport à l'éclairage, ce toujours grâce à son caractère différentiel. Ainsi, le fait d'augmenter le flux de lumière utilisé par le microscope ne change pas fondamentalement la nature de l'image recueillie, mais augmente de manière uniforme son éclairage. Ceci garantit la répétabilité des mesures et leur robustesse par rapport à l'éclairage. A l'inverse, une telle augmentation de flux incident améliore grandement la sensibilité de la technique Dark-Field mais influence nettement l'image qui est faite des défauts. Ce phénomène illustré dans la Figure II.4 interdit un dimensionnement fiable des défauts par la microscopie Dark-Field.

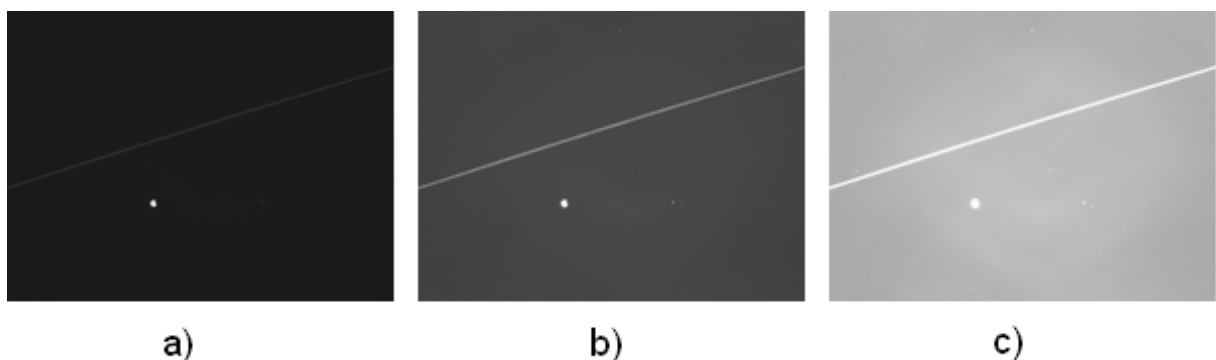


Figure II.4 : Le même champ imagé à l'aide de la microscopie Dark-Field avec différents temps d'exposition : a) 200 ms ; b) 500 ms ; c) 1s.

Enfin, la microscopie Nomarski fournit des données tridimensionnelles (on visualise en effet les pentes de la surface), même si l'information de profondeur n'est pas disponible directement. Ce dernier point est particulièrement intéressant dans la mesure où cette information supplémentaire pourrait permettre de caractériser la profondeur des défauts ou la rugosité de la pièce étudiée, mais aussi faciliter la décision quant à la nature des imperfections détectées.

II.2 Mise en œuvre de la microscopie Nomarski pour l'étude des pièces optiques

II.2.1 Techniques d'exploitation d'images Nomarski existantes

Comme nous l'avons vu, la microscopie Nomarski permet d'obtenir des informations robustes et de nature tridimensionnelle à l'aide d'un montage optique relativement simple. Ceci explique que cette technique soit utilisée traditionnellement dans l'industrie optique, d'une part pour contrôler de manière informelle la rugosité en cours de process (les opérateurs sont capables d'évaluer approximativement la rugosité d'une pièce à partir d'une image Nomarski de celle-ci), d'autre part pour le contrôle d'aspect de certaines pièces dont les spécifications sont très sévères. Cependant, dans ces deux cas, seule l'expertise humaine permet d'accéder à l'information recherchée.

Pour ce qui est des méthodes qui s'affranchissent de cette expertise (les seules qui nous intéressent puisque l'on cherche à construire un système autonome), elles se répartissent en deux types : les techniques qui visent à reconstruire une représentation 3-D de la surface à partir d'images Nomarski, et celles dont le but est d'en extraire directement une information (rugosité, présence d'un défaut, ...) sans utiliser une telle reconstruction.

Techniques de reconstruction 3-D de surface :

Etant donné qu'une image Nomarski est en fait une représentation graphique des pentes de la pièce imagée selon un axe donné, on peut appliquer les techniques classiques de reconstruction de surface, modale ou zonale (Tian, Itoh, and Yatagai 1995), ou des algorithmes de traitement de l'image (Arnison et al. 2000) (Arnison et al. 2004). Il est à noter que la plupart de ces méthodes utilisent la connaissance des pentes selon deux axes

perpendiculaires et donc nécessitent implicitement d'imager la pièce deux fois. Outre le fait que ces méthodes soient lourdes d'un point de vue calculatoire et assez sensibles au bruit, elles ne peuvent fournir que des données qualitatives. En effet, si la reconstruction d'une cartographie de la surface est possible, aucune valeur quantitative de profondeur ne peut être déterminée dans cette carte : en d'autres termes, le relief peut être reconstruit mais sans en connaître l'« altitude » caractéristique. Pour accéder à cette information, il est nécessaire d'utiliser le concept de phase-shift, ce qui n'est possible qu'en ajoutant un degré de liberté dans le montage (translation du prisme (Hartman, Gordon, and Lessor 1980)(Gebbinck et al. 1993), ou rotation de l'analyseur (Li et al. 2001)).

La littérature fournit donc un certain nombre de techniques qui permettent de reconstruire une surface 3-D à partir d'images Nomarski, ce avec une très bonne précision (meilleure que le nanomètre). Elles fonctionnent malheureusement au prix d'un alourdissement conséquent de la charge de calcul, du temps de mesure et de la complexité du montage optique utilisé. Ceci explique d'une part que nous ayons écarté l'utilisation de telles techniques pour notre système, et d'autre part que plusieurs travaux, présentés plus bas, s'affranchissent d'une telle reconstruction.

Autres techniques d'exploitation :

On peut citer tout d'abord (Young and Gray 1996), (Young et al. 1998) et (Yasuda, Bannai, and Onami 1999) dont le but est de détecter des cellules dans des images Nomarski. Ces travaux bien qu'intéressants sont trop spécialisés pour être adaptés à nos besoins.

Pour en revenir à des problématiques plus proches des nôtres, dans (Jabr 1985), l'auteur propose une méthode permettant de déterminer de manière précise la rugosité d'une pièce à partir d'images Nomarski de cette pièce. Outre le fait qu'elle nécessite une procédure de calibration, le principal inconvénient de cette méthode qui paraît très intéressante pour l'étude de pièce optique, est qu'elle utilise un grand nombre d'images Nomarski.

Finalement, dans (Reschke, Bäumer, and Kross 1993), une analyse statistique de la répartition des niveaux de gris dans une image Nomarski permet la détection de défauts d'aspect sur des surfaces optiques polies. Cette technique qui pourrait répondre à notre besoin souffre néanmoins de deux inconvénients majeurs : premièrement elle nécessite de fixer de manière arbitraire des seuils statistiques de détection (ce qui pose le problème du dilemme entre sensibilité de la détection et robustesse aux hétérogénéités d'éclairage) ;

deuxièmement, elle travaille avec des carrés de 16 pixels de côté et donc ne peut fournir des contours précis pour les défauts détectés.

Ainsi, si certains travaux retiennent l'attention et proposent de « bonnes » idées, la littérature ne fournit pas de solutions satisfaisantes directement intégrables au système à concevoir. Par conséquent, les techniques adéquates ont dû être développées : elles sont présentées dans la suite de ce chapitre.

II.2.2 Méthode de détection des défauts d'aspect proposée

II.2.2.1 Présentation

L'objectif de la méthode proposée est de détecter dans une image Nomarski tous les défauts d'aspect significativement plus profonds que la taille caractéristique de la rugosité et de définir de façon robuste leurs contours (Voiry et al. 2005). Elle permet également de construire des représentations image des défauts détectés utilisées pour leur catégorisation ultérieure.

Le traitement de l'image Nomarski se compose de plusieurs phases :

- *Une phase de prétraitement*, qui permet de réduire l'influence des défauts d'éclairage et d'améliorer la visibilité des défauts,
- *Une phase de seuillage adaptatif*, qui permet d'effectuer la détection des défauts proprement dite,
- *Une phase de filtrage et de segmentation*, qui permet de supprimer le bruit résultant de la phase précédente et de trouver les contours des défauts,
- *Une phase d'extraction d'« imagerie »*, qui permet de fournir une représentation image de chaque défaut détecté.

Ces différentes phases sont décrites dans la suite de ce paragraphe, en s'appuyant sur l'exemple de l'image Nomarski présentée dans la Figure II.5. Cet exemple illustre bien la difficulté de la détection des défauts d'aspect, due principalement à la nature de la représentation de ces défauts en microscopie Nomarski (défauts constitués d'une partie « noire » et d'une partie « blanche ») et de la présence de nombreux artefacts dans la trame de rugosité.



Figure II.5 : Image originale issue du mode DIC.

Prétraitement

Le prétraitement préconisé, tiré de (Bourgeat, *et al.*, 2003), est utile à deux titres. D'une part, il permet d'égaliser la dynamique de l'image : on corrige ainsi les inhomogénéités d'éclairage (dus à l'illumination non uniforme), et les modifications locales de contraste (dus aux variations d'épaisseur des couches de matériaux et aux variations de focus). D'autre part, il permet de mettre en évidence les défauts à détecter d'une façon constante : les contours des défauts ne sont plus caractérisés par un passage du gris (correspondant au « fond ») au noir ou du gris au blanc, mais toujours d'un passage du gris au noir (Figure II.6).

Le principe de ce prétraitement est très simple : l'amplitude de chaque pixel est divisée par la somme de la valeur moyenne M et de l'écart-type σ pris dans son voisinage 5X5 :

$$M = \frac{1}{25} \sum_{i=-2}^2 \sum_{j=-2}^2 P(x+i; y+j) \quad (\text{II.3})$$

$$\sigma^2 = \frac{1}{25} \sum_{i=-2}^2 \sum_{j=-2}^2 (P(x+i; y+j) - M)^2 \quad (\text{II.4})$$

où P est le tableau à deux dimensions contenant les valeurs des pixels de l'image à traiter, et $(x ; y)$ les coordonnées du pixel considéré dans ce tableau.



Figure II.6 : Image de la Figure II.5 après prétraitement.

La nouvelle valeur associée à ce pixel est donnée par :

$$P'(x; y) = \frac{P(x; y)}{M + \sigma} \quad (\text{II.5})$$

Seuillage adaptatif

On cherche maintenant à détecter les défauts. Etant donné que ceux-ci sont caractérisés par des frontières de niveau de gris faible (noir), il paraît judicieux d'effectuer un seuillage par rapport au niveau de gris pour réaliser la détection visée. La difficulté est alors de trouver un seuil adéquat, qui soit assez haut pour détecter convenablement les défauts présents mais assez bas pour éviter de détecter les éléments constitutifs du « fond » (rugosité). Ce problème, illustré par la Figure II.7, ne trouve pas de solution satisfaisante dans les choix de seuil automatique usuels (isodata, somme de deux gaussiennes, maximisation de la variance) proposés par la littérature (Gonzalez and Woods 2002).

La méthode proposée dans ce travail consiste à augmenter le seuil jusqu'à apparition de la rugosité. On commence par filtrer le bruit présent dans l'image à l'aide d'un filtre gaussien 3x3. On cherche ensuite un critère caractérisant l'apparition de la rugosité. L'évaluation directe de la population des pixels de chaque niveau de gris ne convient pas. En effet, un pixel d'un niveau de gris donné, s'il est dans le voisinage de pixels de niveaux de gris plus bas a des chances d'appartenir à un défaut plutôt qu'à un motif de rugosité.

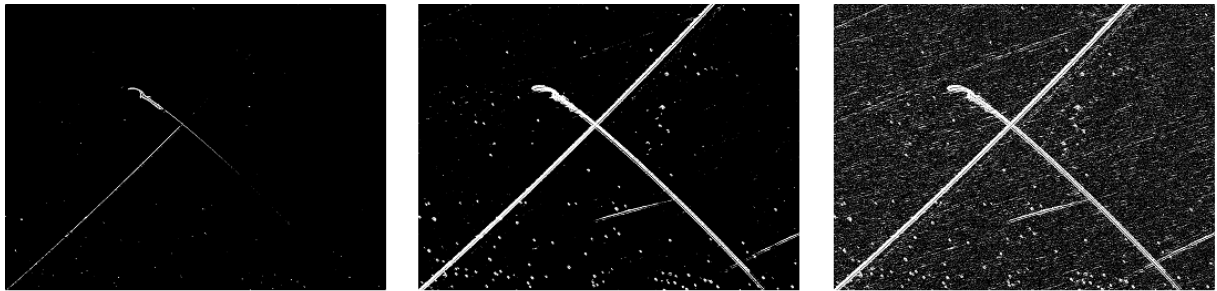


Figure II.7 : Image pré-traitée seuillée avec $s=78$, $s=92$, $s=112$.

Pour éviter ce problème, l'image est tout d'abord découpée en carré de 8×8 pixels. Ensuite, pour chaque niveau de gris g , le nombre de carrés dont le pixel de niveau de gris minimum correspond à g est évalué. On obtient alors, après lissage par une gaussienne, une courbe caractéristique, dont un exemple est donné en Figure II.8. A notre connaissance, toutes les images Nomarski traitées donnent une courbe de cette forme, on peut donc s'en servir pour fixer un seuil robuste qui servira à la détection des défauts.

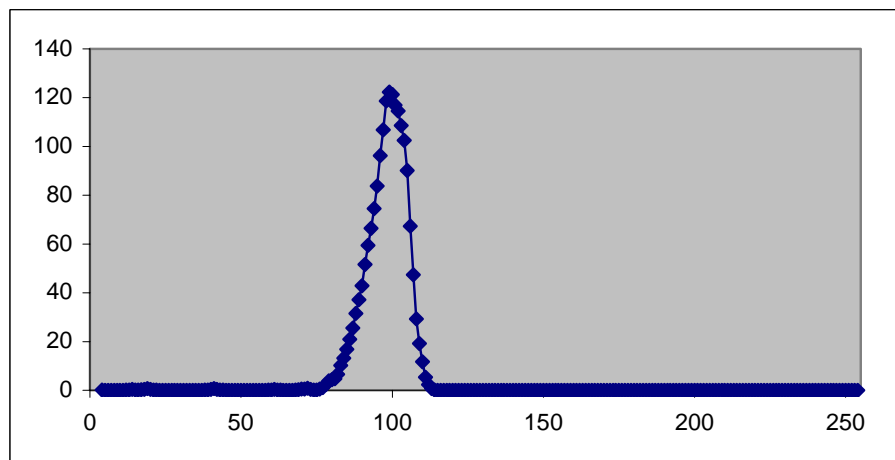


Figure II.8 : Nombre de carrés 8×8 dont le pixel de niveau de gris minimum est du niveau de gris reporté en abscisse.

Après expérimentations, il a été décidé de choisir comme seuil, le seuil S tel que :

$$S = Max - H/2 \tag{II.6}$$

où Max est le niveau de gris correspondant au maximum de la courbe considérée, et H désigne la largeur à demi-hauteur de cette même courbe.

La pièce imagée pouvant avoir des caractéristiques physiques très différentes dans un même champ, on pourra calculer au besoin 12 seuils différents correspondant aux 12 parties de l'image (au format 4/3) définies dans la Figure II.9.

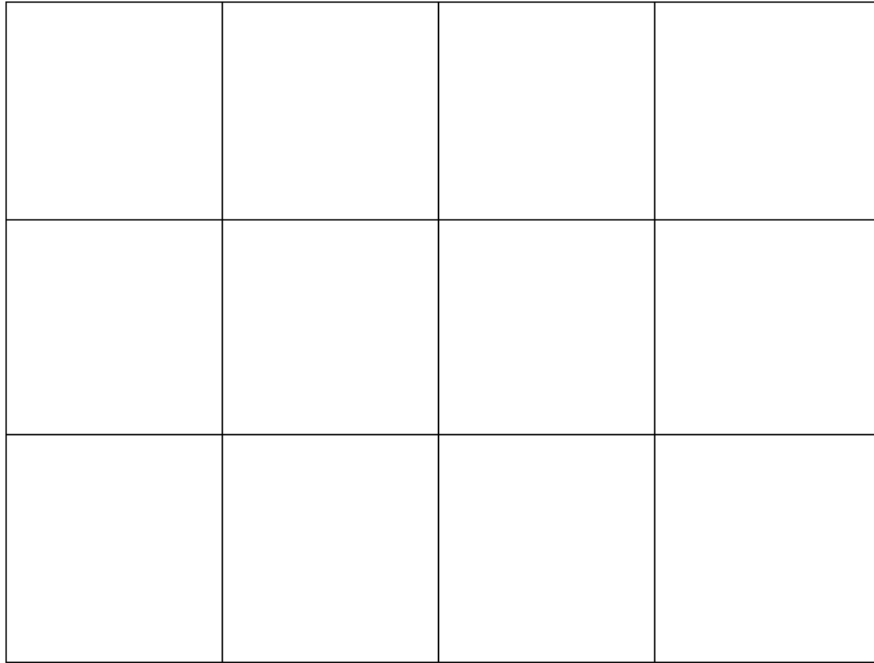


Figure II.9 : Découpage de l'image en 12 parties de taille équivalentes.

La méthode employée ici permet d'assurer une sensibilité importante et constante, puisque le critère de détection est lié à l'apparition de la rugosité. Par conséquent, tous les défauts significativement plus profonds que la taille caractéristique de la rugosité sont détectés. Cependant, on ne détecte que les contours des particules, certains défauts peu marqués sont détectés comme des nuages de points, et enfin, on remarque quelques détections parasites de faible étendue (Figure II.10).

Filtrage et segmentation

Cette phase qui a pour but de régler ces problèmes et de recenser les défauts détectés est constituée de trois opérations distinctes :

- l'application d'un filtre de rang à l'image seuillée,
- l'application d'une érosion morphologique à l'image obtenue.
- l'identification des différents défauts dans cette image par Flood-Fill

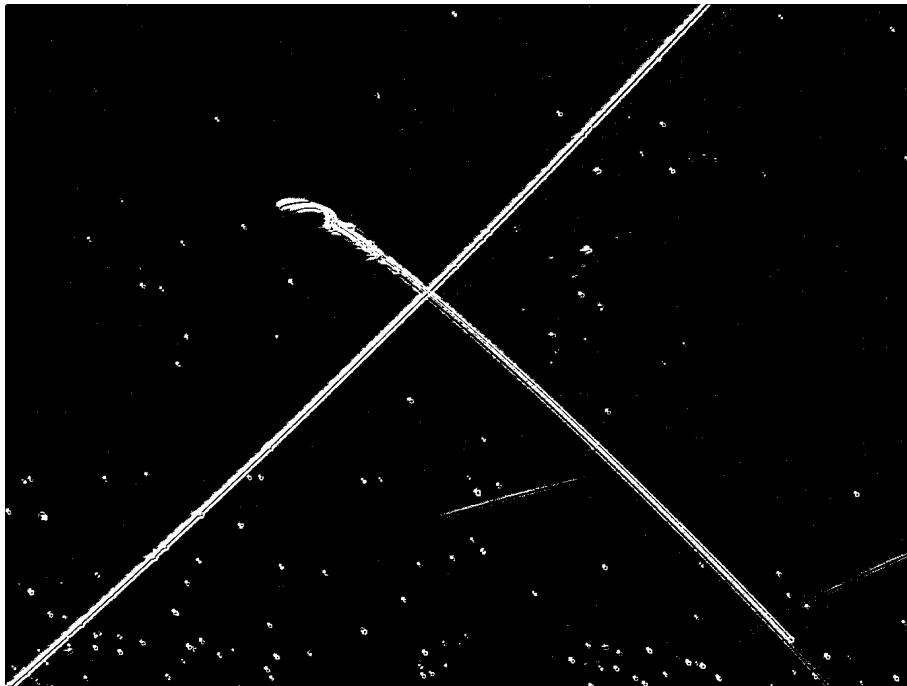


Figure II.10 : Image de la Figure II.5 après seuillage.

Appliquer un filtre rang de dimension D et de rang p (exprimé en pourcentage) à une image consiste à remplacer le niveau de gris de chaque pixel de l'image par le niveau de gris du $E(pD^2)$ -ième pixel dans la liste ordonnée de manière croissante (du point de vue des niveaux de gris) des D^2 pixels appartenant au voisinage du pixel considéré. Dans les faits, pour des valeurs importantes du paramètre p , cela est presque équivalent à appliquer une dilatation morphologique, avec la différence non négligeable que les points blancs isolés (qui correspondent à du bruit) sont éliminés. Cette opération est suivie par une érosion morphologique : on a ainsi effectué une sorte de fermeture morphologique, qui permet de rassembler les nuages de points constituant un seul défaut, et de « boucher » les trous de certains défauts, tout en éliminant du bruit. Le résultat de cette dernière opération est représenté dans la Figure II.11, où la précision des contours des défauts détectés (en rouge) peut être appréciée.

Il est à noter qu'il existe, à travers le choix des paramètres du filtre de rang et de l'érosion, un dilemme entre la sensibilité de la méthode et la justesse des contours fournis. Si l'on choisit d'atteindre la sensibilité maximale, la taille des défauts est surestimée (voir paragraphe II.2.2.2, p.50) ; si l'on choisit d'accéder à une taille de défaut très juste, certains défauts superficiels ne seront pas détectés.

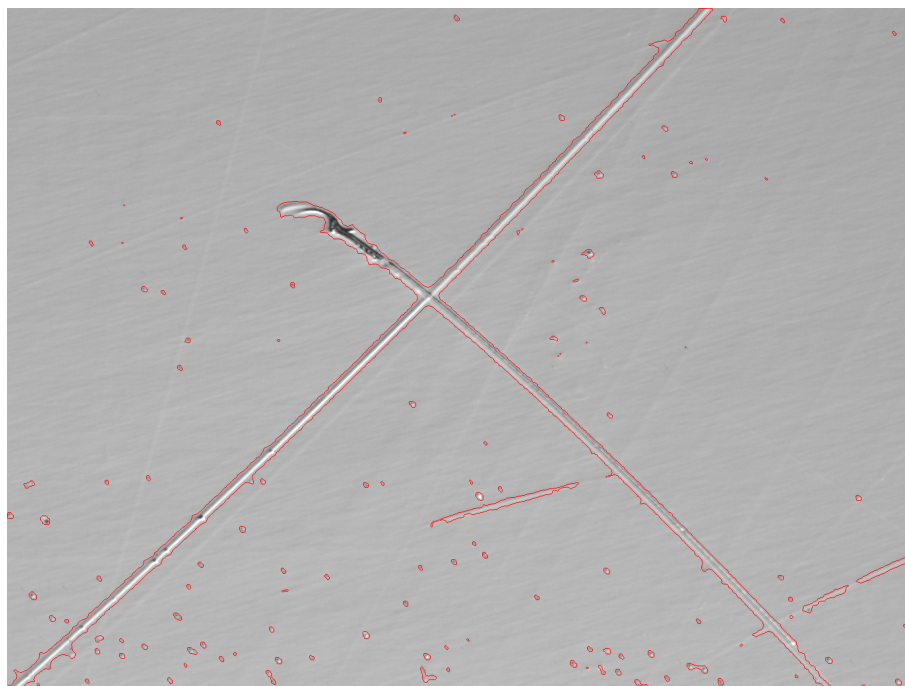


Figure II.11 : Image de la Figure II.5 avec les contours détectés.

La dernière étape consiste à faire la détection des défauts proprement dite. Pour cela, on applique l'algorithme classique du Flood-Fill (remplissage par « germe »), qui permet d'extraire les différentes composantes connexes de l'ensemble des pixels détectés comme faisant partie d'un défaut. L'image obtenue à la précédente étape est parcourue, et une nouvelle « étiquette » est assignée à chaque pixel non-étiqueté rencontré ; cette étiquette est ensuite propagée aux pixels voisins de ce pixel « germe », puis itérativement à tous les pixels qui lui sont connexes. On obtient finalement un ensemble de groupes de pixels connexes avec des étiquettes distinctes, correspondant aux différents défauts détectés.

Extraction d'« imagettes » de défauts

A partir de ces informations, sont formées des représentations de défauts dans leur environnement immédiat. Pour chaque défaut, on forme une « imagette » en considérant les pixels de l'image Nomarski correspondant à ce défaut ainsi qu'une bande de 10 pixels autour de ces pixels. Un exemple synthétique de cette opération est présenté dans la Figure II.12. Dans cet exemple, les quatre imagettes extraites de l'image Nomarski représentent des défauts usuellement rencontrés sur les pièces optiques : une raie (a), une piqûre (b) (qui sont des « vrais » défauts), une poussière (c), et une trace de nettoyage (d) (qui sont des artefacts).

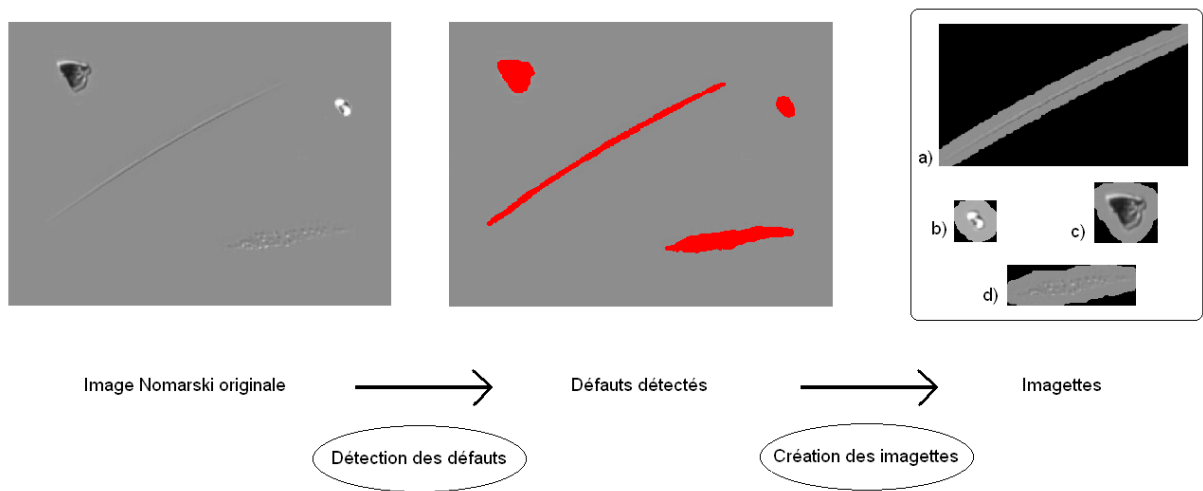


Figure II.12 : Extraction d'imagettes de défauts à partir d'une image Nomarski brute.

II.2.2.2 Robustesse et répétabilité de la méthode

Les exemples de la Figure II.13 illustrent la robustesse de la méthode. Dans cette figure, les images Nomarski originales sont représentées à gauche et les défauts détectés à droite (pour une meilleure visibilité, ces défauts ont été dilatés).

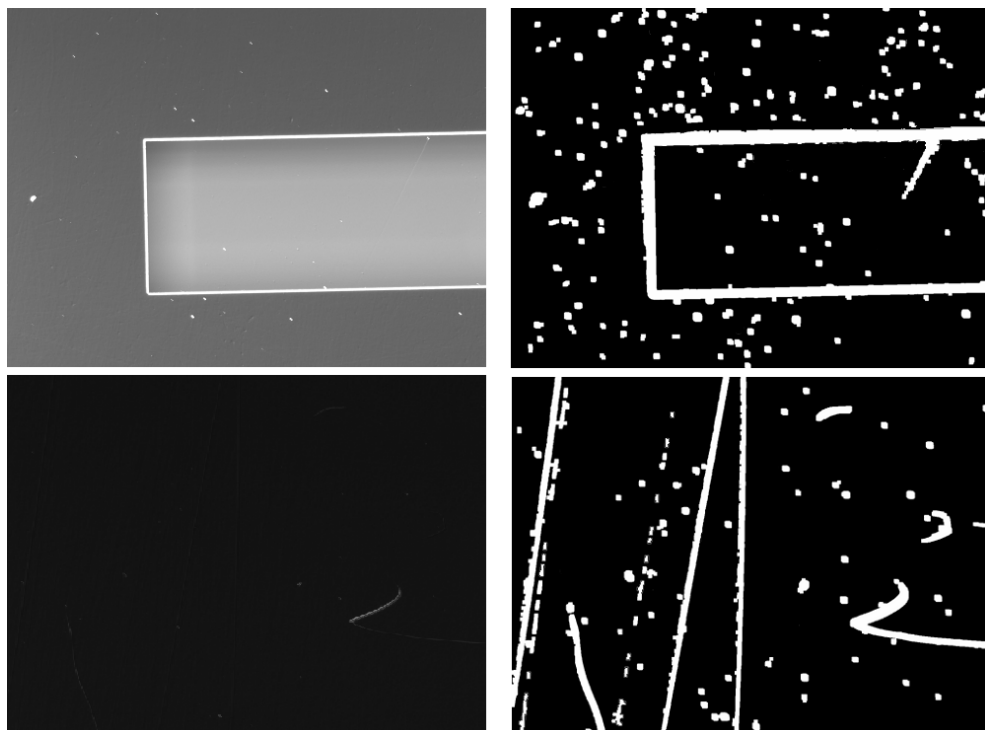


Figure II.13 : Image Nomarski originale (à gauche) et représentation des défauts détectés (à droite).

Le premier exemple montre que la méthode proposée est robuste aux conditions d'éclairage, grâce au prétraitement employé. Ainsi, sous réserve que l'image conserve une dynamique suffisante, la détection des défauts est possible même si l'image est sous/surexposée. Ceci est très intéressant pour notre système, car les conditions d'éclairage peuvent être difficiles à maîtriser (reflets parasites, présence de surfaces aux propriétés réfléchissantes différentes au sein de la même pièce. . .).

Le deuxième exemple représente une bande de traitement déposée sur un substrat. Premièrement, cette bande est constituée d'un matériau différent du substrat, ce qui explique sa différence de teinte et de texture. D'autre part, étant déposée sur le substrat, elle forme un relief, imagé avec un défaut de focus. Malgré cela, les défauts sont détectés avec une sensibilité haute et constante dans les deux parties de l'image. La méthode proposée est donc également robuste à de petits défauts de focus et à la présence de surfaces aux propriétés physiques différentes dans le même champ, ce qui présente un avantage évident pour l'application visée.

Si l'on désire détecter les défauts d'aspect de façon robuste, on doit également les dimensionner avec une répétabilité et une justesse acceptable. Afin d'évaluer la répétabilité et la justesse de la méthode du point de vue des mesures dimensionnelles qu'elle fournit, des expériences ont été menées. Elles ont mis en œuvre trois défauts étalon en forme de disque, de tailles connues et certifiées. Ces défauts ont été imagés grâce à un microscope Nomarski (paragraphe II.2.4.1, p.61), utilisant un grossissement $\times 50$, équipé d'une caméra de résolution 1376×1032 pixels, ce qui amène à une résolution de $1.29 \mu\text{m}$ par pixel. Après traitement des images obtenues, on a calculé le diamètre des défauts détectés, en utilisant différents temps d'exposition de la caméra (influençant ainsi grandement la nature des images récoltées). Les résultats obtenus sont consignés dans les Tableaux II.2, II.3 et II.4. Pour ces expériences, les paramètres de la méthode ont été fixés dans la configuration où la sensibilité est maximale, au détriment de la justesse des contours détectés, ce qui nous place dans la pire situation.

L'écart-type des mesures obtenues est très faible dans les trois cas (bien en dessous du microns, à comparer avec la résolution caméra de $1.29 \mu\text{m}$ par pixel) et ce alors que le temps de pose a été doublé. La mesure est donc très robuste par rapport aux conditions d'éclairage.

| Diamètre du défaut-étalon | Diamètre mesuré (μm) | Temps d'exposition (ms) |
|---------------------------------------|----------------------|-------------------------|
| 16 μm | 19.03 | 45 |
| | 18.87 | 45 |
| | 19.20 | 45 |
| | 18.53 | 55 |
| | 19.53 | 55 |
| | 19.37 | 55 |
| | 18.24 | 65 |
| | 18.47 | 65 |
| | 18.47 | 65 |
| | 18.64 | 75 |
| | 18.75 | 75 |
| | 19.15 | 75 |
| | 18.87 | 90 |
| | 19.03 | 90 |
| 19.09 | 90 | |
| Moyenne | 18.77 | |
| Ecart-type | 0.51 | |
| Surestimation de la taille (%) | 17.31 | |

Table II.2 : Mesures dimensionnelles d'un défaut étalon de 16 μm de diamètre imagé avec différents temps d'exposition.

| Diamètre du défaut-étalon | Diamètre mesuré (μm) | Temps d'exposition (ms) |
|---------------------------------------|----------------------|-------------------------|
| 63 μm | 74.04 | 45 |
| | 74.15 | 45 |
| | 73.76 | 45 |
| | 74.05 | 55 |
| | 74.08 | 55 |
| | 73.98 | 55 |
| | 73.92 | 65 |
| | 73.94 | 65 |
| | 74.09 | 65 |
| | 74.14 | 75 |
| | 74.12 | 75 |
| | 74.05 | 75 |
| | 73.84 | 90 |
| | 73.86 | 90 |
| | 74.09 | 90 |
| Moyenne | 74.00 | |
| Ecart-type | 0.12 | |
| Surestimation de la taille (%) | 17.46 | |

Table II.3 : Mesures dimensionnelles d'un défaut étalon de 63 μm de diamètre imagé avec différents temps d'exposition.

| Diamètre du défaut-étalon | Diamètre mesuré (μm) | Temps d'exposition (ms) |
|---------------------------------------|-----------------------------------|-------------------------|
| 100 μm | 111.83 | 45 |
| | 111.66 | 45 |
| | 111.43 | 45 |
| | 111.58 | 55 |
| | 111.48 | 55 |
| | 111.60 | 55 |
| | 111.25 | 65 |
| | 111.17 | 65 |
| | 111.49 | 65 |
| | 111.56 | 75 |
| | 111.69 | 75 |
| | 111.69 | 75 |
| | 111.78 | 90 |
| | 111.80 | 90 |
| | 111.77 | 90 |
| Moyenne | 111.58 | |
| Ecart-type | 0.20 | |
| Surestimation de la taille (%) | 11.58 | |

Table II.4 : Mesures dimensionnelles d'un défaut étalon de 100 μm de diamètre imagé avec différents temps d'exposition.

La taille des défauts est surestimée d'environ 15%, ce qui est normal, compte tenu du choix des paramètres de la méthode. Cette valeur reste acceptable et elle peut être corrigée en partie puisqu'elle est systématique.

Nous disposons donc maintenant d'une technique permettant de détecter dans une image Nomarski tous les défauts d'aspect significativement plus profonds que la taille caractéristique de la rugosité. Il a été montré, de plus, que cette méthode est robuste vis-à-vis des conditions d'éclairage, de la mise au point et des hétérogénéités de matières. Elle permet en outre, d'obtenir des mesures dimensionnelles correctes et très répétables et d'isoler une représentation image de chaque défaut détecté.

II.2.3 Mesure de la rugosité

Etant donné que la rugosité d'une surface diffuse une partie de la lumière incidente et réduit le contraste des images issues du système dans lequel est intégrée cette surface, ce paramètre physique est d'une grande importance dans l'industrie optique. La micro-

scopie Nomarski, puisqu'elle apporte une information de profondeur, permet d'évaluer ce paramètre.

II.2.3.1 Définition de la rugosité

Il convient tout d'abord de donner une définition précise du terme de « rugosité ». Cette notion a la désagréable particularité d'être définie de manière différente selon le corps de métier qui l'utilise. On se cantonnera donc ici à la définition qui est employée dans le monde de l'optique, qui correspond au paramètre Rq défini comme suit :

$$Rq = \sqrt{\frac{1}{N} \sum_{i=1}^N (\bar{Z} - Z_i)^2} \quad (\text{II.7})$$

Dans cette définition, on dispose de N points de mesures d'altitude Z_i , et d'altitude moyenne \bar{Z} . Le Rq est en fait l'écart-type (Root Mean Square) des altitudes des points mesurés.

Pour avoir un sens, cette définition doit être accompagnée d'une spécification des longueurs d'onde de défauts de surface à prendre en compte.

II.2.3.2 Méthodes de mesure existantes pour évaluer la rugosité

Il existe deux familles de méthodes permettant la mesure de la rugosité (Marioge 2000)(Baker 2004) :

- Les méthodes déterministes : utilisant un profilomètre mécanique, un microscope interférométrique, à effet tunnel (STM) ou à force atomique (AFM), elles fournissent une mesure directe de la microtopographie de la surface. Connaissant l'altitude de nombreux points, on peut appliquer directement la définition précédente pour obtenir le Rq .
- Les méthodes paramétriques : elles fournissent une mesure indirecte de la rugosité, en la déduisant d'un autre paramètre physique. Par exemple c'est la densité spectrale de puissance qui est mesurée réellement ; on utilise dans ce cas un goniodiffusomètre. Un autre paramètre physique qui peut être évalué dans ce but est la quantité de lumière diffusée (Total Integrated Scattering) ; on utilise alors une sphère intégrante.

On peut ajouter à ces méthodes l'utilisation de la microscopie Nomarski. Cette technique est toujours appréciée dans l'industrie optique pour évaluer la qualité du poli d'une surface. L'instrument est facile à utiliser, sans contact et moins coûteux que les autres profilomètres. Il ne donne cependant pas directement de valeur chiffrée : les opérateurs évaluent la rugosité d'une surface en comparant les contrastes d'une image de cette surface avec ceux obtenus avec des étalons de rugosité.

Dans (Jabr 1985), l'auteur propose une méthode permettant d'évaluer de manière précise et automatique la rugosité d'une surface en utilisant ce matériel. Cette technique s'appuie sur le résultat suivant, également démontré dans (Wyant 1985) :

$$\sigma(R) = K\sigma(S) \quad (\text{II.8})$$

où $\sigma(R)$ désigne l'écart-type de l'intensité des pixels d'une image Nomarski de la surface étudiée, $\sigma(S)$ l'écart-type (RMS) de l'altitude de cette surface, et K une constante.

Cette formule est valable si l'on considère que l'intensité mesurée est corrigée par rapport aux variations d'éclairage du fond et au bruit de mesure (et que quelques hypothèses, détaillées plus loin, sont vérifiées). Pour se mettre dans ces conditions, l'auteur propose tout d'abord d'utiliser un grand nombre d'images (256), de les moyenner afin de minimiser le bruit. Dans un deuxième temps, il effectue la même opération sur des images obtenues après un léger défocus, ce qui permet d'évaluer l'intensité du fond. En soustrayant ces deux images, il s'affranchit des variations d'intensité dues aux défauts de type poussière, à la non-uniformité du fond et aux réflexions parasites. Dans l'article est également proposée une validation expérimentale de la théorie ; des mesures sont effectuées sur un microscope interférométrique Wyko et comparées à celles obtenues grâce à la méthode décrite précédemment. On obtient alors une courbe traduisant une bonne corrélation entre $\sigma(R)$ et $\sigma(S)$, et qui permet de déterminer la constante K .

Malgré ces résultats encourageants, cette méthode n'a, à notre connaissance, jamais été utilisée pour des mesures de rugosité, sans doute à cause du nombre d'images qu'elle exige.

II.2.3.3 Méthode de mesure proposée

On se propose ici d'utiliser le même résultat théorique (donnée par l'Equation II.8) mais en évitant l'utilisation de multiples images : on essaye d'évaluer simplement $\sigma(R)$.

Pour cela, on va considérer que l'intensité I d'un pixel d'une image Nomarski est donné par :

$$I = FR + n \quad (\text{II.9})$$

où F désigne l'éclairage du fond, R l'intensité due au relief de la pièce et n le bruit de mesure. Ce formalisme est conforme à la théorie (Lessor, Hartman, and Gordon 1979) : le fond à un effet multiplicatif sur l'intensité mesurée.

On démontre facilement que :

$$V(FR) = V(F)V(R) + V(F)E(R)^2 + V(R)E(F)^2$$

Soit :

$$V(R) = \frac{V(FR) - V(F)E(R)^2}{V(F) + E(F)^2} \quad (\text{II.10})$$

où E désigne l'espérance mathématique et V la variance.

On doit donc déterminer les quantités $V(FR)$, $V(F)$, $E(F)$ et $E(R)$.

– A cause de l'indépendance de n et FR , on a :

$$V(I) = V(FR) + V(n)$$

Soit :

$$V(FR) = V(I) - V(n) \quad (\text{II.11})$$

Or on peut calculer $V(I)$ directement à partir d'une image Nomarski du champ à étudier. On peut obtenir $V(n)$ en étudiant deux images (1) et (2) du même champ acquises dans les mêmes conditions. Pour cela on crée l'image Δ de la différence pixel à pixel des images 1 et 2. L'intensité I d'un pixel dans cette image est donnée par :

$$I_{\Delta} = I_1 - I_2 = FR + n_1 - FR - n_2 = n_1 - n_2 \quad (\text{II.12})$$

On a alors :

$$V(I_{\Delta}) = V(n_1) + V(n_2) = 2V(n) \quad (\text{II.13})$$

car n_1 et n_2 sont indépendants.

On a donc :

$$\boxed{V(FR) = V(I) - \frac{V(I_\Delta)}{2}} \quad (\text{II.14})$$

- Pour trouver les caractéristiques de F , on utilise la technique de filtrage homomorphique. On considère le logarithme d'une image Nomarski du champ à étudier. On a alors :

$$\log I = \log(FR + n) = \log F + \log R + \log\left(1 + \frac{n}{FR}\right) \quad (\text{II.15})$$

Dans cette expression, seul $\log F$ est constitué de composantes basses fréquences (défaut d'illumination) alors que les autres termes sont constitués de hautes et moyennes fréquences (bruit et rugosité). On peut donc séparer ce terme en appliquant un filtrage passe bas au logarithme de l'image. En prenant l'exponentiel de l'image obtenue, on reconstitue l'image de F et on peut alors calculer $V(F)$ et $E(F)$.

- On a :

$$E(FR) = E(R)E(F)$$

car F et R sont indépendants. Et donc :

$$E(R) = \frac{E(FR)}{E(F)}. \quad (\text{II.16})$$

Or on a :

$$E(I) = E(FR + n) = E(FR) + E(n) = E(FR) \quad (\text{II.17})$$

en supposant que l'on a un bruit à moyenne nulle.

On peut écrire alors :

$$\boxed{E(R) = \frac{E(I)}{E(F)}} \quad (\text{II.18})$$

De toutes ces considérations, on peut finalement tirer un schéma d'évaluation de la rugosité d'une pièce sur un champ donné, à partir de deux images :

- Acquérir deux images du champ dans les mêmes conditions.
- Former l'image différence des deux images acquises ; en calculer la variance $V(I_\Delta)$
- Sur une des images, calculer la variance $V(I)$ et la moyenne $E(I)$.
- Prendre le logarithme d'une des images, lui appliquer un passe-bas et enfin repasser à l'exponentielle. Calculer la variance $V(F)$ et la moyenne $E(F)$ du résultat.
- A l'aide des expressions II.10, II.14 et II.18 calculer $V(R)$.
- Finalement calculer la rugosité grâce à l'Equation II.8.

On peut rajouter en préambule de ce schéma, la détection des défauts significativement plus profonds que l'ordre de grandeur de la rugosité (paragraphe II.2.2, p.42), et ne plus considérer ensuite les pixels correspondant à des défauts. La présence de défauts peut en effet influencer grandement le résultat obtenu, problème qui existe d'ailleurs avec les autres procédés de mesure de la rugosité. D'autre part, la dernière étape du schéma, sous-entend que l'on connaisse la valeur de la constante K , d'où le besoin d'une *calibration expérimentale* de la méthode.

Enfin, on peut récapituler ici les quelques hypothèses qui sont faites sur les images à traiter :

- Bruit blanc, à moyenne nulle.
- Indépendance du bruit, de l'éclairage et du relief de la pièce.
- Rugosité à fréquences significativement plus hautes que les défauts d'éclairage.
- Hypothèses de (Jabr 1985) :
 - Surface à moyenne nulle
 - Shearing (décalage dû au prisme de Wollaston) s assez grand pour annuler l'autocorrélation entre $Z(x)$ et $Z(x + s)$

II.2.3.4 Répétabilité de la méthode

Pour évaluer la robustesse et la répétabilité de la méthode, des mesures ont été effectuées, d'une part pour des temps de pose différents (Tableau II.5), d'autre part pour 5 positions différentes de la pupille d'éclairage, ce qui produit 5 conditions d'éclairage différentes (Tableau II.6).

| Temps de Pose | RMS Mesuré |
|-------------------------------|-------------------|
| 20 | 0.005128285 |
| 26 | 0.005028242 |
| 30 | 0.005053057 |
| 36 | 0.004989438 |
| 40 | 0.004969354 |
| 46 | 0.00490609 |
| 50 | 0.004879957 |
| 56 | 0.004893262 |
| 60 | 0.004833605 |
| Moyenne | 4,964587778 |
| Ecart-Type | 0,09496309 |
| Moyenne/Ecart-Type (%) | 1,912809162 |

Table II.5 : Répétabilité par rapport au temps de pose.

| Position de la pupille | RMS mesuré |
|---------------------------------|-------------------|
| 1 | 0,006765738 |
| 2 | 0,006381274 |
| 3 | 0,006397052 |
| 4 | 0,00641958 |
| 5 | 0,006283768 |
| Moyenne | 0,006449482 |
| Ecart-Type | 0,000184238 |
| Ecart-Type / Moyenne (%) | 2,856628787 |

Table II.6 : Répétabilité par rapport à la position de la pupille.

Ces résultats sont concluants, la méthode semble robuste au changement d'intensité et de forme de l'éclairage. Il faudrait cependant effectuer plus de mesures pour mieux caractériser cette propriété.

II.2.3.5 Etalonnage de la méthode

Dans le but d'effectuer la calibration nécessaire de la méthode (déterminer expérimentalement le coefficient K), des mesures ont été effectuées sur des étalons de rugosité. Les dix étalons utilisés couvrent la gamme de rugosités de 0,3 nm à 1,1 nm RMS. On a d'une part mesuré la rugosité sur trois champs différents de chaque échantillon avec un microscope interférométrique Wyko 1100. On a d'autre part acquis grâce à un microscope Nomarski (paragraphe II.2.4.1, p.61) des séries de deux images sur trente champs des mêmes spécimens et ensuite appliqué la méthode proposée. Les résultats obtenus sont présentés dans le Tableau II.7 et la Figure II.14.

| Echantillon | Mesure Wyko | | | | Calcul par la méthode proposée | |
|-------------|-------------|-------|-------|---------|--------------------------------|---------|
| | $M1$ | $M2$ | $M3$ | Moyenne | Ecart-type | Moyenne |
| 5003 | 11,78 | 11,18 | 11,60 | 11,52 | 9.36 | 189.07 |
| 5004 | 7,29 | 6,62 | 7,17 | 7,03 | 5.08 | 99.72 |
| 5005 | 7,56 | 8,07 | 7,99 | 7,87 | 4.18 | 122.19 |
| 5006 | 5,83 | 6,34 | 6,25 | 6,14 | 3.72 | 100.64 |
| 5007 | 3,83 | 4,10 | 4,14 | 4,02 | 7.13 | 73.89 |
| 5008 | 2,99 | 2,76 | 2,77 | 2,84 | 8.72 | 51.43 |
| 5013 | 10,48 | 10,57 | 12,10 | 11.05 | 8.39 | 176.02 |
| 5014 | 6,65 | 7,45 | 6,80 | 6,97 | 3.53 | 104.44 |
| 5016 | 5,74 | 5,96 | 5,66 | 5,79 | 8.65 | 100.35 |
| 5018 | 3,38 | 3,08 | 3,03 | 3,16 | 10.75 | 47.29 |

Table II.7 : Mesures de rugosité (en Å) et valeur calculée par la méthode proposée.

La Figure II.14 est comparable à la figure de corrélation de (Jabr 1985), ce qui indique que la méthode proposée est pertinente (on obtient des résultats d'une qualité équivalente avec une méthode beaucoup moins contraignante au niveau de l'utilisation : 2 images à acquérir au lieu de 512, pas de défocus nécessaire). En outre, le coefficient de corrélation élevé (environ 0,98) entre le paramètre que l'on calcule à partir de deux images Nomarski et la rugosité nous semble suffisant pour permettre des mesures de rugosité correctes.

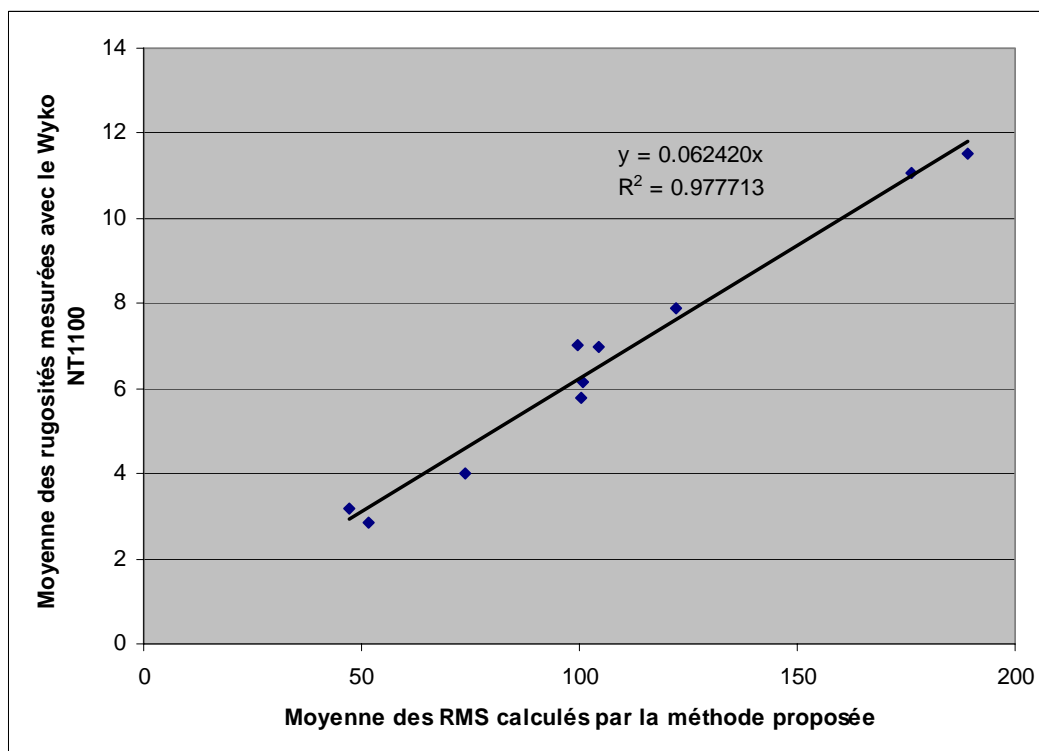


Figure II.14 : Graphique correspondant aux données du Tableau II.7, avec la régression linéaire correspondante.

Néanmoins, cette courbe de calibration pâtit de plusieurs imperfections expérimentales qui pourraient être corrigées. Premièrement, les champs étudiés conjointement avec le microscope Nomarski et le Wyko n'étaient pas en correspondance directe ; en effet on s'est contenté de choisir des champs dans la même zone de l'échantillon. On se trouve alors devant le problème de variation locale de la rugosité sur les échantillons (cf Tableau II.7 : des mesures Wyko dans la même zone peuvent varier de 10%). D'autre part, les longueurs d'onde accessibles par le Wyko et le microscope Nomarski ne sont pas les mêmes (période minimale de $3.9 \mu\text{m}$ pour le Wyko, période minimale de $2.6 \mu\text{m}$ pour le microscope Nomarski). La rugosité mesurée dans les deux cas n'est de toute façon pas tout à fait la même. Il paraît donc pertinent d'effectuer la calibration dans de meilleures conditions pour améliorer les résultats obtenus : il faudrait repérer des champs bien précis sur chaque échantillon, et trouver un moyen d'étudier la même gamme de fréquence (en sous-échantillonnant correctement les images données par le microscope par exemple, ou en employant un autre moyen que le Wyko NT1100).

On dispose donc maintenant d'une technique permettant de mesurer la rugosité d'une surface à partir de deux images Nomarski de celle-ci. Ceci nous permet d'atteindre un paramètre physique très important pour la fabrication optique, à un moindre coût. On a montré de plus que cette méthode est robuste vis-à-vis des conditions d'éclairage (temps de pose caméra et forme de la pupille d'éclairage). Elle permet d'obtenir des résultats corrects, même si l'étape de calibration pourrait être améliorée. Il est à noter sur ce point que cette calibration dépend des réglages du microscope Nomarski (position du prisme de Wollaston, orientation de l'analyseur).

II.2.4 Implémentation et validation des méthodes

II.2.4.1 Présentation du prototype

Afin de tester et valider les méthodes proposées dans le cadre de contrôles de pièces optiques dans les conditions réelles de production, un prototype a été réalisé. Ce système se base sur un microscope Olympus BX-52 (Figure II.15).

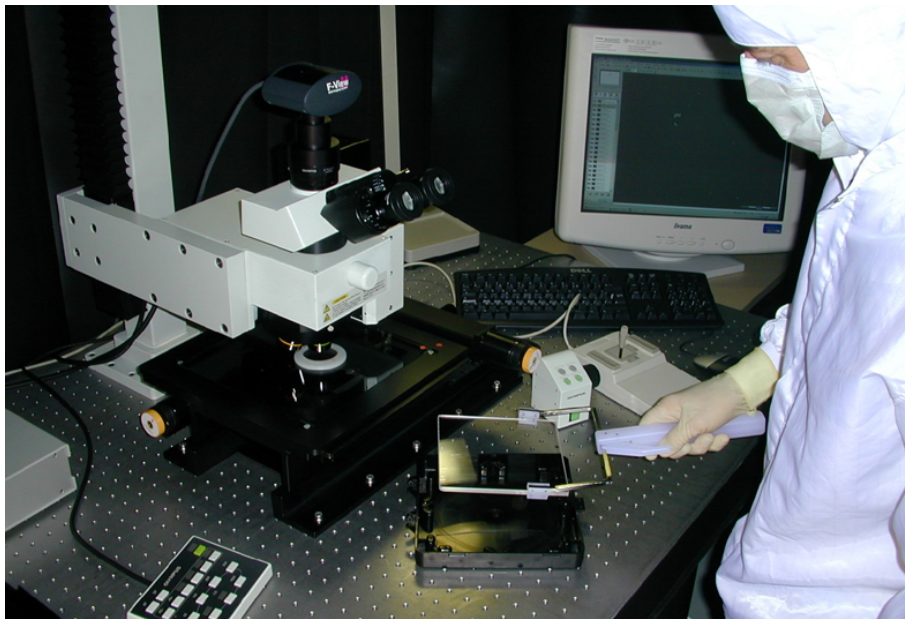


Figure II.15 : Le microscope Olympus BX-52.

Ce microscope permet d'utiliser différents modes d'éclairage : Bright Field, Dark Field et Nomarski. Il dispose d'objectifs de grossissement x5, x10, x20, x50 et x100, ainsi que d'un oculaire fixe d'un grossissement x10. Une caméra CCD de 1376x1032 pixels lui est associée et permet d'acquérir des images numériques (codées sur 12 bits). Typiquement,

on utilisera ici l'objectif de grossissement x5, qui amène à une résolution caméra de 1,29 μm par pixel, résolution suffisante pour l'application visée. Dans ce cas, chaque image acquise représente un champ microscopique d'environ 1,77mm x 1,33mm.

Afin d'étudier des pièces optiques entières, ce microscope est associé à une platine, qui permet de translater la pièce étudiée selon les axes X et Y. Le tout est géré par un logiciel qui donne accès à l'utilisateur aux fonctions du microscope de manière ergonomique, en particulier à la définition de scan automatique de pièce. Cette configuration matérielle existante a été modifiée afin d'intégrer les méthodes de traitement proposées plus haut et d'exploiter les mesures qu'elles fournissent pour le contrôle d'aspect de pièces optiques de taille limitée. La Figure II.16 présente le diagramme du flux de données lors d'un contrôle.

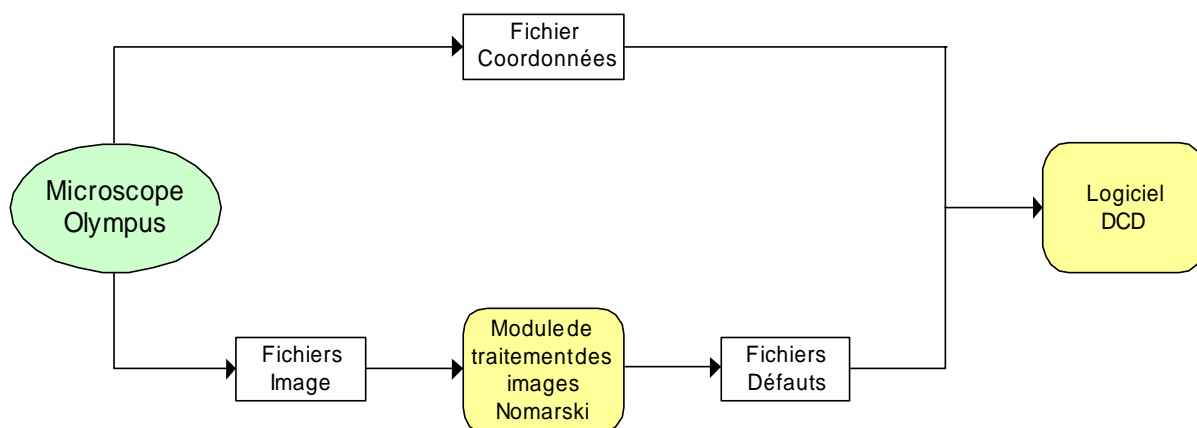


Figure II.16 : Flux de données lors de la mesure d'une pièce optique avec le prototype.

Les trois entités créant ou traitant l'information sont :

- le microscope Olympus qui fournit des fichiers « Image » (où sont enregistrées les représentations Nomarski des champs microscopiques explorés lors du scan) et un fichier « Coordonnées » (qui contient les coordonnées de ces différents champs),
- le module de traitement des images Nomarski qui traite les fichiers « Image » pour donner des fichiers « Défauts » (où sont sauvegardées les descriptions de tous les défauts détectés),
- le logiciel DCD qui exploite le fichier « Coordonnées » et les fichiers « Défauts » et met en forme les informations recueillies pour l'opérateur (son fonctionnement est détaillé dans le paragraphe II.2.4.2).

On souhaite étudier une pièce optique en partie ou dans sa totalité, c'est-à-dire un « champ global » constitué de nombreux champs microscopiques. Dans ce but, le logiciel qui pilote le microscope Olympus permet de créer et d'exécuter simplement un programme de « scan » : la définition par l'utilisateur d'une zone d'intérêt engendre automatiquement une série de positions platine qui seront « visitées » par le microscope. Ces positions décrivent un certain nombre de champs, qui couvrent l'intégralité de la zone d'intérêt sans se chevaucher ; le logiciel permet par ailleurs de définir une série d'actions qui seront réalisées après chaque déplacement de la platine.

Dans notre cas, à chaque position du programme de scan, on appelle :

- Une fonction qui lance le module de traitement des images Nomarski (dans les faits, ce module qui s'exécute comme un processus indépendant n'est lancé qu'une seule fois)
- Une fonction qui permet d'acquérir une image au format caméra (12 bits),
- Une fonction qui convertit cette image au format 8 bits,
- Une fonction qui sauvegarde cette image dans un dossier prédéfini (fichier « Image »),
- Une fonction qui met à jour le fichier « Coordonnées » en lui ajoutant les coordonnées X, Y et Z du champ courant.

Au final, on obtient un certain nombre de fichiers « Image » (autant que de champs étudiés) et un fichier « Coordonnées ». Les fichiers « Image » sont des fichiers contenant des images au format brut TIFF 8 bits, enregistrés sous le nom IMAGEXXX.tif, où XXX désigne le numéro du champ dont l'image est sauvegardée. Le fichier « Coordonnées » est un fichier contenant les coordonnées en X, Y et Z de tous les champs étudiés. Ce fichier, après avoir précisé en entête le nombre de champs qu'il décrit adopte la syntaxe suivante :

n ° Image CX CY CZ <fin de ligne>

où CX, CY, et CZ désignent respectivement les coordonnées en X, Y et Z du champ considéré.

Parallèlement à cela, le module de traitement des images Nomarski est exécuté. Il implémente de manière performante (en termes de temps d'exécution) la méthode de détection des défauts proposée dans le paragraphe II.2.2. Ainsi les fichiers « Image » sont

traités en ligne pour donner des fichiers « Défauts » et sont ensuite détruits. Les fichiers « Défauts » sont des fichiers contenant les descripteurs de tous les défauts détectés dans le champ associé. Ils sont enregistrés dans un dossier prédéfini, sous le nom IMAGEXXX.txt, où XXX désigne le numéro du champ considéré. Ces fichiers, après avoir précisé en entête le nombre de défauts qu'ils décrivent adoptent la syntaxe suivante :

```
Defaut
X x1 Y y11 y12
X x2 Y y21 y22
X x3 Y y31 y32
....
Defaut
X ...
```

où xi désigne une abscisse dans l'image dont les points d'ordonnées situées entre yi1 et yi2 font partie du défaut considéré.

L'implémentation du module de traitement des images Nomarski permet son exécution en temps masqué : l'image d'un champ est traitée pendant que la platine se déplace vers la prochaine position. Ainsi, à la fin de l'opération de scan, les données sont directement disponibles pour le logiciel de post-traitement qui va permettre leur interprétation.

Il est à noter que si l'on veut étudier simultanément la rugosité de la pièce, le programme de « scan » est légèrement modifié (il faut acquérir deux images au lieu d'une).

II.2.4.2 Le logiciel de post-traitement

Les fichiers résultant du scan, les fichiers « Défauts » et le fichier « Coordonnées », ont un format brut. Le logiciel DCD (Détection et Caractérisation de Défauts) que nous avons développé a pour but de traiter ces données brutes pour fournir un listing des défauts présents sur la pièce étudiée ainsi que différentes informations les concernant. Il permet également de ne conserver que les défauts répondant à des contraintes de taille ou de positionnement et d'afficher une cartographie de la pièce étudiée, c'est-à-dire de fournir des informations directement exploitables par les opérateurs.

Le traitement des défauts bruts par le logiciel DCD s'effectue en plusieurs étapes :

- Conversion des coordonnées des points des défauts en coordonnées réelles (dans le repère platine)
- Tri des défauts par rapport au masquage spatial des données demandé
- Rassemblement des défauts appartenant à des champs différents mais faisant partie de la même entité
- Calcul des paramètres morphologiques pour chacun des défauts
- Tri des défauts par rapport à leurs paramètres morphologiques.
- Génération de la liste des défauts et de la cartographie de la pièce.

Il est à noter qu'une seconde phase de rassemblement des défauts visant cette fois à réunifier des défauts de type « rayure discontinue » est envisagée dans une évolution future. D'autre part, nous décrivons ici uniquement les points de ce traitement, délicats ou présentant un intérêt particulier.

Rassemblement des défauts partagés sur plusieurs champs :

Il peut arriver qu'un défaut (par exemple une rayure longue) soit imagé sur plusieurs champs et donc soit représenté dans les fichiers « Défauts » par plusieurs éléments ; il faut alors rassembler ces divers éléments pour reconstituer le défaut.

Pour chaque défaut en contact avec le bord du champ auquel il appartient, on considère les 4 champs adjacents. On teste si les défauts en contact avec le bord de ces champs sont contigus avec le défaut considéré. Si cela est le cas, les deux défauts n'en sont en fait qu'un : il faut les assembler ; on place pour cela le couple constitué par leurs deux identifiants dans une liste des rassemblements à effectuer. Une fois tous les défauts examinés, on traite les différentes fusions de défauts.

Discrimination entre défauts de type « piqûre » et défauts de type « rayure » :

La plupart des normes (voir Annexe A) concernant les défauts d'aspect traitent de manière différente les défauts ponctuels (les « piqûres ») et les défauts longilignes (les « rayures »). Dans le logiciel DCD, la distinction entre ces deux types de défauts est basée sur le ratio d'aspect. Si ce paramètre, calculé comme le rapport entre longueur a et largeur b du défaut (voir Figure II.17), est plus grand qu'un seuil fixé par l'opérateur (typiquement 3), le défaut est considéré comme une « rayure », autrement c'est une « piqûre » .

Calcul des paramètres morphologiques d'un défaut :

– Orientation :

L'orientation d'un défaut est donnée par la formule :

$$O = \arctan\left(\frac{\lambda - Cxx}{Cxy}\right) \quad (\text{II.19})$$

avec

$$\lambda = \frac{1}{2}(Cxx + Cyy + \sqrt{(Cxx - Cyy)^2 + 4Cxy^2})$$

où :

$$Cxx = \frac{1}{n} \sum_i (x_i - \bar{x})^2$$

$$Cyy = \frac{1}{n} \sum_i (y_i - \bar{y})^2$$

$$Cxy = \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})$$

avec n le nombre de points de coordonnées $(x_i; y_i)$ composant le défaut et $(\bar{x}; \bar{y})$ les coordonnées de son barycentre dans l'image.

– Largeur, Longueur :

Une fois l'orientation d'un défaut calculée, on détermine le plus petit rectangle qui l'englobe, tenant compte de son orientation, comme décrit dans la Figure II.17.

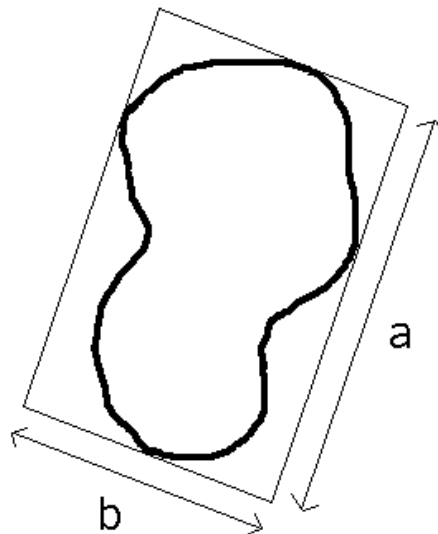


Figure II.17 : Rectangle englobant orienté d'un défaut.

La largeur l et la longueur L du défaut sont alors données comme suit :

- si le défaut est considéré comme une piqure :

$$L = a$$

$$l = b$$

- si le défaut est considéré comme une rayure (défaut longiligne) :

$$L = \sqrt{a^2 + b^2}$$

$$l = \frac{A}{L}$$

Il est à noter qu'un calcul plus précis de ces deux paramètres est envisageable. On pourrait par exemple effectuer une opération de squelettisation du défaut puis calculer sa longueur en évaluant l'abscisse curviligne de ce squelette et enfin en déduire sa largeur.

II.2.4.3 Complément au prototype : recalage de cartes

Grâce au prototype, on est capable de générer des cartographies de pièces optiques entières. Il peut alors être intéressant de pouvoir confronter deux cartographies de la même pièce obtenues à deux instants différents. En effet, on peut juger, en comparant ces deux cartes, de la qualité d'un process (apparition ou disparition de défauts entre les deux mesures), de l'efficacité d'une étape de nettoyage, etc. . .

Cependant, le positionnement des défauts dans les cartographies générées par le logiciel par DCD se fait dans le repère absolu de la platine. Ainsi pour que deux cartes de défauts soient directement comparables, il faut que la pièce ait été exactement dans la même position lors des deux mesures. Comme ceci est impossible, on a besoin d'une méthode qui permette de retrouver, à partir de deux cartographies de pièce, les paramètres de la transformation mettant en correspondance les repères liés à la pièce pour les deux mesures. En appliquant cette transformation, la comparaison directe des deux cartes est alors possible.

A noter que l'une des applications particulières de cette technique concerne les travaux exposés dans la suite de cette thèse : en comparant deux cartographies de la même pièce avant et après nettoyage (et en faisant l'hypothèse que ce nettoyage soit assez efficace pour

éliminer ou faire « bouger » toutes les poussières et autres artefacts), on pourrait, sans l'intervention d'experts, déterminer la nature d'une partie des défauts détectés. En effet, sous l'hypothèse exposée plus haut, les défauts communs aux deux cartographies seraient des « vrais » défauts, les autres des artefacts. Cela faciliterait énormément la constitution d'une base d'items labellisés de grande taille, pouvant être utilisée pour l'apprentissage du système de discrimination des défauts.

La difficulté principale de la mise en correspondance de deux cartographies d'une même pièce à deux instants différents est la présence, à côté des défauts qui leurs sont communs, d'items qui n'existent que dans une des deux cartographies. Ces items peuvent correspondre à des poussières, des traces de nettoyage (qui ont disparues, apparues, ou changées de place entre les deux mesures) ou à des défauts qui sont apparus/ont disparus pendant le process subi par la pièce entre les deux mesures. La technique proposée fonctionne donc en trois étapes : la première consiste à déterminer les couples de correspondance entre items communs aux deux cartographies, la deuxième a pour but d'optimiser les paramètres de rotation et de translation de la transformation liant ces couples, et la troisième d'appliquer effectivement cette transformation à l'une des cartographies pour obtenir enfin deux cartographies dans le même repère.

Couples de correspondance entre items

Les données de départ sont des défauts appartenant à deux cartographies distinctes, constitués d'un certain nombre de points dont les coordonnées sont exprimées dans le repère de la platine (ces points correspondent en fait aux pixels caméra issus du microscope, détectés comme faisant partie d'un défaut).

Pour chaque item de chaque cartographie, on cherche l'item le plus « ressemblant » dans l'autre cartographie. Pour cela, chaque défaut est représenté par un vecteur constitué de ses 6 premiers moments de Hu (Hu 1962) : ces moments, invariants en translation et en rotation sont faciles à calculer et représentent correctement les formes complexes. On effectue une première sélection par rapport à la taille des défauts : le défaut cherché doit avoir une taille (en nombre de pixels) très proche du défaut considéré. Ensuite on trie les défauts restant par rapport à leur distance de Mahalanobis au défaut considéré. Cette distance, qui mesure en fait la similarité entre deux vecteurs \vec{x} et \vec{y} , est définie comme

suit :

$$d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})} \quad (\text{II.20})$$

où S désigne la matrice de covariance de l'ensemble des vecteurs représentant les défauts. Si le défaut le plus proche au sens de cette distance, l'est suffisamment, un nouveau couple de correspondance hypothétique est créé (ce couple met en correspondance les centres de gravité des deux défauts).

Malheureusement, si les deux cartographies ont dans la réalité peu de points communs, la plupart des couples constitués ne doivent pas être pris en compte. Afin de retrouver les couples de correspondance authentiques, on forme tous les triplets de couples hypothétiques possibles. A l'aide de ces triplets (trois points de la cartographie 1 correspondant à trois points de la cartographie 2), on peut former deux vecteurs distincts et leurs images dans l'autre cartographie. Pour chaque triplet, il est alors possible de calculer l'unique déplacement (une rotation et une translation) transformant les deux vecteurs en leurs images. On obtient trois paramètres pour chaque triplet : une composante de translation en X, une composante de translation en Y, et un angle de rotation. Les « faux » couples vont produire une distribution aléatoire de déplacements, tandis que les « vrais » couples vont engendrer des déplacements aux paramètres très similaires, on peut donc les identifier.

Optimisation des paramètres

Cependant le précédent calcul des paramètres des déplacements est très sensible numériquement, ce qui explique que les vrais couples ne donnent pas exactement les mêmes paramètres. Une optimisation est nécessaire, elle est effectuée par la méthode de la décomposition en valeurs singulières (SVD : Singular Value Decomposition). L'algorithme SVD se résume de la façon suivante (Tarel 1996) :

- Calculer la matrice de covariance croisée Q des deux ensembles de points :

$$Q = \frac{1}{n} \sum_{i=1}^n Y'_i Y_i^T$$

où Y_i et Y'_i désignent les n points X_i et X'_i à mettre en correspondance, exprimés dans les repères centrés sur leurs barycentres.

- Trouver la décomposition en valeurs singulières de Q :

$$Q = ZDW^T$$

où Z et W sont des rotations et D est diagonale.

- La meilleure rotation est alors $R = ZW^T$ et la meilleure translation $T = \bar{X}'_i - R\bar{X}_i$.

Il suffit alors d'appliquer cette rotation et cette translation à l'une des cartes, sans oublier de s'adapter au pas de discrétisation utilisé par l'autre carte : on obtient alors deux cartographies directement comparables point à point.

II.2.4.4 Validation du prototype

On présente ici en guise de validation du prototype, la cartographie des défauts d'aspect d'une pièce de 100mm x 100mm réalisée à l'aide de notre prototype (Figure II.18).

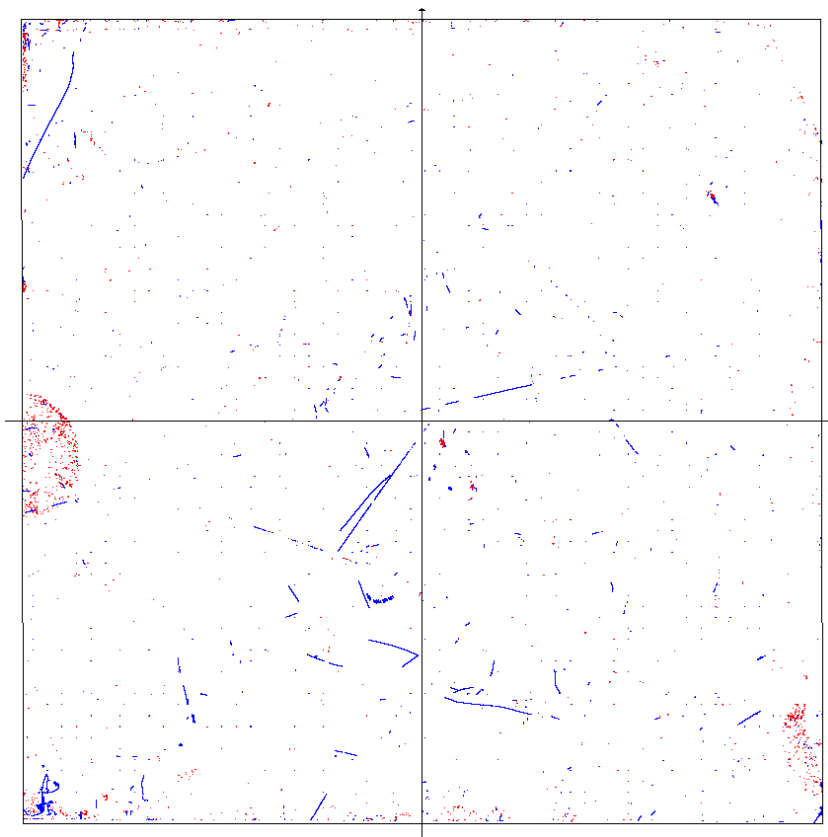


Figure II.18 : Cartographie des défauts d'aspect d'une pièce optique de 100mm x 100mm réalisée avec le prototype présenté.

Elle est issue de l'exploitation de 4070 champs microscopiques de dimension 1,77mm x 1,33mm. La phase de mesure correspondante a duré environ 2H45 : une surface d'environ un mm² est explorée en une seconde (pour l'objectif x5). Dans cette cartographie, les défauts de type « piqûre » sont représentés en rouge et les défauts de type « rayure » en bleu.

Le deuxième exemple (Figure II.19) montre une cartographie de rugosité (mesurée grâce à la méthode présentée au paragraphe II.2.3.2) de la même pièce. On peut apprécier sur cette figure, l'environnement logiciel ergonomique, permettant aux opérateurs d'effectuer simplement de nombreuses opérations d'exploitation de données mais également d'exporter ces données pour la rédaction d'un procès verbal de contrôle.

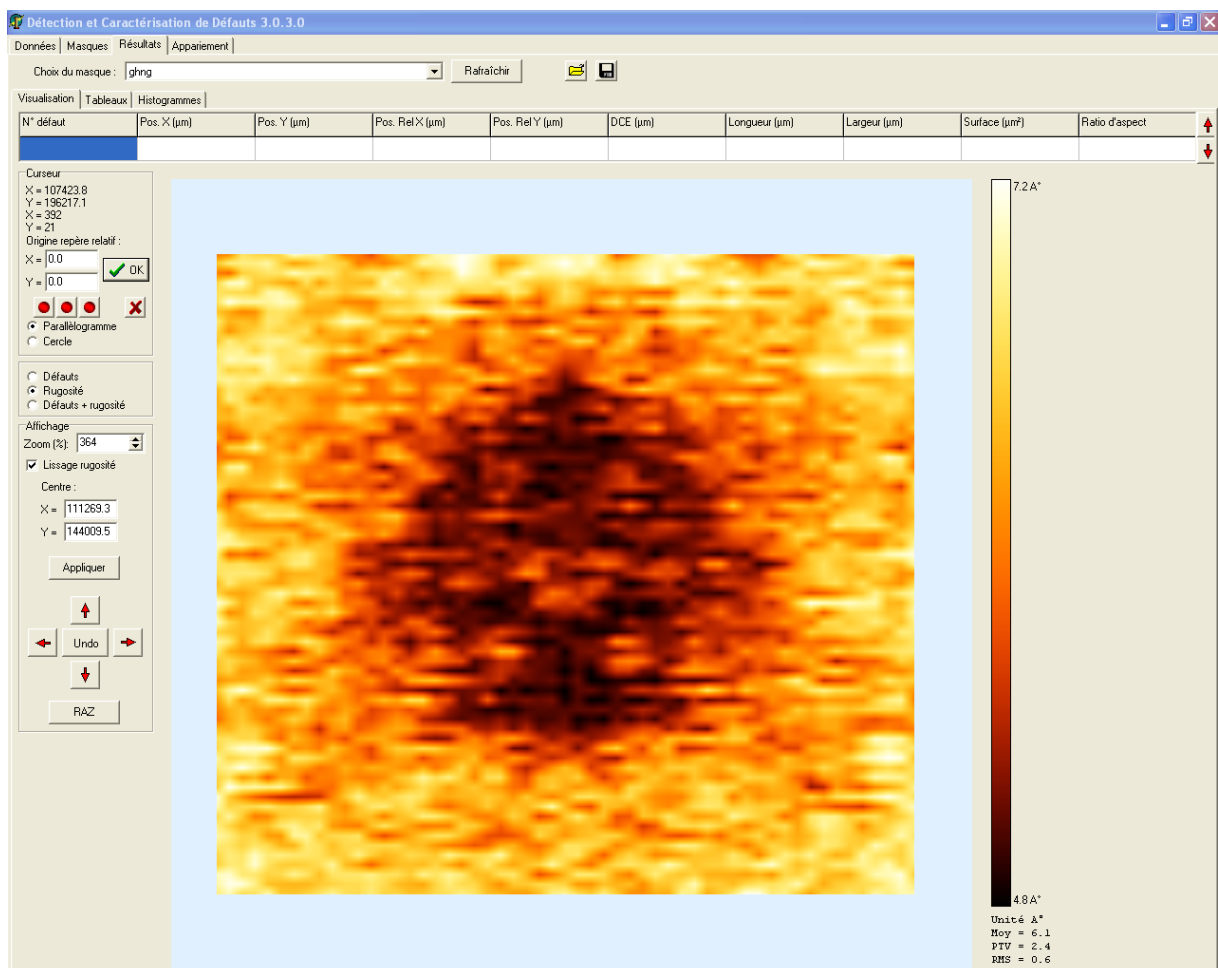


Figure II.19 : Cartographie de rugosité d'une pièce optique de 100mm x 100mm réalisée avec le prototype présenté.

De plus, afin d'illustrer le fonctionnement de la méthode de recalage de cartes, on a étudié une pièce optique circulaire de diamètre 42mm, ce qui correspond à 620 champs microscopiques avec l'objectif x5. A l'aide du prototype, une première mesure du miroir est effectuée sans nettoyage préalable. On peut voir sur la cartographie correspondante (Figure II.20) un certain nombre de défauts et surtout des traces de « slurry » (liquide de polissage).

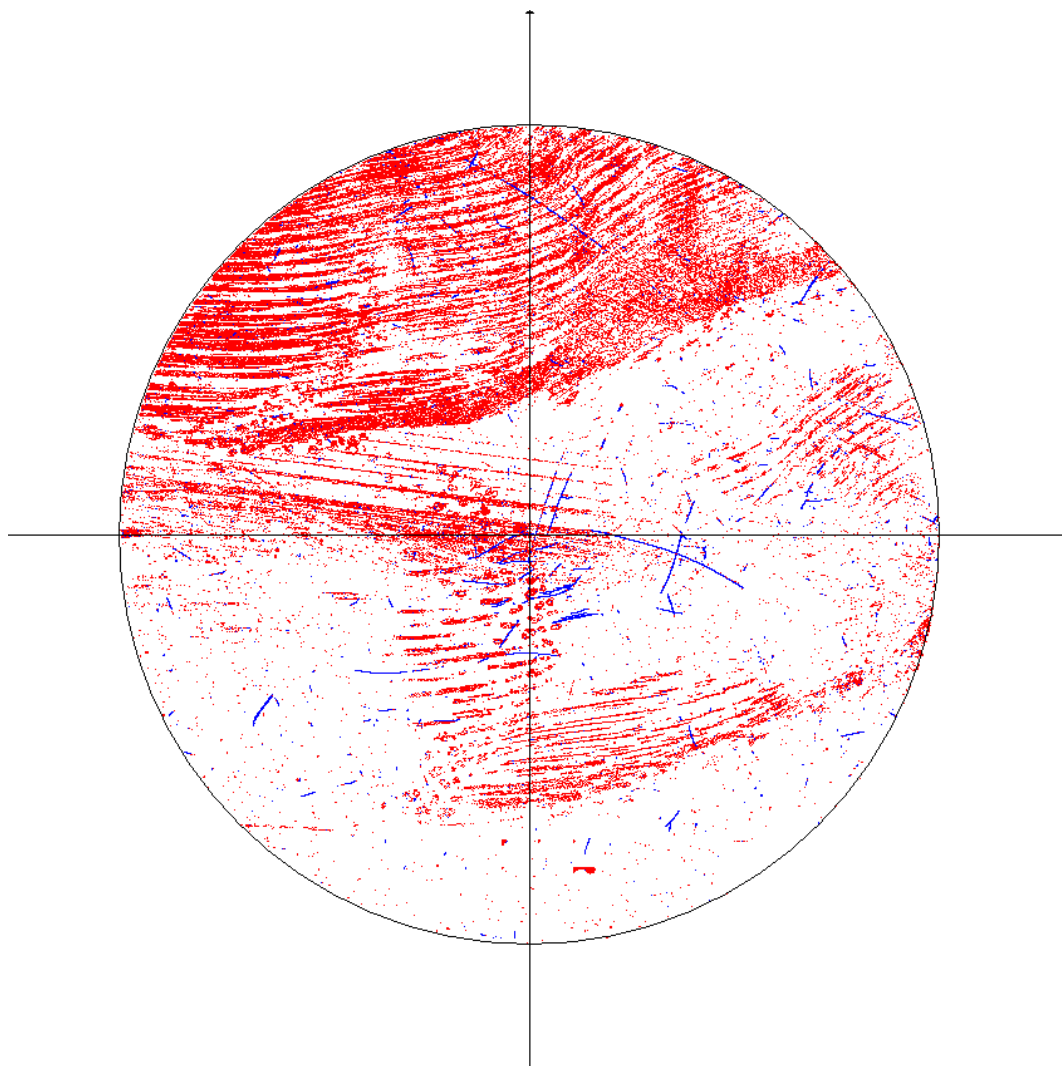


Figure II.20 : Cartographie d'une pièce optique de diamètre 42mm.

Après un nettoyage assez grossier (il reste des traces!), on effectue une deuxième mesure (Figure II.21). Pour cette deuxième mesure, le miroir a été tourné d'une trentaine de degrés et a été déplacé globalement sur la platine. Les traces de « slurry » ont pour la plupart disparues mais on remarque certains défauts déjà présents dans la première cartographie (en particulier des raies au centre du champ).

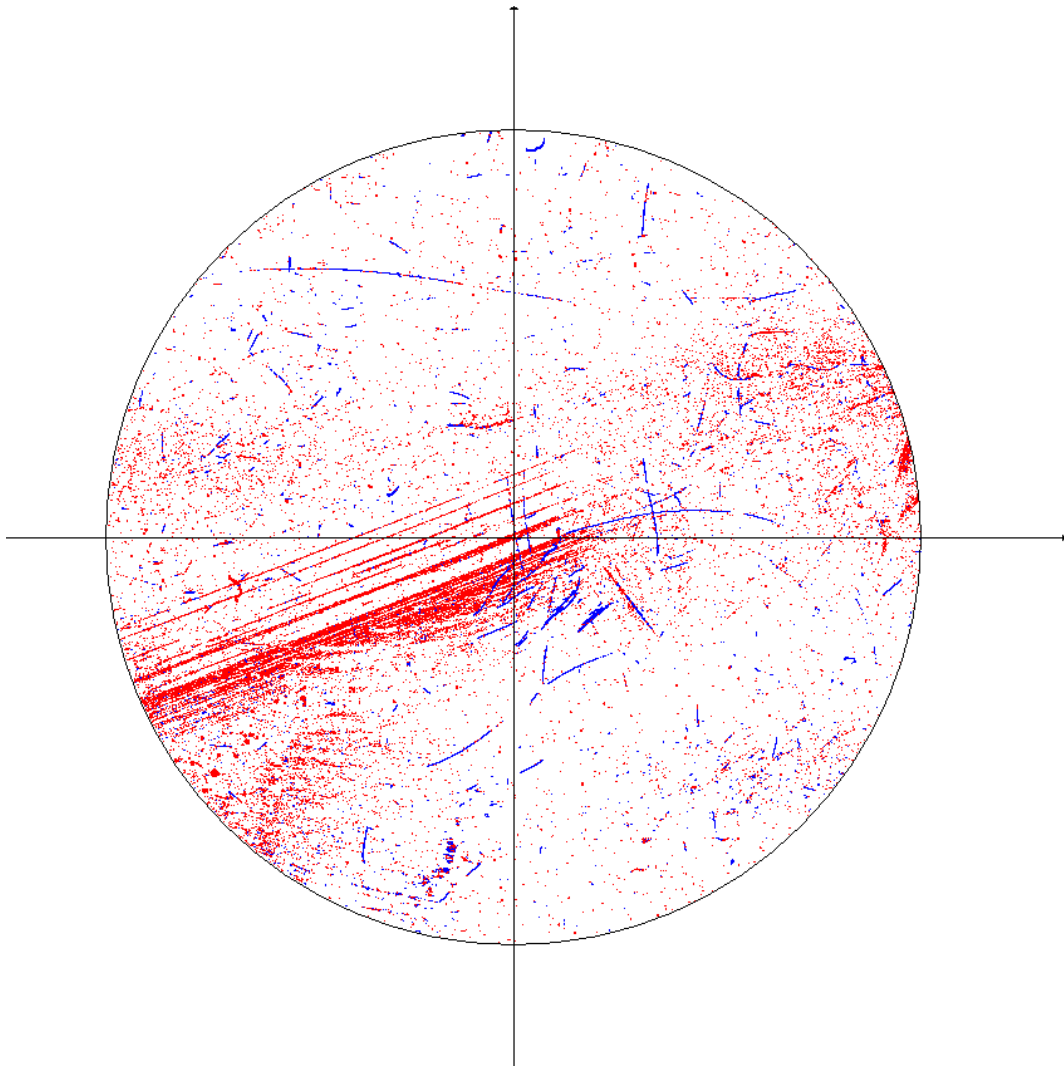


Figure II.21 : Cartographie de la même pièce (Figure II.20) après nettoyage.

En appliquant la technique de recalage de carte proposée, on est capable de retrouver les paramètres du déplacement entre les deux mesures et ainsi de recalcr les deux cartes dans le même repère. Ceci est illustré dans la Figure II.22 qui représente la première cartographie recalée, la deuxième cartographie originale et leur mise en correspondance.

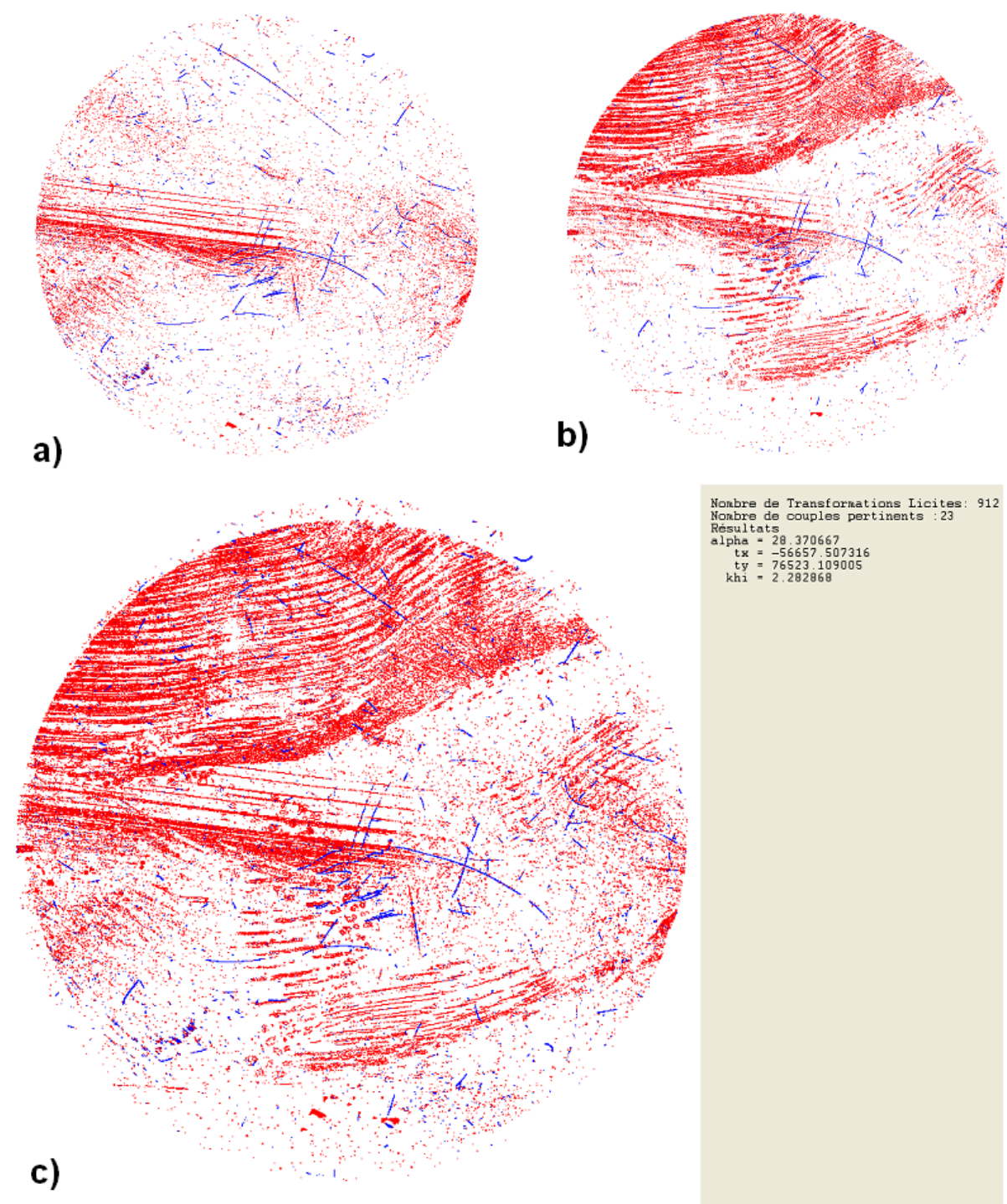


Figure II.22 : a) Cartographie de la Figure II.21 recalée par rapport à la cartographie de la Figure II.20; (b) Cartographie de la Figure II.20; (c) Mise en correspondance des cartographies représentées en a) et b).

II.3 Conclusion du chapitre

Le choix des capteurs utilisés par le système est primordial car il influence directement la nature et la qualité des informations qui seront traitées pour détecter et caractériser les défauts d'aspect.

Afin de récolter les renseignements nécessaires à un tel choix, un travail conséquent de recherche bibliographique a été entrepris. Les fruits de ce travail, basé sur la littérature d'optique instrumentale, les brevets industriels et les contacts noués avec des entreprises spécialisées, sont présentés en Annexe C. L'analyse de ces informations a permis d'arriver à trois conclusions :

- Il est nécessaire d'utiliser *deux capteurs distincts* :
 - un capteur de *détection grossière*
 - un capteur de *détection fine et de caractérisation*.
- Les meilleurs techniques de détection des défauts en termes de rapport coût / performance sont :
 - la strioscopie
 - l'éclairage par les tranches.
- Toujours en termes de compromis coût / performance, la meilleure technique de détection fine et de caractérisation des défauts est la microscopie Nomarski.

On s'est ensuite intéressé plus particulièrement à cette technique de microscopie Nomarski dont le principe et la mise en œuvre technologique ont été exposés ainsi que ses avantages sur les autres microscopies.

La littérature ne fournit pas les outils nécessaires à l'exploitation des images Nomarski pour le contrôle de pièces optiques. Partant de ce constat, nous avons proposé des méthodes permettant, à partir de ces données :

- de détecter tous les défauts d'aspect significativement plus profonds que la taille caractéristique de la rugosité,
- de mesurer simplement la rugosité de la surface considérée, qui est un paramètre d'une grande importance dans l'industrie optique.

D'autre part, la robustesse et la répétabilité de ces méthodes ont été étudiées, afin de garantir les propriétés métrologiques d'un système qui les utiliserait.

Dans le but de valider les techniques proposées dans les conditions de production, un prototype les mettant en œuvre a été mis au point. Ce prototype permet l'étude de pièces optiques de petite taille en termes de défauts d'aspect et de rugosité et fournit un environnement d'exploitation des mesures par les opérateurs. Notre approche a par ailleurs été validée par l'expérience puisque de nombreuses pièces différentes, fabriquées au sein de REOSC, ont été contrôlées avec succès grâce à ce prototype.

De plus, l'implantation performante des méthodes de traitement des images Nomarski autorise leur exécution en temps « masqué » : l'image d'un champ est traitée pendant que la platine se déplace vers la prochaine position. Ceci permet, associé avec la stratégie multi-résolution proposée, d'envisager le respect des objectifs concernant le temps de contrôle fixés par les contraintes de production.

Cependant, si tous les défauts de la pièce contrôlée sont détectés d'une manière robuste, leurs natures restent inconnues. Ainsi des artefacts (poussières, traces de nettoyage) présents sur la pièce sont également détectés et donc considérés comme des défauts d'aspect de façon erronée. Au cours des contrôles évoqués plus haut, ce phénomène a été limité par la maîtrise de l'environnement : le prototype était installé en salle propre et les pièces soigneusement nettoyées avant leur examen. Dans les faits, le système à concevoir doit fonctionner en atelier, sans préparation spéciale des pièces. On doit par conséquent lui donner la capacité de distinguer entre « vrais » défauts et artefacts. Dans la suite de ce travail, on étudiera cette phase de caractérisation qui met en œuvre la véritable intelligence du système. Toutes les données traitées dans le chapitre suivant proviennent de la microscopie Nomarski.

Extraction de données et classification des défauts

Sommaire

| | |
|---|------------|
| III.1 Représentation des défauts pour le système de classification | 80 |
| III.1.1 Paramétrisation des données : état de l'art | 81 |
| III.1.1.1 Descripteurs de contour | 82 |
| III.1.1.2 Descripteurs d'image | 82 |
| III.1.1.3 Descripteurs relevant d'une heuristique | 84 |
| III.1.2 Implantation et validation du codage par les invariants de Fourier Mellin | 84 |
| III.1.2.1 Implantation | 85 |
| III.1.2.2 Validation | 90 |
| III.1.2.2.1 Présentation rapide des Self Organizing Map (SOM) | 90 |
| III.1.2.2.2 Protocole expérimental | 91 |
| III.1.2.2.3 Résultats et discussion | 93 |
| III.2 Réduction de dimension et classification neuronale | 102 |
| III.2.1 Dimension intrinsèque d'une distribution | 102 |
| III.2.2 Le « fléau de la dimension » | 106 |
| III.2.3 Les techniques de réduction de dimension | 111 |
| III.2.3.1 Techniques par sélection d'un sous-ensemble de paramètres | 111 |

| | | |
|--------------|--|------------|
| III.2.3.2 | Techniques projectives | 113 |
| III.2.3.2.1 | Techniques Linéaires | 114 |
| III.2.3.2.2 | Techniques locales | 116 |
| III.2.3.2.3 | Techniques globales | 117 |
| III.2.4 | Réduction de dimension et classification neuronale des défauts | 121 |
| III.2.4.1 | Etude des techniques SOM, CCA et CDA | 122 |
| III.2.4.2 | Classification neuronale : discrimination des défauts de type « poussière » | 129 |
| III.2.4.2.1 | Prologue : dimension intrinsèque des données | 130 |
| III.2.4.2.2 | Protocole expérimental | 131 |
| III.2.4.2.3 | Résultats et discussion | 135 |
| III.3 | Classification neuronale modulaire | 141 |
| III.3.1 | Classification modulaire : définition et rapide état de l'art . . . | 141 |
| III.3.2 | La technique Treelike Divide To Simplify (T-DTS) | 143 |
| III.3.3 | Technique de décomposition proposée | 146 |
| III.3.3.1 | Fondement principal | 146 |
| III.3.3.2 | Estimation locale de la dimension intrinsèque | 149 |
| III.3.3.3 | Formation de groupes de dimensions intrinsèques ho- mogènes | 151 |
| III.3.4 | Validation de la méthode sur un problème « jouet » | 153 |
| III.3.4.1 | Présentation du problème | 153 |
| III.3.4.2 | Protocole expérimental | 156 |
| III.3.4.3 | Résultats et discussion | 156 |
| III.3.5 | Classification neuronale modulaire des défauts d'aspect | 158 |
| III.4 | Conclusions du chapitre | 161 |

Nous nous intéressons dans ce chapitre à la partie « intelligente » du système : le mécanisme qui va permettre de déterminer la nature d'un défaut détecté. Une telle étape de classification est essentielle pour garantir la fiabilité du diagnostic puisqu'elle permet de différencier les défauts réels des différents artefacts dus à l'environnement et à la préparation de la pièce.

Pour remplir cette tâche de catégorisation complexe, nous mettons en œuvre des techniques connexionnistes. Ces méthodes jouissent en effet de caractéristiques attrayantes pour ce type d'applications (Zhang 2000)(Egmont-Petersen, de Ridder, and Handels 2002). Leur capacité à reproduire l'expertise humaine nécessaire à la discrimination des défauts, sans utiliser un modèle explicite de celle-ci, a particulièrement motivé leur utilisation.

Les données présentées en entrée du système de classification sont des imagerie de défauts (introduites dans le paragraphe II.2.2.1, p.48) issues de la phase de détection développée dans le Chapitre II. Pour différentes raisons, ces imagerie ne peuvent être décrites directement sous une forme vectorielle, représentation pourtant privilégiée par les réseaux de neurones artificiels. Une phase de codage de l'information est donc nécessaire. La première partie de ce chapitre traite de cette problématique et présente la solution retenue.

La représentation obtenue est d'une taille assez importante. Ceci peut entraîner des problèmes, liés au « fléau de la dimension », que des techniques de réduction de dimensionnalité peuvent limiter. La deuxième partie de ce chapitre présente ces techniques et compare trois d'entre elles. En particulier, l'impact de leur utilisation sur les performances d'un classifieur neuronal de type Multi Layer Perceptron (MLP) est étudié. Les expériences menées impliquent des données réelles (des imagerie de défauts) issues de mesures de composants optiques et visent à discriminer les défauts de type « poussière ».

Afin d'améliorer les performances obtenues, nous avons mis en œuvre un traitement neuronal modulaire, exposé dans la troisième partie de ce chapitre. Dans ce cadre, nous proposons un nouveau mécanisme de décomposition basé sur les dimensions intrinsèques locales et la topologie des données. Le concept est validé sur un problème jouet puis sur les données réelles issues de notre problème industriel.

III.1 Représentation des défauts pour le système de classification

Les données d'entrée du système de classification sont des images en niveau de gris avec un nombre important de pixels. Il n'est pas souhaitable et presque impossible de manipuler directement de tels objets et ce pour trois raisons :

- La taille importante des vecteurs représentant les défauts (constitués des niveaux de gris de leurs pixels) conduit à des calculs lourds et peu efficaces. En effet, la complexité calculatoire des classifieurs usuels augmente au mieux linéairement avec la taille des vecteurs d'entrée. D'autre part, la multiplication des paramètres d'entrée pénalise lourdement l'efficacité des classifieurs, à cause du « fléau de la dimension » (paragraphe III.2.2).
- La taille des images n'étant pas constante (typiquement, elle peut varier de 3 ordres de grandeur), le nombre de pixels les décrivant ne l'est pas non plus. Ainsi, une description brute n'a pas une taille invariante, ce qui pose de gros problèmes (comment gérer des vecteurs de tailles différentes avec un même système ?). Une solution consiste à ramener toutes les images à une taille fixe. Elle a le gros désavantage de décrire les défauts de grandes tailles avec une très faible résolution, les rendant difficilement identifiables.
- L'information portée par les niveaux de gris des pixels est très redondante (par exemple la corrélation entre niveaux de gris de pixels voisins est importante). Ceci est peu efficace et de plus une corrélation importante entre composantes du vecteur peut « cacher » une partie intéressante du signal.

Il est donc nécessaire de représenter ces données d'une manière plus compacte. Cependant, il est impératif de conserver leur représentativité afin de maintenir l'intégrité de l'information qu'elles véhiculent et ainsi, la séparabilité des différentes classes (si l'on se place dans l'optique d'une classification).

III.1.1 Paramétrisation des données : état de l'art

Afin d'assurer une forte puissance descriptive des signaux projetés dans un espace de représentation de dimension beaucoup plus petite, il convient de choisir une transformation possédant un certain nombre de qualités :

- *Orthogonalité et complétude des composantes* : les composantes du vecteur descriptif obtenu doivent être autant que faire se peut orthogonales et complètes. Ceci assure d'une part l'efficacité du codage d'un point de vue informatif, et d'autre part l'unicité de la représentation (ce qui autorise par la même une transformation inverse dont il peut être intéressant de disposer).
- *Robustesse aux bruits* : la description obtenue doit être résistante aux différents bruits pouvant entacher les données brutes. Cette propriété fondamentale garantit que seule l'information pertinente sera codée. Dans notre cas, les imagerie sont soumises à différents bruits : non homogénéités d'éclairage, bruit de caméra et différence de rugosité du fond.
- *Invariance à certaines transformations* : les paramètres descriptifs doivent présenter des propriétés d'invariance par rapport à certaines transformations (celles qui laisse le problème considéré invariant). Par exemple, dans notre cas, la représentation d'un défaut doit être insensible aux rotations et aux translations (une poussière, même après un tel déplacement reste une poussière!). On souhaite également qu'elle soit invariante avec les homothéties car ceci nous permettrait de travailler indifféremment avec différents grossissements. En bref, des imagerie de défauts qui ne diffèrent que par ces transformations doivent être représentées par le même jeu de paramètres, ce afin d'éviter la multiplication des problèmes à traiter (par exemple le problème de discrimination des poussières peut être vu comme un ensemble de problèmes de discrimination de poussières orientées à 0° , 15° , $30^\circ \dots$).

Enfin, on peut rajouter à cette liste le fait que la représentation finale des défauts optimise la compacité et la séparabilité des classes recherchées. Ceci nécessite néanmoins la connaissance de la classe de chaque défaut et n'est pas facile à mettre en œuvre.

III.1.1.1 Descripteurs de contour

Une première façon de caractériser les images de défauts est d'en considérer uniquement les contours. Plusieurs descripteurs de contours, ayant des propriétés différentes, sont utilisés de façon usuelle ; une bonne revue de ces descripteurs peut-être consultée dans (Chettaoui et al. 2005). De plus, il est également possible d'utiliser les descripteurs d'images décrits plus loin (l'image considérée est alors réduite aux contours que l'on veut décrire).

Il existe dans la littérature plusieurs exemples où ce type de codage est couplé avec des classifieurs (souvent neuronaux) : (Chettaoui et al. 2005), (Rajpal, Chaudhury, and Banerjee 1999), (Veropoulos et al. 1998), (Veropoulos, Campbell, and Learmonth 1998). Dans notre cas, il apparaît cependant clair que l'étude des seuls contours ne peut pas permettre une classification fiable des défauts, car ils ne véhiculent pas assez d'information sur leurs natures. Par exemple, il est impossible par ce biais de discriminer une poussière d'un défaut de type piqûre de forme identique.

III.1.1.2 Descripteurs d'image

Une autre solution est de coder l'image dans son ensemble afin de conserver le plus d'information possible. Les descripteurs d'image utilisés dans la littérature sont beaucoup plus nombreux que les descripteurs de contour ; ils sont recensés dans (Choksuriwong, Laurent, and Emile 2005), (Derrode 1999) et (Adam et al. 2001) :

- Moments invariants de Hu (Hu 1962).

Ces moments, obtenus par combinaisons des moments réguliers de l'image sont invariants par rapport aux translations, rotations et homothéties. Ils sont très utilisés dans la littérature pour la description de formes en vue d'une classification ou d'une indexation, mais sont assez sensibles aux bruits. Par ailleurs cette famille de descripteurs n'est ni orthogonale, ni complète.

- Moments invariants de Zernike (Chong, Raveendran, and Mukundan 2003).

Les moments de Zernike sont construits en utilisant des polynômes complexes qui forment un ensemble orthogonal complet défini sur le disque unité. Ils sont invariants

par rotation et changements d'échelles et présentent des propriétés intéressantes en termes de résistance aux bruits, efficacité informative et possibilité de reconstruction.

- Descripteurs de texture usuels (Gonzalez and Woods 2002), ou basés sur l'analyse multirésolution (Deguillemont, Lecoeuche, and Dubus 1998).

Ces descripteurs sont variés et donc présentent des propriétés différentes. Cependant l'information de texture seule paraît très limitée pour la classification envisagée et le calcul de ces paramètres pour des images de taille différentes peut poser des problèmes. Il pourrait cependant être intéressant d'ajouter quelques uns de ces descripteurs à un vecteur obtenu par d'autres moyens.

- Descripteurs issus des ondelettes (Ducottet 1994) et des filtres de Gabor.

Ces descripteurs ont un excellent pouvoir de représentation (efficacité informationnel, représentation temps-fréquence, robustesse au bruit), bénéficient d'une justification biologique (certains neurones du cortex visuel des mammifères ont le comportement de ce type de filtres) et se prêtent bien à une implémentation efficace (Kutil and Uhl 2001). On citera ici (Cohen, Raz, and Malah 1995a) et (Cohen, Raz, and Malah 1995b) où est présentée une technique intéressante permettant de trouver une base d'ondelettes adaptée à un phénomène donné, tout en gardant des propriétés d'orthogonalité et d'invariance par translation. Ces familles de descripteurs sont complètes et invariantes en translation et changement d'échelle. Malheureusement, dans le cas général, ces descripteurs ne sont pas invariants en rotation (sauf les ondelettes non séparables), et les fonctions de Gabor ne sont pas orthogonales. Par ailleurs le choix de l'ondelette-mère parmi les nombreuses options possibles peut être délicat.

- Descripteurs issus de la transformée de Fourier-Mellin (Ghorbel 1994).

Les problèmes d'interprétation d'image étant répandus, il existe beaucoup d'exemples où ces différentes techniques de codage sont utilisées conjointement à des classifieurs. Dans (Gleason et al. 1998), (Nath et al. 1994) et (Recce et al. 1996) des objets à classer sont représentés par le biais de moments, dans (Sheng and Lejeune 1991) par des invariants

issus de la transformée de Fourier-Mellin, dans (Siegel and Gunatilake 1998) et (Kalman and Kwasny 1998) par une décomposition en ondelettes ; dans (Garris, Wilkinson, and Wilson 1991), (Lampinen and Oja 1995), (Lampinen 1992) et (Lampinen, Smolander, and Korhonen 1995) est utilisé un banc de filtres de Gabor ; enfin, dans (Teoh Beng Jin, Ngo Chek Ling, and Thian Song 2004) est décrite une application de reconnaissance d'empreintes utilisant les invariants de Fourier-Mellin.

III.1.1.3 Descripteurs relevant d'une heuristique

Les signaux bruts peuvent être également décrits par un certain nombre de paramètres choisis par un expert suivant une heuristique. Ces paramètres (par exemple ratio d'aspect, longueur, largeur du défaut, descripteurs issus de considérations plus « pointues » ...) sont extraits facilement de l'image brute. L'utilisation de tels paramètres est à double tranchant : d'une part ils peuvent avoir une bonne qualité de représentation du problème de part leur caractère expérimental, d'autre part ils souffrent d'un manque d'exhaustivité et d'objectivité (Oukhellou 1997). Ainsi, dépendant entièrement de la qualité de l'expertise, ils peuvent ne pas balayer l'ensemble de l'espace des paramètres et ainsi engendrer une faible capacité de généralisation. Il peut être cependant judicieux d'ajouter de tels paramètres à un vecteur descriptif obtenu par d'autres moyens.

III.1.2 Implantation et validation du codage par les invariants de Fourier Mellin

Plusieurs raisons nous ont amené à choisir une représentation des défauts par des invariants issus de la transformée de Fourier-Mellin. Premièrement, ces invariants possèdent les caractéristiques basiques requises pour un codage de qualité : ils sont complets et invariants par rapport à la translation, la rotation et l'homothétie. Ensuite, ils paraissent particulièrement adaptés à notre problématique dans la mesure où le type d'images considéré a un fort contenu informationnel fréquentiel. Ils bénéficient également d'une justification neurophysiologique : des études (Gallant, Braun, and Essen 1993) ont en effet montré que les cellules du cortex visuel des macaques répondaient préférentiellement à des stimuli polaires (rappelons ici que la transformée de Fourier Mellin est la transformée de Fourier dans le domaine log-polaire).

Mais ce qui est sans doute le plus marquant est leur supériorité par rapport aux autres descripteurs, en particulier les moments de Hu et de Zernike, dans des tâches de reconnaissance de lettres (Adam et al. 2001)(Grace and Spann 1991), d'indexation d'images (Derrode, Mezhoud, and Ghorbel 2000) ou de classification d'objets (Choksuriwong, Laurent, and Emile 2005). Cette supériorité trouve par ailleurs une justification théorique dans le fait que les moments de Hu et de Zernike peuvent être exprimés comme une combinaison de moments de Fourier-Mellin d'ordres radiaux élevés (Ping et al. 2000)(Sheng and Lejeune 1991). Ainsi l'utilisation directe des descripteurs de Fourier-Mellin permet de considérer des moments d'ordres limités, ce qui se traduit par une résistance accrue au bruit et une meilleure transcription de l'information (les moments radiaux d'ordres élevés ont en effet tendance à favoriser exagérément la contribution des bords de l'image par rapport aux zones proches de son centroïde).

III.1.2.1 Implantation

La transformée de Fourier-Mellin (TFM) :

La transformée de Fourier-Mellin d'une fonction $f(r; \theta)$, définie en coordonnées polaires, est donnée par la relation suivante (Derrode 1999) :

$$M_f(q; p) = \frac{1}{2\pi} \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} r^{ip} \exp(-iq\theta) f(r; \theta) \frac{dr}{r} d\theta \quad (\text{III.1})$$

avec $q \in \mathbb{Z}$ et $p \in \mathbb{R}$.

On peut tout d'abord remarquer que la TFM est la combinaison de la représentation par coefficients de Fourier d'une fonction 2π -périodique (pour la variable angulaire θ) et de la transformée de Mellin (pour la variable radiale r). Ainsi, cette transformation est séparable en deux parties, l'une orthogonale (Fourier) et l'autre qui ne l'est pas. Il est cependant possible d'orthogonaliser les descripteurs obtenus (Sheng and Shen 1994).

De plus, cette transformation fournit une représentation unique de la fonction f , ce qui sous-tend l'existence d'une transformée inverse.

Cette définition pose néanmoins deux problèmes :

- Elle suppose l'existence d'un centre des coordonnées. Ce point doit par ailleurs être invariant en translation, pour assurer que la représentation issue de la TFM le soit aussi.
- Elle utilise une intégrale qui peut ne pas être définie (!). En particulier, si la fonction f décrit une image en niveau de gris, cette intégrale diverge et la TFM ne peut être calculée.

Le premier de ces problèmes est réglé par la définition d'un point fixe de l'image comme centre des coordonnées. Un choix raisonnable consiste à utiliser le centre de gravité de l'image. Le deuxième problème est résolu par la définition d'un prolongement analytique de la TFM.

Le transformée de Fourier-Mellin analytique (TFMA) :

Cette solution proposée dans (Ghorbel 1994) consiste à remplacer $f(r; \theta)$ par $r^\sigma f(r; \theta)$ avec $\sigma > 0$. Cette transformation réversible assure l'existence de l'intégrale et on peut alors définir la transformée de Fourier-Mellin analytique par :

$$M_{f_\sigma}(q; s) = \int_{r=0}^{\infty} \int_{\theta=0}^{2\pi} r^{s-1} \exp(-iq\theta) f(r; \theta) dr d\theta \quad (\text{III.2})$$

avec $q \in \mathbb{Z}$, $s = \sigma + ip \in \mathbb{C}$, $\sigma > 0$.

En pratique le paramètre de convergence σ est choisi dans l'intervalle $[0, 5; 1, 5]$; il module l'influence respective des pixels situés au centre des coordonnées par rapport aux pixels situés à la périphérie de l'objet.

D'autre part, la fonction $f(r; \theta)$ peut être reconstruite à partir de la TFMA en utilisant la transformation inverse suivante :

$$f(r; \theta) = \int_{-\infty}^{\infty} \sum_{q \in \mathbb{Z}} r^{-\sigma-ip} \exp(-iq\theta) M_{f_\sigma}(q; s) dp \quad (\text{III.3})$$

avec $s = \sigma + ip \in \mathbb{C}$.

Enfin, dans le cas d'images, comme le support de f est borné, la transformée de Fourier-Mellin s'étend à l'infini. Cependant, dans la pratique, nous ne travaillons qu'avec un nombre fini de coefficients donc une partie de l'information originale sur l'objet est perdue. Néanmoins, l'amplitude de la TFM décroît lorsqu'on s'éloigne du centre du repère

et s'annule à l'infini. En prenant en compte un spectre plus important, l'information perdue par troncature du spectre peut donc être aussi faible que souhaité. Afin d'illustrer ce fait et le type d'information codée par la TFM, on reprend, dans la Figure III.1, un exemple étudié dans (Derrode 1999). Cet exemple met en jeu une image en niveaux de gris représentant un papillon. Dans un premier temps, l'image est décomposée grâce à la TFMA et dans un deuxième temps, elle est reconstruite en utilisant la transformation inverse. La Figure III.1 représente les différentes reconstructions obtenues en utilisant des tailles de spectre de plus en plus importantes. De gauche à droite, P , qui module la taille du spectre de Mellin ($p \in [-P; P]$) augmente; de haut en bas, Q , qui module quant à lui la taille du spectre de Fourier ($q \in [-Q; Q]$) augmente. On constate que, au fur et à mesure que les dimensions des spectres augmentent, les contours et la texture des ailes du papillon se précisent pour aboutir finalement à une image reconstruite de bonne qualité.

Les invariants issus de la TFMA :

Dans (Ghorbel 1994), l'auteur construit par normalisation un ensemble de primitives invariante aux changements d'échelle et d'orientation données par la relation suivante :

$$I_f(q; s) = M_f(q; s) [M_f(0; \sigma)]^{\frac{-s}{\sigma}} [M_f(1; \sigma)]^{-q} |M_f(1; \sigma)|^q \quad (\text{III.4})$$

avec $q \in \mathbb{Z}$, $s = \sigma + ip \in \mathbb{C}$.

Il est important de noter que $I_f(0; \sigma) = 1$, que $I_f(1; \sigma)$ est un réel pur, et que l'ensemble des invariants vérifie la propriété de symétrie suivante :

$$I_f^*(q; s) = I_f(-q; s^*)$$

On peut donc limiter le nombre de primitives non redondantes à l'ensemble correspondant à :

$$\{(q, p)/(q = 0; 0 \leq p \leq P) \cup (1 \leq q \leq Q; -P \leq p \leq P)\}$$

Ces invariants jouissent de la propriété de *complétude* : on peut calculer à partir de leurs valeurs les coefficients de la TFM et donc reconstruire $f(r; \theta)$ en utilisant la TFM inverse. Il n'y a donc pas d'ambiguïté de représentation : un jeu de paramètres est associé de manière unique à une image (aux transformations laissant ces paramètres invariants près, bien entendu).

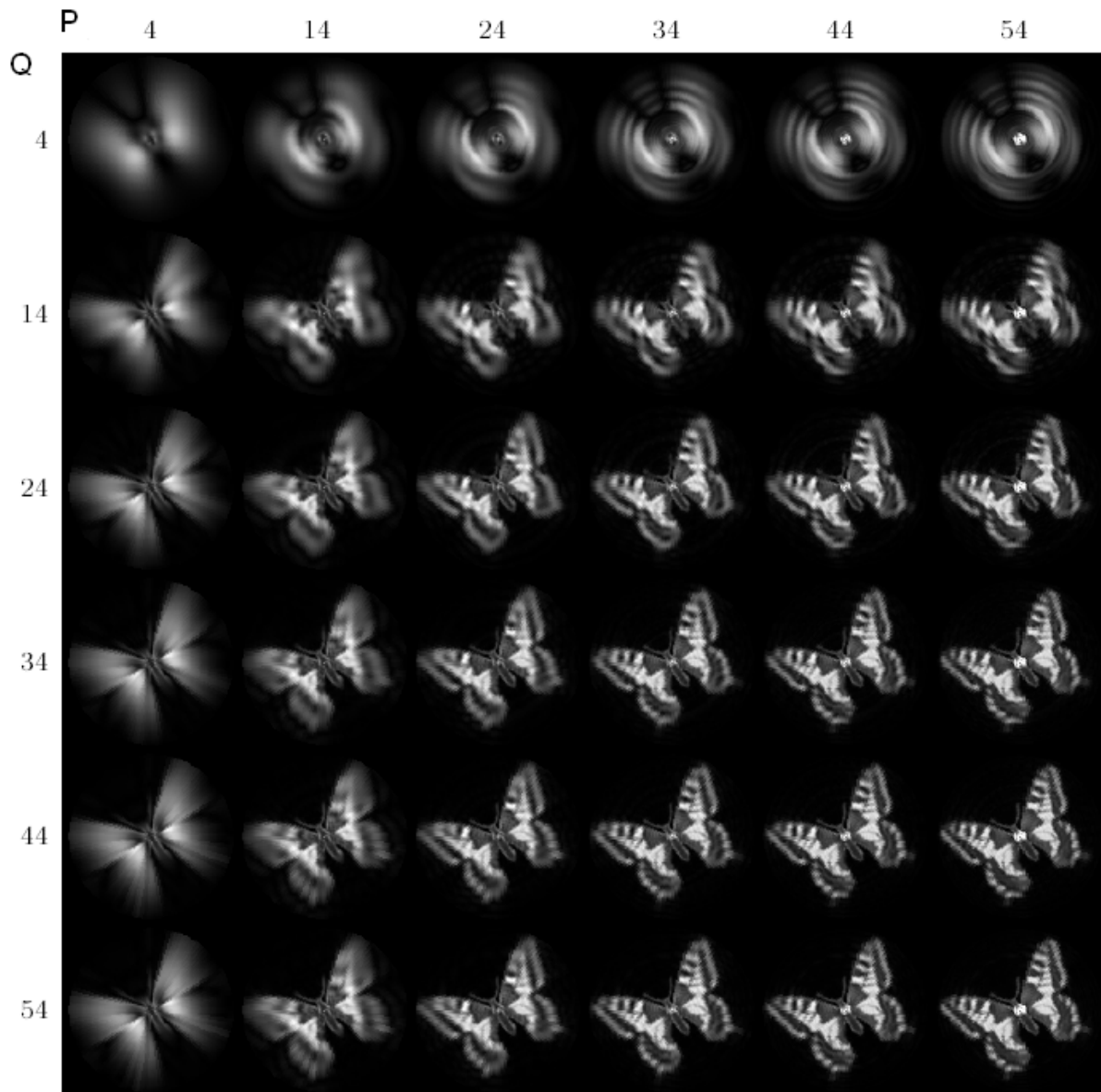


Figure III.1 : Reconstruction d'une image de papillon à partir de sa TFM pour des tailles variables de spectre.

Enfin, on peut montrer que la famille des invariants est sommable (Ghorbel 1998), ce qui implique l'existence d'une métrique dans l'espace des formes décrites : cette distance mesure le degré de similarité des objets indépendamment de leur orientation et de leur taille.

Dans (Derrode and Ghorbel 2001), sont proposées trois méthodes d'approximation de la TFMA d'images en niveaux de gris : une approximation par échantillonnage polaire, une approximation par échantillonnage log-polaire et une approximation cartésienne sans rééchantillonnage. Dans (Grace and Spann 1991), le calcul des coefficients de la TFM utilise la FFT.

Dans ce travail, nous avons, pour notre part, choisi d'utiliser une technique qui permet d'évaluer les différents coefficients de Fourier-Mellin en n'utilisant qu'une simple convolution. Cette technique permet donc de travailler directement en utilisant les coordonnées cartésiennes discrètes de l'image et permet une implémentation efficace du calcul des invariants. L'approximation recherchée s'obtient en utilisant un banc de filtres proposés dans (Ravichandran and Trivedi 1995) :

$$M_f(q; \sigma + ip) \approx \sum_{\substack{k \\ 1 \leq (k^2 + l^2) \leq r_{\max}^2}} \sum_l h_{p,q}(k, l) f(k_0 - k, l_0 - l) \quad (\text{III.5})$$

où $f(i; j)$ est le niveau de gris du pixel de coordonnées $(i; j)$,

$(k_0; l_0)$ sont les coordonnées du centre de gravité de l'image

r_{\max} est le rayon maximal de l'image et où :

$$h_{p,q}(k, l) = \frac{\exp\left(i \left[\frac{p}{2} \ln(k^2 + l^2) - q \cdot \arctan\left(\frac{l}{k}\right) \right]\right)}{(k^2 + l^2)^{1 - \frac{\sigma}{2}}}$$

Une fois les $M_f(q; \sigma + ip)$ calculés, il suffit d'appliquer la relation III.4 pour obtenir les invariants recherchés.

Il est à noter que l'utilisation de cette technique amène à utiliser des bancs de filtres qui ne sont pas exactement orthogonaux. Par conséquent les invariants calculés par ce biais ne le sont pas non plus. Ceci n'est néanmoins pas spécifique à cette technique, la discrétisation inhérente aux images manipulées rend en effet l'orthogonalité stricte impossible.

III.1.2.2 Validation

Afin d'évaluer la pertinence de l'utilisation des invariants de Fourier-Mellin pour coder les imagerie de défauts, différentes expériences de « clustering » à partir de données réelles ont été menées. Une opération de clustering consiste à mettre en évidence des groupes de données (ici des imagerie de défauts) présentant des caractéristiques communes. La philosophie de ces expériences est alors la suivante : si l'on est capable, en utilisant le codage proposé, de former de manière robuste des clusters de données pertinents du point de vue du problème à traiter, on aura démontré que ce codage est satisfaisant.

Pour ces expériences, nous avons utilisé la technique des Self Organizing Map (présentée dans le paragraphe suivant), et de nombreuses imagerie de défauts provenant de pièces optiques réelles (Voiry et al. 2006a).

Outre l'étude de la pertinence du codage employé, ces expériences ont permis d'obtenir des informations précieuses sur la complexité de l'espace des données et sur la capacité des SOMs à décrire cet espace (Voiry et al. 2006b).

III.1.2.2.1 Présentation rapide des Self Organizing Map (SOM)

Le but n'est pas ici de décrire l'algorithme mettant en œuvre cette technique bien connue (Kohonen 2001) mais plutôt d'en rappeler succinctement le principe et les caractéristiques importantes.

Cette méthode a pour objectif de projeter un espace multidimensionnel dans une représentation de faible dimension, de façon non supervisée. Typiquement, une SOM consiste en une grille de neurones de dimension 2 (plus rarement 1 ou 3) qui constitue cette représentation. Les neurones entretiennent une relation de voisinage en tant que nœud de la grille et un vecteur « prototype » est associé avec chacun d'entre eux. Durant la phase d'apprentissage, ces vecteurs sont ajustés afin de représenter au mieux les données : cela correspond à une quantification de l'espace des données. De plus, des données similaires (au sens d'une mesure prédéfinie) sont projetées vers le même neurone (ou vers des neurones très proches dans la grille), alors que des données différentes sont projetées vers des neurones éloignés les uns des autres : cela correspond à un partitionnement des données (clustering). Enfin, en considérant pour un vecteur donné en entrée, la position du neurone gagnant (c'est-à-dire dont le prototype est le plus proche du vecteur considéré) dans

la grille, on obtient une projection non-linéaire de l'espace multidimensionnel des données d'entrée vers l'espace bidimensionnel et discret de la grille de la SOM.

Par conséquent les SOMs sont des outils permettant de quantifier un espace de données et de projeter cet espace dans un espace de faible dimension, tout en conservant sa topologie. Ainsi, cette technique est couramment utilisée par l'industrie pour caractériser ou visualiser des données de hautes dimensions, mais également pour des tâches de classification (Kohonen et al. 1996) (Heikkonen and Lampinen 1999).

Les principaux paramètres à fixer pour mettre en œuvre cette méthode sont :

- La forme, les dimensions et le type de maillage de la grille de neurones
- La distance utilisée au sein de cette grille
- Le type de voisinages entre neurones
- Les paramètres numériques de l'algorithme d'apprentissage.

Ces paramètres influent naturellement sur la qualité de la carte construite.

III.1.2.2.2 Protocole expérimental

Afin d'acquérir un grand nombre d'images de défauts, trois expériences ont été menées. Elles ont mis en œuvre le prototype décrit dans le paragraphe II.2.4.1 (p.61), utilisant un grossissement x50. Le Tableau III.1 décrit les paramètres de ces trois expériences notées A, B et C.

Il est important de souligner que :

- Les données recueillies proviennent de deux composants optiques distincts, réalisés dans des matériaux différents. Ceci assure la présence d'une grande palette de type de défauts différents. D'autre part, les résultats obtenus pour ces deux pièces pourront raisonnablement être étendus à d'autres pièces.
- Les données des expériences A et B proviennent d'une pièce qui n'a pas été nettoyée, ce qui garantit la présence d'un certain nombre de « faux » défauts (poussières, traces de nettoyage...) à côté des défauts permanents. Dans le cas de l'expérience C, la pièce optique étudiée a été préalablement nettoyée minutieusement : les images représentent toutes (ou presque) des défauts permanents.

- Afin d'éviter l'apprentissage de « fausses » classes, les imagerie où sont représentées plusieurs défauts de nature différente ou un bord de champ microscopique ont été écartées.

| Expérience | Identifiant de la pièce optique | Nettoyage | Nombre de champs étudiés | Aire correspondante | Nombre d'imagerie recueillies |
|------------|---------------------------------|-----------|--------------------------|----------------------|-------------------------------|
| A | 1 | NON | 1178 | 28 cm ² | 3865 |
| B | 2 | NON | 605 | 14 cm ² | 1910 |
| C | 2 | OUI | 529 | 12,5 cm ² | 1544 |

Table III.1 : Paramètres des trois expériences pendant lesquelles ont été acquises les imagerie.

Les imagerie obtenues ont été codées grâce aux invariants de Fourier-Mellin avec $\sigma = 1$ et $(q, p) \in \{(q, p)/(q = 0; 0 \leq p \leq P) \cup (1 \leq q \leq Q; -P \leq p \leq P)\}$ où $P = 1$ et $Q = 2$ (Equation III.4). Ces paramètres conduisent à définir pour chaque imagerie, un vecteur descriptif à 13 composantes.

Ensuite, afin que chaque composante ait la même influence sur la construction du modèle, les vecteurs obtenus ont été normalisés en utilisant la transformation classique de centrage-réduction. Chaque composante F_i a été modifiée comme suit :

$$F_i = \frac{F_i - M}{\sigma} \quad (\text{III.6})$$

où M désigne la valeur moyenne de la composante F_i sur les trois bases de données et σ son écart-type.

En ce qui concerne les neurones des SOMs, ils sont arrangés selon une grille au maillage rectangulaire, et la distance entre deux d'entre eux est la distance Euclidienne. Quant aux dimensions de la grille (en termes de nombre de neurones), elles ont été fixées de manière empirique. Ce paramètre est en effet déterminant pour la qualité de la carte obtenue : la forme de la grille doit être adaptée à la forme de l'espace à décrire (Demartines 1994). Après de nombreux essais, les meilleurs résultats ont été obtenus en utilisant : une grille de

15x7 neurones pour l'expérience A, 20x8 pour l'expérience B, et 10x8 pour l'expérience C. Pour ces expérimentations, l'implantation des SOMs de la Neural Network Toolbox de Matlab®© a été mise en œuvre. Par ailleurs, les paramètres d'apprentissage proposés par défaut par le logiciel ont été utilisés.

III.1.2.2.3 Résultats et discussion

Les différentes cartes obtenues en utilisant les données des expériences A, B et C sont représentées dans les Figures III.2, III.3 et III.4. Dans ces figures, le défaut affiché pour chaque nœud de la grille est choisi aléatoirement parmi les exemples de la base de données projetés vers le neurone associé ; *la taille de ces images est normalisée donc l'échelle réelle n'est pas respectée.*

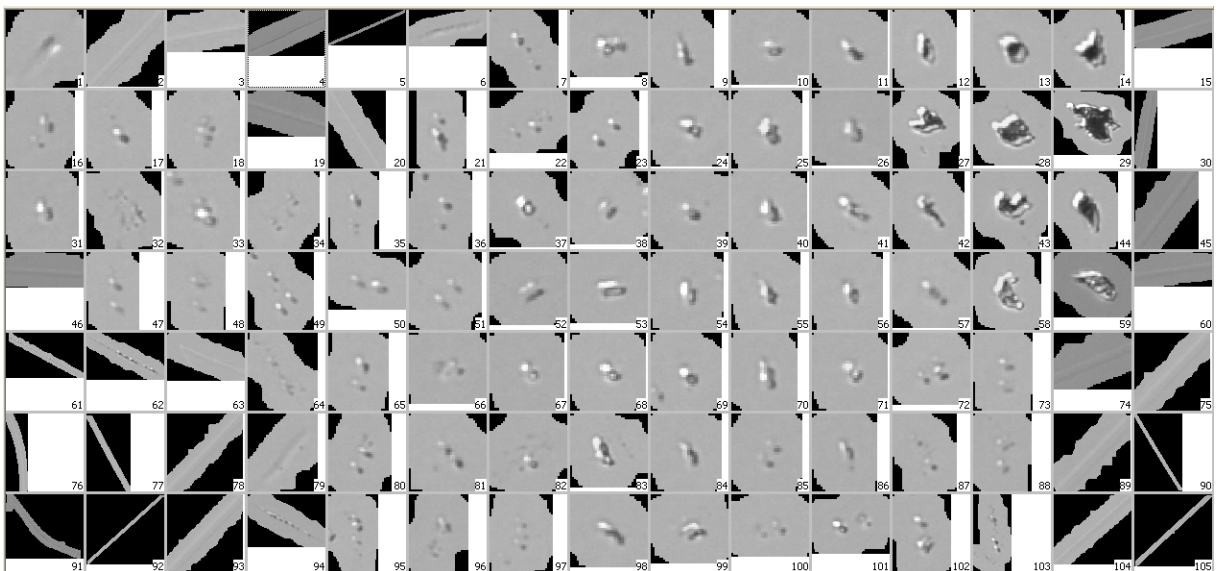


Figure III.2 : SOM obtenue avec les données de l'expérience A.

Existence de groupes de « similarité morphologique » :

Dans les trois cartes obtenues, on peut noter une forte ressemblance entre les images associées avec des nœuds adjacents et donc l'existence de groupes de neurones codant des défauts d'apparence similaire. Ceci traduit le fait que les invariants de Fourier-Mellin manipulés par les SOMs ont une bonne capacité à décrire les formes rencontrées dans les bases de données d'images. Les résultats étant par ailleurs satisfaisants pour les trois expériences, cette capacité ne dépend pas du type de pièce étudiée.

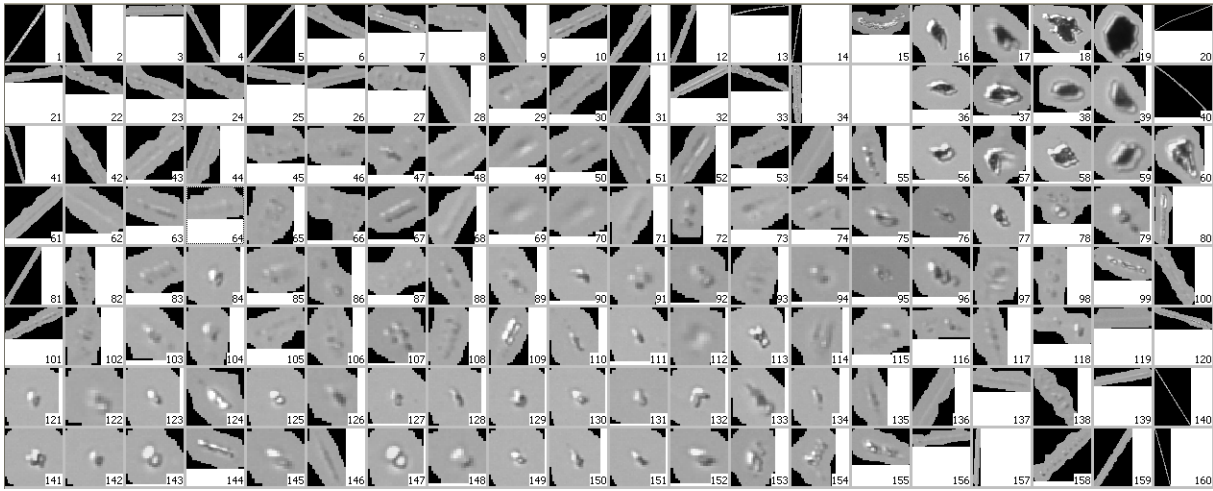


Figure III.3 : SOM obtenue avec les données de l'expérience B.

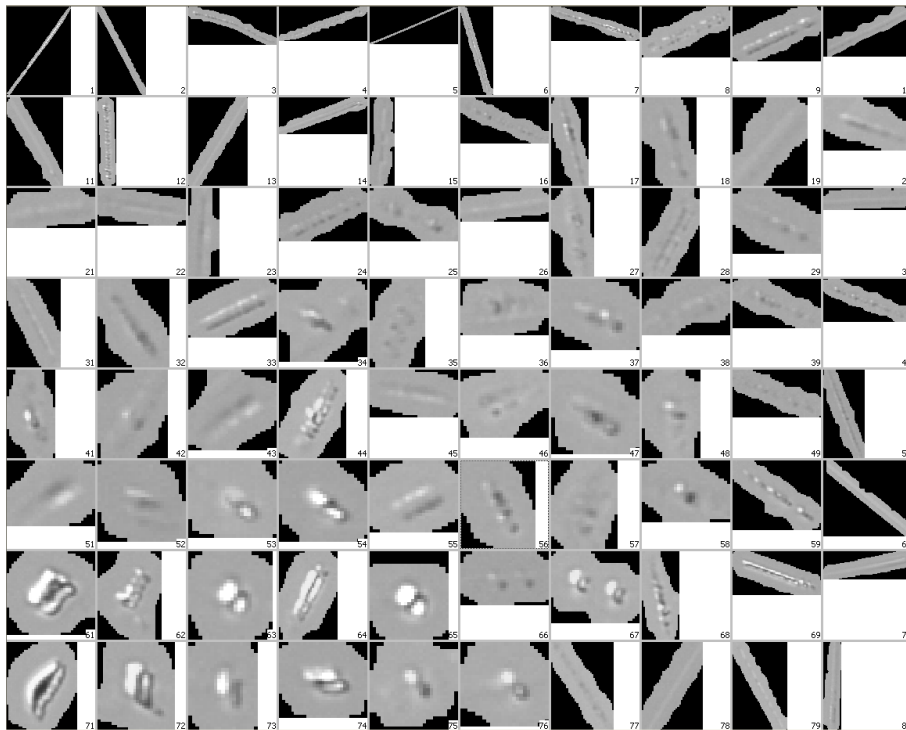


Figure III.4 : SOM obtenue avec les données de l'expérience C.

Homogénéité de la plupart des nœuds :

D'autre part, en étudiant l'ensemble des imageries projetées dans chaque nœud, il apparaît qu'un bon nombre de nœuds ont une composition « homogène » : dans ce cas, toutes les imageries projetées vers le nœud ont la même apparence et donc ont de bonnes chances de correspondre à des défauts de même nature. Cette « homogénéité » des nœuds est illustrée par les Figures III.5 et III.6.



Figure III.5 : Différentes imageries de défauts projetées dans le nœud n°28 de la carte représentée dans la Figure III.2.

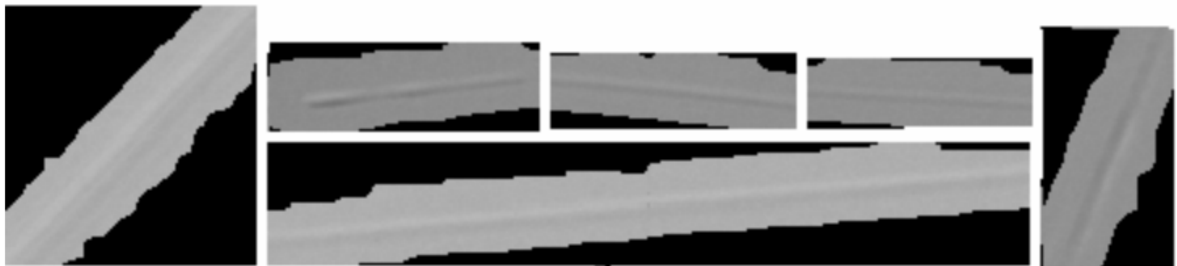


Figure III.6 : Différentes imageries de défauts projetées dans le nœud n°93 de la carte représentée dans la Figure III.2.

Sens physique des groupes de « similarité morphologique » :

D'un point de vue plus physique, une autre propriété des cartes obtenues est intéressante : les groupes de « similarité » qui émergent, correspondent à des groupes différents de défauts d'aspect. Ainsi, un expert a identifié la nature des défauts représentés dans la grille de la Figure III.2. La répartition des 4 labels ainsi définis, c'est-à-dire défaut de type « rayure », défaut de type « piqûre », artefact de type « poussière » et artefact de type « traces de nettoyage », est décrite dans la Figure III.7. On remarque alors que les nœuds associés avec des labels semblables forment des groupes structurés dans la grille. On peut également constater que le groupe identifié par l'expert comme étant constitué

par des artefacts de type « poussière », bien visible dans les cartes des Figure III.2 et III.3, a disparu dans la carte de la Figure III.4, correspondant à la pièce nettoyée. Il semble donc qu'il soit possible de faire un lien entre projection vers un neurone donné de la carte et nature physique du défaut (donc appartenance à une classe donnée).

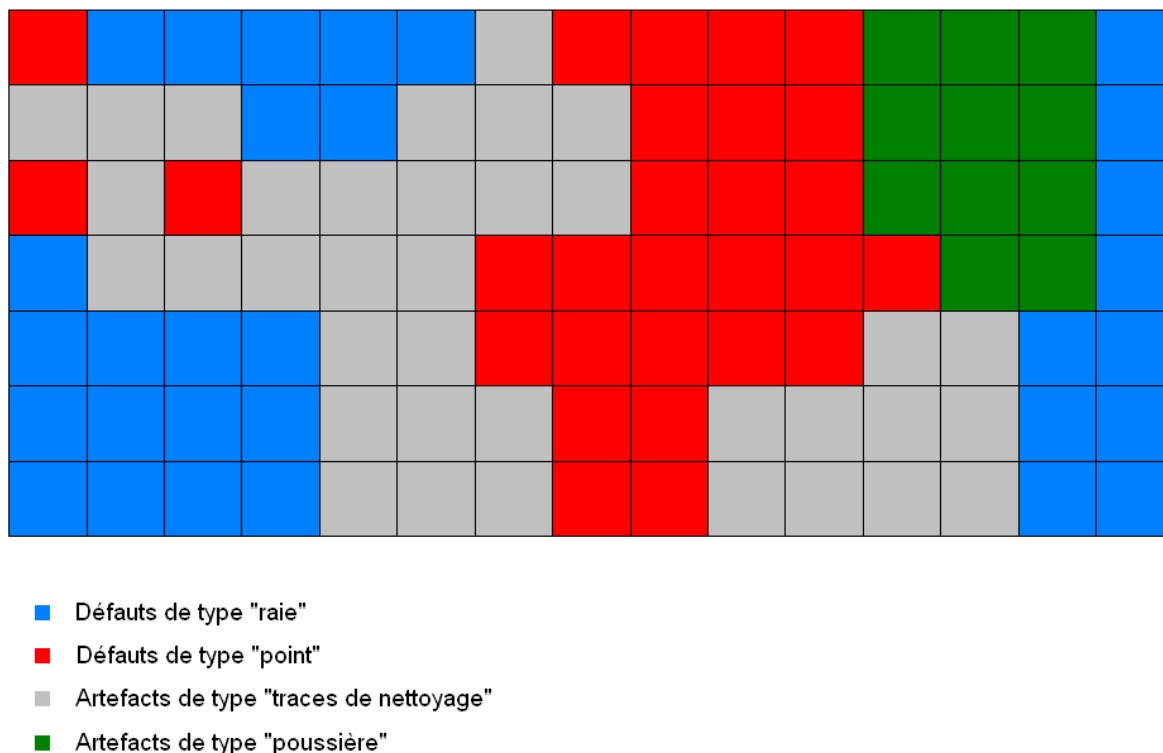


Figure III.7 : Répartition des différents labels définis par un expert dans la grille de la Figure III.2.

Possibilité d'une classification directe et problème d'hétérogénéité aux frontières :

On pourrait envisager d'utiliser ce lien pour obtenir une classification directe des défauts. Il suffirait pour cela de projeter l'imagette à classer dans la grille des Figures III.2 ou III.3, qui comportent toutes les deux des défauts de toutes natures. Le défaut considéré serait alors associé à la classe définie pour le nœud d'arrivée.

Cette technique utilisée dans (Silvén, Niskanen, and Kauppinen 2000), ne peut malheureusement pas être appliquée ici. On constate en effet un certain nombre de problèmes d'hétérogénéités au niveau des neurones situés aux frontières des différentes classes définies par l'expert. En effet, si au centre des régions correspondant à ces classes, les neurones ne reçoivent que des imagettes de défauts appartenant à une classe donnée, ce n'est pas le cas

pour les neurones situés aux frontières entre classes, où sont projetés des défauts appartenant aux différentes classes en présence. Une classification directe n'est par conséquent pas envisageable.

Problème d'hétérogénéité de certaines cellules de la SOM :

Plusieurs raisons, qui ne s'excluent pas mutuellement, peuvent expliquer ces hétérogénéités :

- Soit le codage choisi n'amène pas assez d'information pour que les exemples de classes différentes projetés dans un même neurone de la SOM soient discriminés.
- Soit la SOM n'est pas capable de maintenir cette information lors de la projection non-linéaire qu'elle effectue.
- Soit la SOM ne dispose pas d'assez de cellules pour pouvoir distinguer les deux classes lors de la phase de quantification.

On peut exclure cette troisième explication : les nombreuses expériences menées avec des tailles de grille très différentes ont toujours montré ce genre de problèmes aux frontières de classe. Afin d'apprécier les deux premières, nous avons étudié plus avant la projection effectuée par la SOM

Pour cela, la qualité de la projection réalisée par la SOM, de l'espace de données original (également appelé espace des poids) vers l'espace 2D de sa grille, a été évaluée grâce à la représentation $dy-dx$ proposée dans (Demartines 1992). On étudie, pour chaque paire de neurones, la distance dx entre ces deux neurones dans l'espace des données que l'on compare à la distance dy qui les sépare dans la grille. Pour chaque couple de neurones $(i; j)$, on trace un point de coordonnées $(dy(i, j); dx(i, j))$ où $dx(i, j) = \|\vec{x}_i - \vec{x}_j\|$ et $dy(i, j) = \|\vec{y}_i - \vec{y}_j\|$. \vec{x}_k (resp. \vec{y}_k) est le vecteur représentatif correspondant au k -ième neurone dans l'espace des poids (resp. dans l'espace de la grille). Cette construction est expliquée par le graphique de la Figure III.8, tiré de (Demartines 1994).

L'apparence du graphe obtenu dépend directement de la qualité de la projection réalisée par la SOM. En effet, si la topologie de l'espace des données est bien respectée, dx est très corrélé à dy , et l'on obtient des points arrangés le long d'une ligne droite. Au contraire, si l'organisation des neurones ne reflète pas cette topologie, on obtient un nuage de point diffus.

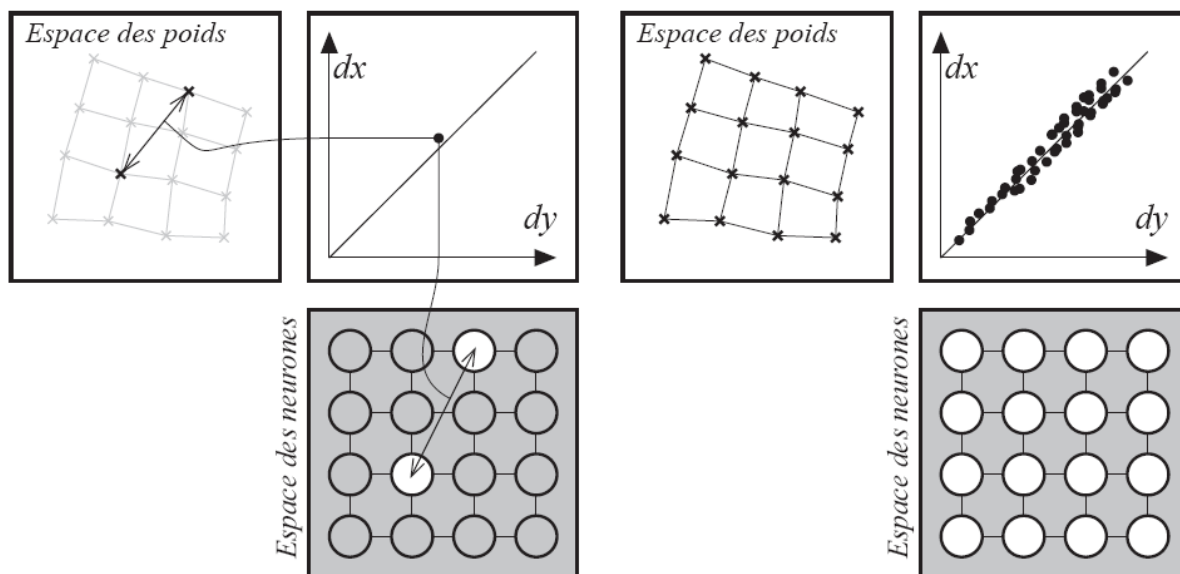


Figure III.8 : Construction de la représentation $dy-dx$: on trace la distribution jointe des distances entre les poids (axe vertical) et des distances de grille (axe horizontal).

Le graphe $dy-dx$ correspondant à l'expérience A est représenté dans la Figure III.9.

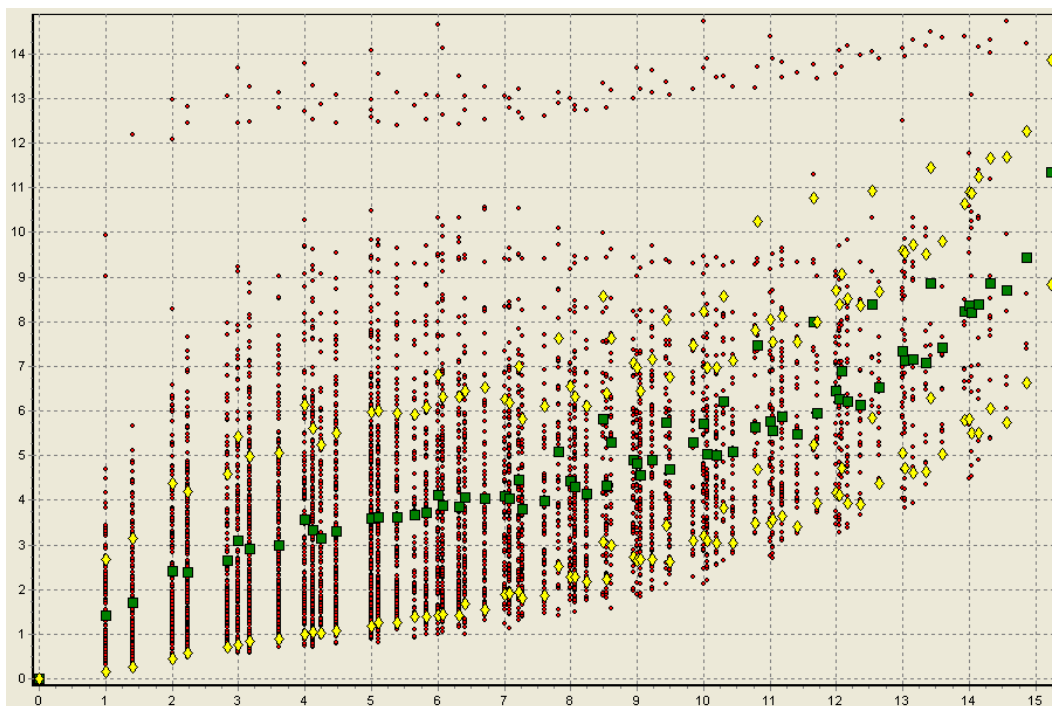


Figure III.9 : Graphique $dy-dx$ correspondant à la SOM obtenue avec les données de l'expérience A. (la moyenne \square et l'écart-type \diamond des dx sont également représentés pour chaque dy)

Ce graphique montre deux choses très différentes :

- La moyenne des dx pour chaque dy suit une courbe relativement rectiligne et monotone, ce qui tend à prouver que l'algorithme a plutôt bien fonctionné.
- Les écarts-types de ces mêmes dx sont assez importants, ce qui traduit le fait que le nuage de points obtenu est assez diffus. Comme nous l'avons plus haut, cela signifie que la projection n'est pas d'une très grande qualité : la topographie de l'espace de départ n'a pas été retranscrite avec beaucoup de fidélité dans la grille.

Pour poursuivre notre analyse, on aimerait visualiser directement la grille dans l'espace des poids (c'est-à-dire un espace à 13 dimensions), ce qui n'est évidemment pas possible dans le plan 2D d'un dessin. La technique de représentation curviligne proposée dans (Demartines 1992) permet de s'affranchir de ce problème. L'idée consiste à considérer que la grille de neurones décrit une surface courbée dans l'espace des poids. On représente alors les neurones selon leurs coordonnées curvilignes sur cette surface.

En pratique, on commence par positionner les neurones des deux lignes médianes de la grille comme les axes sur lesquels on va s'appuyer ; la distance entre neurones dans la représentation est directement leur distance dans l'espace des poids. On place ensuite les autres neurones de la grille, de proche en proche, du centre de la grille vers l'extérieur. Le schéma de la Figure III.10 illustre ce mécanisme décrit en détail dans (Demartines 1994).

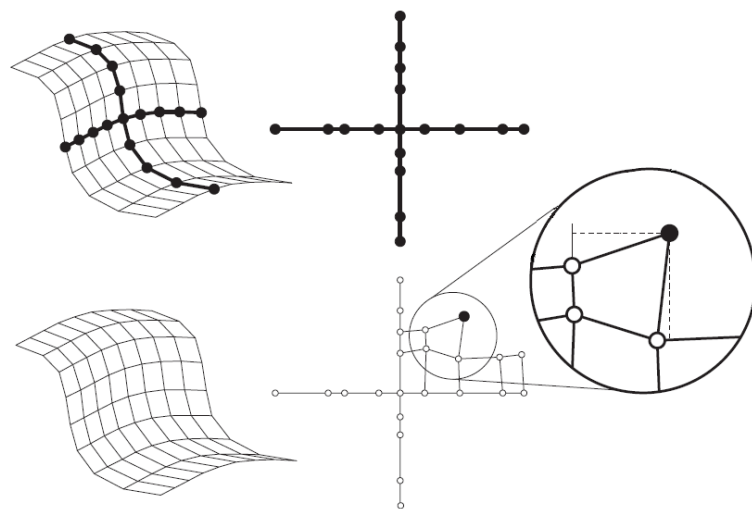


Figure III.10 : Construction de la représentation curviligne à gauche la grille de la SOM ; à droite, sa représentation curviligne

La représentation obtenue à partir de la carte de la Figure III.2 est représentée sur la Figure III.11.

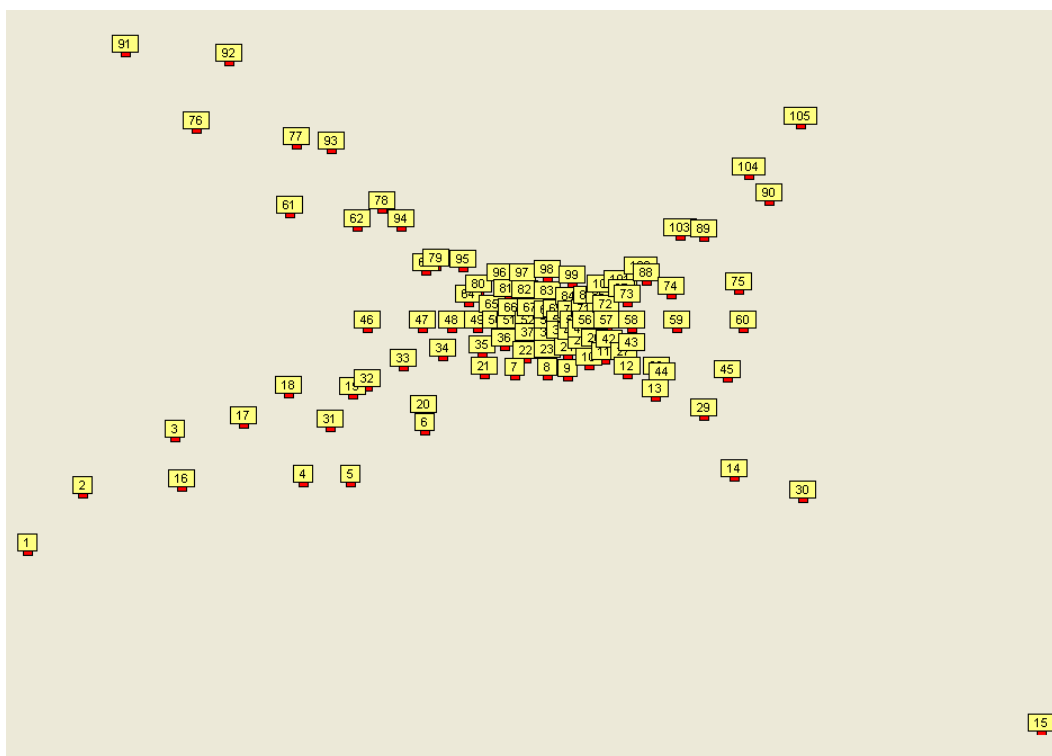


Figure III.11 : Représentation curviligne correspondant à la SOM de la Figure III.2 (les numéros des neurones sont identiques)

Sur cette représentation, on peut remarquer tout d'abord que les neurones correspondant à des défauts de type « rayure » sont clairement séparés des autres. Ceci est compréhensible, puisque leur apparence allongée est franchement différente de celle des autres défauts.

D'autre part, la disposition des neurones devrait être à peu près uniforme, à cause des propriétés de quantification de l'espace et d'organisation des SOMs ; force est de constater que ce n'est pas le cas. Etant donné que la qualité de l'apprentissage n'est pas à remettre en cause, cela signifie plutôt que la surface décrite par la grille dans l'espace des poids n'est pas une surface 2D. Le caractère bidimensionnel de la carte n'est donc pas compatible avec ce qu'elle doit représenter, d'où l'aspect étiré de la représentation.

On peut donc conjecturer que la qualité moyenne de la projection de l'espace des données (13 dimensions) dans l'espace 2D de la grille de la SOM est due au fait que cet espace d'arrivée est trop « petit ». Deux dimensions n'étant pas assez pour transcrire

intégralement l'information contenue dans la base de données de départ, la projection vers l'espace 2D de la grille se fait avec perte, et la topologie ne peut être reproduite avec précision. Ceci rejoint le concept de dimension intrinsèque, qui représente en fait le nombre de degré de liberté d'une distribution (le nombre de paramètres nécessaires à sa description complète, présenté au paragraphe III.2.1, p.102). Celle-ci a été évaluée pour les données de l'expérience A ; elle est comprise entre 6 et 7 (supérieure à 2 en tout cas), confirmant ainsi notre hypothèse.

En ce qui concerne l'hétérogénéité de certains neurones de la SOM, on est donc en droit de penser (même si cela ne peut être confirmé pour l'instant) que l'information permettant de lever ce problème est effectivement codée par les invariants de Fourier-Mellin mais perdue lors de la projection vers la grille de la SOM. Les résultats obtenus avec les SOMs restent en effet satisfaisants alors que l'on réduit d'un facteur 3 l'information disponible sur les images.

Dans le prochain paragraphe, nous allons étudier plus avant la problématique de réduction de dimension introduite ici. Une telle transformation peut en effet être bénéfique pour le système de classification en termes de performance brute et de temps d'exécution.

III.2 Réduction de dimension et classification neuro- nale

Lorsque l'on travaille avec des données de haute dimensionnalité, c'est-à-dire avec des données décrites par des vecteurs formés de nombreuses composantes, un certain nombre de problèmes spécifiques se posent ; ils sont désignés dans la littérature par l'expression générique « curse of dimensionality », le « fléau de la dimension ». Ce phénomène, que tente de limiter les techniques de réduction de dimension, est lié de manière subtile à la notion de dimension intrinsèque d'une distribution de données.

III.2.1 Dimension intrinsèque d'une distribution

Cette notion s'oppose à celle de dimension « brute » d'un espace. Considérons un problème mettant en jeu N formes décrites par des vecteurs à n composantes réelles. Ces données sont plongées dans \mathbb{R}^n , un espace à n dimensions : leur dimension brute est donc n . Leur dimension « intrinsèque » est par contre le nombre $m \leq n$ de degrés de liberté de l'espace où elles évoluent, c'est-à-dire le nombre minimum de variables libres nécessaires pour les représenter sans perte d'information. Cette dimension m est en général plus petite que la dimension brute n , à cause des relations linéaires et non-linéaires qui existent entre les composantes des vecteurs décrivant les données.

En termes plus généraux, pour (Fukunaga 1982), un ensemble de données inclus dans \mathbb{R}^n a une dimension intrinsèque égale à m si tous ses éléments sont situés dans un sous-espace de \mathbb{R}^n , de dimension m . Dans le cas des problèmes réels où les données sont toujours bruitées, il est cependant plus juste de considérer que les éléments issus de l'hypersurface de dimension m contiendront du bruit orthogonal à cette surface.

Si cette notion est intuitive, il existe plusieurs définitions formelles qui ramènent à elle :

- la dimension topologique d'un espace (Kégl 2007),
- la dimension fractale d'un espace (Grassberger and Procaccia 1983)
- la dimension d'une sous-variété (Demartines 1994).

La Figure III.12 illustre la différence entre dimension brute et intrinsèque. Cet exemple typique, tiré de (Demartines 1994), montre une distribution particulière (« dite en fer à cheval »), obtenue par échantillonnage d'une surface 2D « enroulée ». Cette surface

est représentée en gris sur la figure et les points de la distribution obtenue en noir. Cette distribution est plongée dans un espace 3D : sa dimension brute est donc 3. Par contre, elle est issue d'une surface 2D, sa dimension intrinsèque est donc 2. On peut s'en convaincre en projetant les points de la distribution sur cette surface (figure de droite).

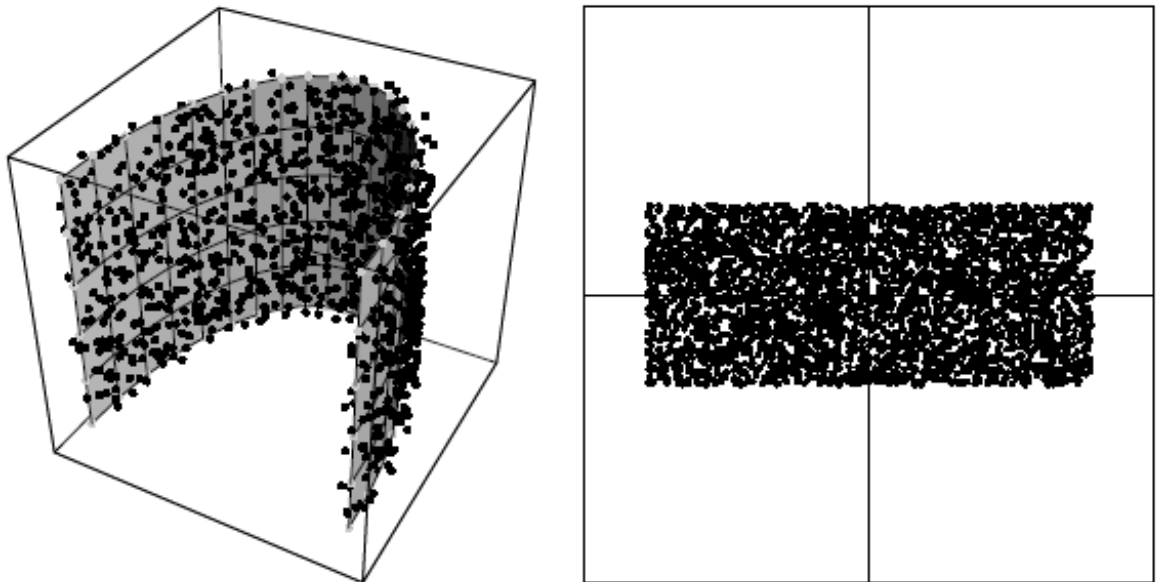


Figure III.12 : A gauche, la distribution « en fer à cheval » et la surface curviligne qui la sous-tend (en gris) ; à droite, projection de la distribution sur cette surface.)

Les méthodes permettant d'estimer la dimension intrinsèque d'une distribution peuvent être grossièrement divisées en deux groupes : celles qui s'appuient sur une projection des données, et celles qui se basent sur leur géométrie.

Méthodes basées sur une projection :

Ces approches, qui étudient les valeurs propres caractérisant les données considérées, se basent sur l'Analyse en Composantes Principales (voir paragraphe III.2.1, p.114). La dimension intrinsèque est dans ce cas déterminée par le nombre de valeurs propres plus grandes qu'un seuil fixé, ces valeurs propres correspondant aux axes supportant la plus grande partie de la variance observée.

Suivant ce principe, (Fukunaga and Olsen 1971) proposent d'effectuer une ACP sur l'ensemble des données. Cette technique souffre de deux inconvénients majeurs : le premier concerne le choix du seuil qui est fait a priori ; le deuxième tient au fait que l'ACP est une méthode par essence linéaire, et donc qui ne permet pas de révéler les dépendances

non linéaires entre les données.

Afin de passer outre ce deuxième problème, dans (Bruske and Sommer 1997) l'ACP est utilisée localement. Tout d'abord un graphe appelé Optimal Topology Preserving Map est construit en effectuant une quantification vectorielle de l'espace puis en reliant les prototypes voisins grâce à la règle d'apprentissage hebbienne proposée dans (Martinetz and Schulten 1991). Une ACP est ensuite appliquée à chaque nœud du graphe ; enfin, en moyennant le nombre de valeurs propres significatives pour tous les nœuds, on obtient l'estimation de la dimension intrinsèque recherchée. Cette évolution s'avère beaucoup plus robuste au bruit que la technique globale et elle est capable de prendre en compte les non linéarités de la distribution étudiée. Malheureusement, elle fait l'hypothèse d'une linéarité locale et donc est très sensible à la définition des voisinages, qui dépend elle-même fortement du nombre de prototypes dans la phase de quantification. Par ailleurs le problème du choix du seuil à prendre en compte n'est pas réglé et pourtant tout aussi déterminant pour la qualité de l'estimation obtenue.

Méthodes basées sur la géométrie :

Ces méthodes, qui exploitent la géométrie de la distribution de données, étudient en général une dimension fractale ou les distances entre points.

La plus populaire des dimensions fractales est la dimension de corrélation donnée par :

$$D = \lim_{r \rightarrow 0} \frac{\ln(C_n(r))}{\ln(r)} \quad (\text{III.7})$$

$$\text{avec } C_n(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} I(\|\vec{x}_i - \vec{x}_j\| \leq r)$$

où $I(H)$ vaut 1 si H est vraie et 0 sinon.

Dans les faits, (Grassberger and Procaccia 1983) préconisent de tracer la courbe des $\log(C_n(r))$ en fonction des $\log(r)$. La dimension de corrélation est alors estimée en calculant la pente de sa partie linéaire. Malheureusement, pour que cette estimation soit exacte à 5% près, il faut disposer de N points avec $N = 42^D$ (Camastra and Vinciarelli 2001). Pour une application réelle, ce nombre devient rapidement inaccessible, même pour de petites dimensions intrinsèques. La conséquence directe de ceci est que l'estimation fournie par cette technique est sous-évaluée. Afin de pallier à ce problème, on peut corriger la valeur obtenue par une procédure de calibration (Camastra and Vinciarelli 2002). On

calcule des estimations de dimension pour des distributions uniformes de N points de différentes dimensions intrinsèques connues ; on utilise ensuite cette courbe de calibration pour compenser le biais négatif de la méthode. Cette correction est néanmoins loin d'être parfaite car elle suppose que la distribution étudiée est également uniforme.

Une méthode relativement similaire consiste à estimer une dimension fractale différente, la dimension de capacité, par la technique des « packing numbers » (Kégl 2007). A noter que cette méthode constitue un cas limite de la technique présentée dans (Raginski and Lazebnik 2007), où est estimée une dimension dite de quantification.

Si ces méthodes s'intéressent aux distances entre voisins les plus proches de manière implicite, certaines méthodes l'utilisent directement. Elles trouvent leur justification dans le fait que, si l'on considère des variables (les points) $X_1 \dots X_n$ indépendantes et identiquement distribuées avec une densité $f(x)$ dans \mathbb{R}^m et que $T_k(x)$ est la distance Euclidienne d'un point x à son k -ième plus proche voisin, on a :

$$\frac{k}{n} > f(x) V_m [T_k(x)]^m \quad (\text{III.8})$$

où V_m désigne le volume d'une hypersphère de rayon unité dans \mathbb{R}^m . Autrement dit, la proportion de points de la distribution se trouvant à l'intérieur d'une boule autour de x est à peu près $f(x)$ fois le volume de la boule. (Levina and Bickel 2007) utilisent directement ce résultat pour calculer un estimé au maximum de vraisemblance de la dimension intrinsèque. Pour cela, ils examinent le nombre de voisins inclus dans des sphères dont le rayon est choisi tel que la densité de points peut être considérée comme constante et qu'assez de voisins y sont inclus. Ces hypothèses conduisent à une sous-estimation de la dimension intrinsèque dans les cas réels de grande dimension.

Dans (Costa and Hero 2004), la même idée est mise en œuvre de manière plus sophistiquée. La dimension intrinsèque est estimée en étudiant la longueur d'un Minimal Spanning Tree sur les distances géodésiques (qui sont calculées grâce à ISOMAP – voir paragraphe III.2.3.2.3, p.120). Cette technique nécessite l'utilisation d'ensembles de données issus d'un bootstrap aléatoire et donc les mesures de dimension ne sont pas répétables (!).

Finalement, le principal avantage de ces deux méthodes est qu'elles fournissent une garantie sur les propriétés statistiques (biais et variance) de l'estimation.

III.2.2 Le « fléau de la dimension »

Ce paragraphe qui discute des différents problèmes relatifs aux données de grandes dimensions est basé sur (Verleysen 2001) et (Demartines 1994).

La plupart des difficultés dues à la dimensionnalité des données ne sont pas spécifiques à l'utilisation des réseaux de neurones artificiels ou à une autre technique particulière mais plutôt à la nature des données elle-même. De même, elles sont communes à toutes sortes de traitements (clustering, classification, approximation...), qui peuvent être mis en œuvre dans un large éventail d'applications industrielles.

Ces difficultés proviennent toujours du fait que l'on dispose de trop peu de données par rapport à leur dimension. Le terme de « curse of dimensionality » s'applique au départ au domaine de l'approximation et de l'optimisation de fonction. Par exemple, pour optimiser ou approximer une fonction de d variables avec une erreur ε , il est nécessaire d'évaluer cette fonction un nombre de fois de l'ordre de $\left(\frac{1}{\varepsilon}\right)^d$, ce qui devient rapidement prohibitif même pour des dimensions modérées. Dans le domaine de l'apprentissage, ce fait est connu sous le nom de phénomène de l'« espace vide ». En effet, pour un nombre constant d'exemples, lorsque la dimension de l'espace augmente, il semble de plus en plus « vide ». Pour s'en convaincre, on peut calculer le volume $V_n(r)$ d'une hypersphère en n dimensions en fonction de son rayon r :

$$V_n(r) = \frac{\pi^{n/2} r^n}{\Gamma(1 + n/2)} \quad (\text{III.9})$$

où Γ désigne la fonction gamma d'Euler.

A cause de l'exposant n , ce volume augmente très rapidement avec l'accroissement de la dimension. Ainsi, on comprend intuitivement que le nombre d'exemples nécessaires à un apprentissage efficace augmente exponentiellement avec la dimension de l'espace des caractéristiques. Dans le cas particulier de l'approximation d'une distribution Gaussienne à l'aide d'un noyau Gaussien donné, il a été prouvé que cette intuition était exacte (Comon, Voz, and Verleysen 1994). D'autre part, il faut souligner que le nombre d'exemples requis devient rapidement bien trop grand pour être collecté par des systèmes réels.

Ce problème de base s'accompagne de plusieurs phénomènes surprenants qui peuvent être tout aussi nuisibles.

Concentration de mesure :

Si l'on considère des vecteurs dont les composantes sont aléatoires et suffisamment nombreuses, on remarque que, lorsque la dimension augmente, la variance de leurs mesures (distance, norme) reste presque constante alors que la moyenne de ces mêmes mesures augmente. En particulier, on montre dans (Demartines 1994), que, si \vec{x} est un vecteur à n composantes aléatoires, l'espérance E et la variance V de sa norme Euclidienne suivent la loi suivante :

$$E(\|\vec{x}\|) = \sqrt{an - b} + O(1/n) \quad (\text{III.10})$$

$$V(\|\vec{x}\|) = b + O(1/\sqrt{n})$$

où a et b sont des constantes dépendant de la loi suivie par les composantes du vecteur considéré. Pour des dimensions importantes, la moyenne de la norme Euclidienne augmente donc comme la racine carrée de la dimension, alors que sa variance reste quasi constante. Une des conséquences de ce phénomène est qu'en grande dimension, des vecteurs à composantes aléatoires semblent normés.

Perte des propriétés locales des noyaux gaussiens :

Les fonctions gaussiennes sont souvent utilisées pour leurs propriétés locales : la majeure partie de l'intégrale de telles fonctions est située dans un volume limité autour de leurs centres. Par exemple, 90% de la distribution issue d'une fonction gaussienne scalaire et normalisée tombe dans l'intervalle $[-1,65; 1,65]$. Cette propriété bien connue et intuitive n'a pourtant plus cours lorsque l'on évolue dans des espaces de grande dimension. Pour s'en convaincre, on étudie le pourcentage des échantillons issus d'une distribution gaussienne qui appartiennent à la sphère de rayon 1,65, par rapport à la dimension de l'espace (Figure III.13). Ce pourcentage décroît rapidement vers 0 : quand la dimension augmente, la plus grande partie du volume d'une fonction gaussienne est contenue dans sa « queue » plutôt que près de son centre.

Sur les hypervolumes :

Si le comportement des noyaux gaussiens échappe, pour des dimensions importantes, à notre sens commun, il en est de même pour la structure de l'espace. Par exemple, lorsque l'on compare le volume d'une hypersphère de rayon 1 avec celui d'un hypercube de côté 1,

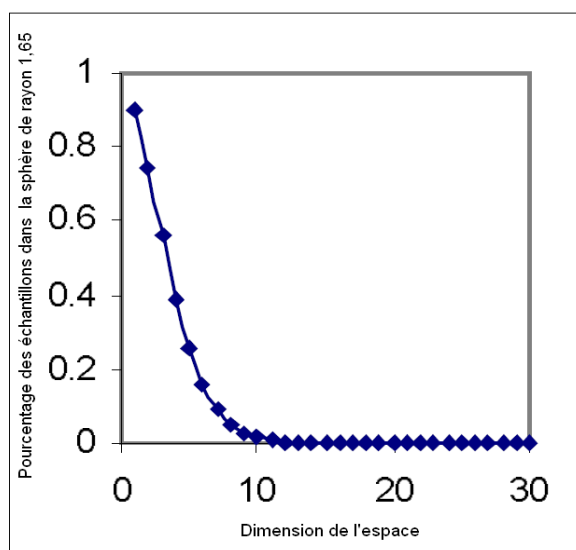


Figure III.13 : Pourcentage des échantillons issus d'une distribution gaussienne qui appartiennent à la sphère de rayon 1,65 en fonction de la dimension de l'espace (tiré de (Verleysen 2001)).

en faisant varier la dimension de l'espace, on obtient la courbe de la Figure III.14. On peut comprendre par notre expérience en 2-D et 3-D, que le ratio entre ces deux volumes diminue avec la dimension de l'espace. Le fait qu'il tende très rapidement vers 0 est beaucoup moins intuitif.

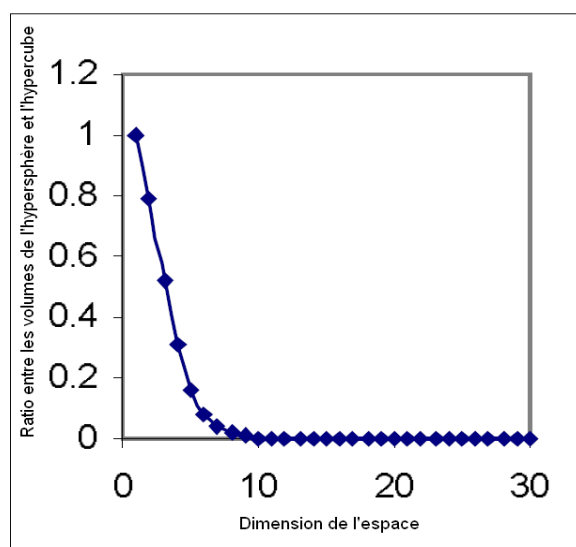


Figure III.14 : Ratio entre les volumes d'une hypersphère de rayon 1 et d'un hypercube de côté 1, en fonction de la dimension de l'espace (tiré de (Verleysen 2001)).

Ce phénomène est également illustré par la Figure III.15. Cette courbe représente le ratio entre les volumes d'hypersphères de rayons 0,9 et 1, en fonction de la dimension. Elle montre que pour des dimensions élevées, la majeure partie du volume d'une hypersphère est contenue dans une enveloppe sphérique dont l'épaisseur est 10% du rayon initial. Tout se passe donc comme si l'espace était vide en son centre puisque les volumes se concentrent dans la « périphérie » (ou dans les « coins ») de l'espace !

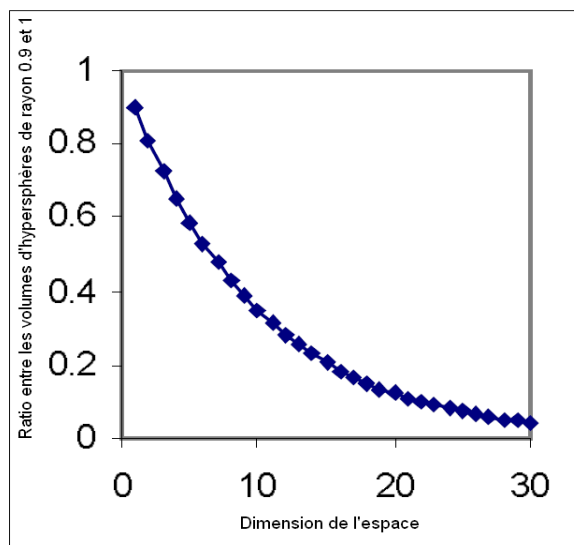


Figure III.15 : Ratio entre les volumes d'hypersphères de rayons 0,9 et 1, en fonction de la dimension de l'espace (tiré de (Verleysen 2001)).

Il est important de noter que tous les phénomènes contre-intuitifs que l'on vient de décrire sont déjà prépondérants pour des dimensions de l'ordre de la dizaine. A cause des différents problèmes exposés, traiter des données de haute dimension peut avoir des effets néfastes sur les performances des techniques neuronales, ce pour plusieurs raisons.

Premièrement, afin de couvrir l'ensemble de ces données, le réseau employé doit mettre en œuvre un grand nombre de neurones, ce qui implique beaucoup de paramètres libres. Ainsi, même si la phase d'apprentissage a été couronnée de succès, souvent, le réseau obtenu exhibe des résultats médiocres sur les ensembles de test. Cette mauvaise généralisation est due aux nombreux paramètres libres qui représentent des informations non pertinentes. Pour éviter ce problème, il faudrait disposer d'un nombre important d'exemples, ce qui est généralement extrêmement difficile. D'autre part, à cause de la lourdeur de l'architecture utilisée, les temps d'apprentissage peuvent devenir rédhibitoires.

Le phénomène de concentration de la mesure décrit plus haut a également un effet pervers sur ce type de techniques. En effet, la plupart des réseaux de neurones utilisent un calcul de distance entre les vecteurs présentés en entrée et leurs vecteurs de poids, que ce soit une norme Euclidienne (RBF, SOM) ou un produit scalaire (MLP). Or, nous avons vu que, pour des espaces de grandes dimensions, la variance des normes de vecteurs aléatoires devenait négligeable devant leur moyenne. Comme il en est de même pour une différence de vecteurs, les distances calculées par le réseau sont presque constantes ; en forçant le trait, on pourrait dire que les sorties du réseau ne dépendent même plus de ses entrées (!). Ce phénomène ne favorise en tous cas pas l'émergence d'une solution stable et robuste. Enfin, dans le cas particulier des RBFs, qui sont basés sur les noyaux gaussiens, la perte des propriétés locales de ceux-ci est évidemment problématique. En effet les noyaux gaussiens sont utilisés pour couvrir l'espace du problème ; malheureusement, pour des grandes dimensions, ils ne sont pas très représentatifs de leurs centres (puisque la plus grande partie de leur activité se fait loin de celui-ci).

Manipuler des données de dimension importante pose donc des problèmes difficiles à gérer. Heureusement, dans le cas d'applications traitant des problèmes réels, la redondance entre les variables observées structure fortement les distributions étudiées. Ainsi, la dimension intrinsèque associée au problème est bien plus petite que sa dimension brute. En outre, seule cette dimension intrinsèque doit être prise en compte dans le cadre du fléau de la dimension. Par exemple, concernant le phénomène de concentration des mesures évoqué plus haut, l'indépendance des composantes des vecteurs est une hypothèse fautive dans le cas général. Le résultat est cependant valide si l'on considère un nombre de composantes indépendantes égal à la dimension intrinsèque de la distribution considérée (pourvu que ce nombre soit assez grand). On peut faire le même constat pour le phénomène de l'espace vide. A ce sujet, un exemple éloquent est proposé dans (Demartines 1994) : si l'on considère une distribution de points en dimension 1000 contenue dans un hyperplan de dimension 2, une simple rotation peut permettre de remplacer les vecteurs coordonnées des points par des vecteurs dont seules les deux premières composantes ne sont pas nulles. Même si on manipule des vecteurs à 1000 composantes, toutes les caractéristiques statistiques de leur distribution sont celles d'une distribution en deux dimensions, où le phénomène d'espace vide n'existe pas.

Ainsi, on peut donc contourner le « fléau » de la dimension ou au moins en limiter les effets. Pour cela, il faut baisser la dimensionnalité du problème considéré. Cette opération de réduction est optimale si l'on se ramène à un espace de dimension égale à la dimension intrinsèque du problème, qui est théoriquement la taille minimale d'un espace permettant la description de la distribution considérée sans perte d'information.

III.2.3 Les techniques de réduction de dimension

Nous venons de voir que, pour limiter les effets négatifs du fléau de la dimension, on doit réduire la dimensionnalité du problème en conservant la majeure partie de l'information portée par les descripteurs originaux du problème. Cette réduction permet également de diminuer la complexité des modèles neuronaux utilisés ultérieurement, ce qui se traduit par un gain en temps de calcul et relâche la contrainte sur la taille de la base d'exemples nécessaires pour assurer un apprentissage correct.

Pour ce faire, deux approches sont couramment employées : la première consiste à sélectionner un sous-ensemble des paramètres disponibles (cette approche ne s'applique que pour des tâches de classification) ; la deuxième projette l'ensemble de ces paramètres vers un espace de dimension plus petite, en essayant de conserver au mieux l'information initiale.

III.2.3.1 Techniques par sélection d'un sous-ensemble de paramètre

Ces méthodes se composent généralement de deux phases : une phase d'ordonnement des paramètres, où les paramètres sont classés suivant leur intérêt par rapport à la classification visée ; et une phase de choix du nombre de paramètres à conserver, guidée par un critère d'arrêt. On ne s'intéressera ici qu'à des problèmes de discrimination de 2 classes, les techniques présentées se généralisant facilement à des problèmes mettant en jeu plus de classes.

En ce qui concerne la phase d'ordonnement, la technique la plus usuelle consiste à utiliser le critère de Fisher F pour évaluer la pertinence d'un paramètre. Ce critère caractérise en effet le pouvoir de discrimination d'un paramètre entre deux classes données.

Pour un problème à deux classes 1 et 2, il est défini comme suit :

$$F = \frac{(\mu_1 - \mu_2)^2}{N_1 V_1 - N_2 V_2} \quad (\text{III.11})$$

où μ_1 et μ_2 désignent les moyennes du paramètre considéré pour les 2 classes, V_1 et V_2 ses variances, et N_1 (resp. N_2) est le nombre d'observations associées à la classe 1 (resp. 2). Le principal inconvénient de ce critère est qu'il considère les paramètres isolément sans prendre en compte leurs éventuelles corrélations.

Afin de remédier à ce problème, il faut étudier directement des groupes de paramètre. Il est néanmoins impossible d'examiner tous les ensembles de paramètres possibles puisque leur nombre devient rapidement énorme (à cause de la combinatoire). Une solution sous optimale consiste alors à partir d'un ensemble vide et à rajouter successivement des paramètres à cet ensemble : c'est l'algorithme de recherche « Sequential Forward Selection » (SFS) (Kittler 1986). On commence par sélectionner le paramètre qui optimise un critère J de séparabilité entre classes ; ensuite, à chaque itération, on choisit parmi les caractéristiques restantes celle qui, en combinaison avec le sous ensemble déjà sélectionné, réalise la meilleure optimisation de J .

Il existe également un pendant à cette méthode, l'algorithme « Sequential Backward Selection » (SBS) : on part d'un ensemble constitué de tous les paramètres disponibles et, à chaque itération, on élimine celui qui provoque la diminution la plus faible du critère J . Plusieurs choix du critère J sont possibles (Oukhellou 1997)(Veropoulos 2001), mais ils se basent toujours sur les concepts de variance des populations à l'intérieur d'une classe (intraclasse) ou de covariance entre les populations de classes différentes (interclasse).

Une fois les paramètres classés selon leur importance par rapport au problème envisagé, il convient de choisir combien en conserver. Ce choix délicat ne peut être guidé par la connaissance de la dimension intrinsèque de la distribution étudiée puisque l'on cherche simplement à éliminer certains paramètres. En général, il est basé sur des seuils portant sur la perte d'information admissible, fixés a priori. A ces solutions dépendant largement des valeurs des seuils, on préférera le critère proposé par (Oukhellou 1997). L'idée est de rajouter aux caractéristiques initiales un paramètre aléatoire : on considère alors que tous les paramètres classés après celui-ci ne sont pas assez pertinents pour la séparabilité des classes et doivent être écartés. Cette approche par sélection d'un sous-ensemble des

paramètres initiaux peut s'avérer efficace (Oukhellou 1997)(Bartkowiak and Gorski 2006). Elle souffre cependant de limitations sévères.

Premièrement, elle n'est applicable que pour des tâches de classification et suppose la connaissance des classes des exemples manipulés, ce qui rend son utilisation parfois impossible. Mais son plus grand inconvénient est de purement et simplement mettre de côté une partie de l'information disponible sur le problème. En effet, d'une part les techniques présentées sont sous optimales, c'est-à-dire qu'elles ne garantissent pas la qualité de la solution trouvée ; et d'autre part, elles dépendent entièrement du critère choisi (séparabilité des classes). Or, si on comprend intuitivement la pertinence d'un tel critère, on conçoit également ses limitations ; par exemple, un paramètre n'étant pas très pertinent au sens de ce critère peut très bien porter seul une information permettant la discrimination des deux classes (il peut ne pas être très pertinent globalement, mais décisif à certains « endroits »). Ainsi il est probable que l'application de telles méthodes entraîne la perte d'informations qui pourraient être utiles pour la classification visée.

III.2.3.2 Techniques projectives

Au contraire, une autre approche consiste à projeter l'espace originel de haute dimension (l'espace des paramètres originaux) vers un espace de plus petite dimension, en conservant le maximum d'information. Elle ne souffre donc pas des limitations que nous venons d'exposer : cette projection est non-supervisée (elle ne nécessite pas la connaissance des classes) et, sous réserve que la dimension de l'espace d'arrivée soit bien choisie, l'information portée par tous les paramètres est sauvegardée.

En fait, on cherche à découvrir une structure de faible dimensionnalité sous-jacente à un espace de haute dimension : la projection des données sur cette structure permet la réduction de leur dimensionnalité. Les techniques permettant de mettre en œuvre cette idée sont désignées dans la littérature par la dénomination de « manifold learning », et, suivant la classification de (Costa and Hero 2004), elles se répartissent en trois catégories : les techniques linéaires, les techniques locales et les techniques globales. On va adopter ici cette typologie et présenter les techniques les plus significatives de chaque famille. Dans ce but, on étudiera la réduction de dimensionnalité d'une distribution de n points \vec{x}_i de dimension brute N (c'est-à-dire $\vec{x}_i \in \mathbb{R}^N$).

III.2.3.2.1 Techniques Linéaires

Analyse en Composantes Principales (ACP) :

Cette technique de réduction de dimension est probablement la plus communément utilisée. L'idée est de chercher les axes autour desquels l'inertie de l'ensemble des vecteurs est la plus faible, c'est-à-dire les axes le long desquels la variance est maximale. Pour ce faire, on forme la matrice de covariance C des n vecteurs $\vec{x}_i \in \mathbb{R}^N$ de la distribution étudiée, après que ceux-ci aient été centrés :

$$C = \frac{1}{n-1} \sum_{j=1}^n (\vec{x}_j - \vec{\mu}) \cdot (\vec{x}_j - \vec{\mu})^T \quad (\text{III.12})$$

$$\text{avec } \vec{\mu} = \frac{1}{n} \sum_{i=1}^n \vec{x}_i$$

Cette matrice est symétrique définie semi-positive, ce qui assure qu'elle est diagonalisable (elle possède N valeurs propres), que les valeurs propres sont toutes réelles, positives ou nulles et que les vecteurs propres sont orthogonaux entre eux. Chaque vecteur propre de la matrice représente un axe principal du nuage de données. Les valeurs propres correspondent aux variances de chacune des nouvelles variables. La réduction de dimension est accomplie en projetant les données sur les axes qui présentent une variance importante et en éliminant les autres. Dans les faits, on conserve en général un certain nombre de valeurs propres qui expliquent un pourcentage donné de la variance des données (classiquement 90 ou 95%).

Cette technique souffre cependant de limitations importantes : premièrement le seuil sur la variance expliquée est fixé a priori, mais surtout l'ACP est par essence linéaire et donc ne peut pas révéler les dépendances non-linéaires entre variables (qui sont pourtant très présentes dans les applications réelles). Une conséquence est que l'on ne peut pas se servir d'estimation de dimension intrinsèque pour choisir le nombre d'axes à conserver. Afin de s'affranchir de ce problème majeur, une version non-linéaire de la méthode a été proposée (Scholkopf, Smola, and Müller 1998). Elle utilise un noyau pour envoyer les données dans un espace de très haute dimension, puis l'ACP est appliquée à ces nouvelles données. Toute la difficulté est de choisir un noyau adéquat...

MultiDimensional Scaling (MDS) :

Le but de cette technique, très utilisée dans le domaine des sciences humaines, est de représenter des mesures de dissemblance entre des paires d'objets sous la forme de distances entre des points dans un espace de faible dimension. On cherche bien entendu à ce que la distance séparant les points dans cette nouvelle représentation respecte au mieux leur degré de similarité.

D'un point de vue opérationnel, cette méthode ressemble fortement à l'ACP à la différence que l'on ne diagonalise pas la matrice des covariances correspondant aux n points \vec{x}_i dans l'espace originel mais leur matrice de Gram (la matrice de leurs produits scalaires). Si l'on dispose des coordonnées des points, on peut former cette matrice $M = [m_{ij}]_{n \times n}$ directement ; dans le cas contraire, on peut extraire les produits scalaires entre points à partir de la matrice de leurs dissemblances $D = [d_{ij}]_{n \times n}$:

$$m_{ij} = \langle \vec{x}_i; \vec{x}_j \rangle = -\frac{1}{2} \left[d_{ij}^2 - \frac{1}{n} \sum_k^n d_{ik}^2 - \frac{1}{n} \sum_k^n d_{kj}^2 + \frac{1}{n^2} \sum_g^n \sum_h^n d_{gh}^2 \right] \quad (\text{III.13})$$

La matrice M est symétrique et définie semi-positive, elle est donc diagonalisable. La représentation \vec{y}_i des points \vec{x}_i dans un nouvel espace de dimension d est alors donnée par :

$$\vec{y}_i = \left(\sqrt{\lambda_1} v_{1i}, \dots, \sqrt{\lambda_k} v_{ki}, \dots, \sqrt{\lambda_d} v_{di} \right)^T \quad (\text{III.14})$$

où les λ_k sont les d plus grandes valeurs propres de M et les $\vec{v}_k = [v_{kl}]$ sont les vecteurs propres qui leur sont associés.

Etant donné que la distance Euclidienne peut être utilisée comme mesure de dissemblance, on peut appliquer cette technique pour obtenir une représentation de faible dimension où la distance entre deux points quelconques correspond autant que faire se peut à leur distance dans l'espace originel. La conservation des distances (si elle est effectivement réalisée) assure que la topologie de l'espace l'est aussi. A noter que cette idée de préservation des distances entre points est utilisée par d'autres techniques présentées plus loin.

Finalement, cette technique souffre des mêmes limitations que l'ACP, dont elle partage la caractéristique principale (linéarité).

III.2.3.2.2 Techniques locales

Linear Local Embedding (LLE) :

On cherche ici à conserver la topologie locale des données originelles lors de la réduction de dimension. Pour cela, on fait l'hypothèse que la distribution étudiée, globalement non-linéaire, est linéaire quand on la considère de manière locale. L'idée est alors d'exprimer, à l'aide de projections linéaires, les relations locales de voisinage existant entre points (ces projections représenteront la géométrie locale des données), puis de préserver ces relations locales lors de la réduction de dimension.

La méthode se compose de trois phases (Roweis and Saul 2000). Premièrement, pour chaque point, on constitue un voisinage composé des k points qui sont les plus proches de lui (les k plus proches voisins ou k -NN). On peut alternativement considérer un ε -voisinage (c'est-à-dire l'ensemble des points situés à l'intérieur d'une sphère de rayon ε autour du point dont on cherche les voisins). Ensuite, on exprime chacun des points de la distribution comme une combinaison linéaire de ses voisins. Pour cela, on minimise l'erreur de reconstruction :

$$E(w) = \sum_{i=1}^n \left| \vec{x}_i - \sum_{j \in V(\vec{x}_i)} w_{ij} \vec{x}_j \right|^2 \quad (\text{III.15})$$

où les \vec{x}_i sont les n points de la distribution considérée, $V(\vec{x}_i)$ le voisinage associé à chacun d'eux et $w = [w_{ij}]_{n \times n}$ la matrice de reconstruction recherchée.

En ajoutant la contrainte suivante :

$$\forall i \in [1 \dots n] \quad \sum_{j \in V(\vec{x}_i)} w_{ij} = 1 \quad (\text{III.16})$$

on peut calculer les poids de reconstruction par une méthode aux moindres carrés. La troisième étape consiste à reproduire au mieux la topologie locale dans le nouvel espace de dimension d ($d \leq N$). Pour cela, on minimise la fonction suivante :

$$\Phi(y) = \sum_{i=1}^n \left| \vec{y}_i - \sum_{j \in V(\vec{x}_i)} w_{ij} \vec{y}_j \right|^2 \quad (\text{III.17})$$

où les \vec{y}_i sont les projections des \vec{x}_i dans le nouvel espace. L'optimisation de cette forme quadratique porte sur les \vec{y}_i (les w_{ij} sont fixés pendant l'étape précédente) et fournit finalement la représentation recherchée.

Il existe quelques autres méthodes locales, non présentées ici pour deux raisons : LLE est de loin la technique de cette famille la plus utilisée ; et d'autre part, ces autres méthodes peuvent être considérées comme des évolutions de celle-ci.

On citera tout de même les Laplacian Eigenmaps, où la représentation de basse dimension est obtenue en diagonalisant la matrice correspondant au Laplacien du graphe de voisinage. L'auteur de cette technique souligne le rapport ténu qu'elle entretient avec LLE (Belkin and Niyogi 2002).

III.2.3.2.3 Techniques globales

Self Organizing Maps (SOM) :

Cette technique classique a déjà été présentée dans le paragraphe III.1.2.2.1 (page 90). Son principal inconvénient est que plusieurs de ses paramètres doivent être fixés a priori, certains d'entre eux étant très sensibles pour la qualité de la projection obtenue (par exemple la forme de la grille doit être adaptée à la distribution étudiée). D'autre part, elle ne fournit qu'une projection discrète (puisque l'espace d'arrivée est l'espace de la grille de neurones), ce qui n'est pas souhaitable dans l'optique de la réduction de dimension utilisée comme prétraitement.

Curvilinear Components Analysis (CCA) :

Le but de la technique Curvilinear Components Analysis (CCA) est de reproduire la topologie d'un espace originel de dimension N dans un espace de dimension P (avec $P \leq N$), dans lequel on souhaite projeter les données, sans fixer aucun paramètre de configuration de la topologie (Demartines and Héroult 1993). Pour ce faire, un critère caractérisant les différences entre les topologies des espaces originel et projeté est calculé à partir des distances entre points (l'idée est de rendre équivalentes, autant que faire se peut, les distances entre les points dans l'espace originel et dans l'espace projeté) :

$$E_{CCA} = \frac{1}{2} \sum_i \sum_{j \neq i} (d_{ij}^n - d_{ij}^p)^2 F(d_{ij}^p) \quad (\text{III.18})$$

où d_{ij}^n (resp. d_{ij}^p) est la distance Euclidienne entre les vecteurs \vec{x}_i et \vec{x}_j de la distribution étudiée dans l'espace originel (resp. dans l'espace projeté), et F est une fonction décroissante qui favorise la conservation des « petites » distances.

Une telle fonction est nécessaire puisque dans le cas général les distances ne peuvent être respectées à toutes les échelles lors de la réduction de dimension : on choisit alors de favoriser la topologie locale par rapport à la topologie globale.

Cette fonction d'énergie est minimisée par une descente de gradient stochastique (Demartines and Hérault 1995) :

$$\forall i \neq j, \Delta x_i^p = \alpha(t) \frac{d_{ij}^n - d_{ij}^p}{d_{ij}^p} u(\lambda(t) - d_{ij}^p)(x_i^p - x_j^p), \quad (\text{III.19})$$

où $\alpha : \mathbb{R}^+ \rightarrow [0; 1]$ et $\lambda : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ sont deux fonctions décroissantes, correspondant respectivement à un paramètre d'apprentissage et à un facteur de voisinage, et u est la fonction de Heaviside.

La méthode CCA fournit également un algorithme d'extrapolation-interpolation, qui permet de projeter, de manière continue, de nouveaux points de l'espace originel dans l'espace projeté, en utilisant la connaissance des N vecteurs déjà projetés. Pour cela, on adapte simplement l'algorithme de projection décrit plus haut ; pour chaque point \vec{x}_i à projeter, on définit une fonction d'erreur que l'on minimise par descente de gradient :

$$E_i = \frac{1}{2} \sum_{j=1}^N (d_{ij}^n - d_{ij}^p)^2 F(d_{ij}^p) \quad (\text{III.20})$$

L' algorithme de descente de gradient utilisé par CCA étant assez lourd, il est souvent précédé d'une phase de quantification, ce qui permet d'alléger considérablement les calculs : seuls des points d'ancrage sont projetés au départ, puis l'algorithme d'extrapolation-interpolation est utilisé pour projeter les autres points de la distribution d'origine.

Curvilinear Distance Analysis (CDA) :

Cependant, CCA rencontre des difficultés lors du « dépliage » de distributions fortement non-linéaires. Afin de minimiser ce problème, une évolution de cet algorithme appelée Curvilinear Distance Analysis (CDA) a été proposée (Lee et al. 2000). Cette technique utilise les distances curvilignes au lieu des distances Euclidiennes, dans le but de mieux approximer les distances géodésiques le long de la surface à déplier.

La différence entre ces deux sortes de distances est illustrée par la Figure III.16, qui montre bien qu'elles peuvent être fortement dissemblables. Dans cette figure, on considère deux points d'un espace non-linéaire, une spirale (a) ; la distance Euclidienne entre ces 2 points (b) est beaucoup plus petite que la distance géodésique qui les sépare (le long de la spirale) (c).

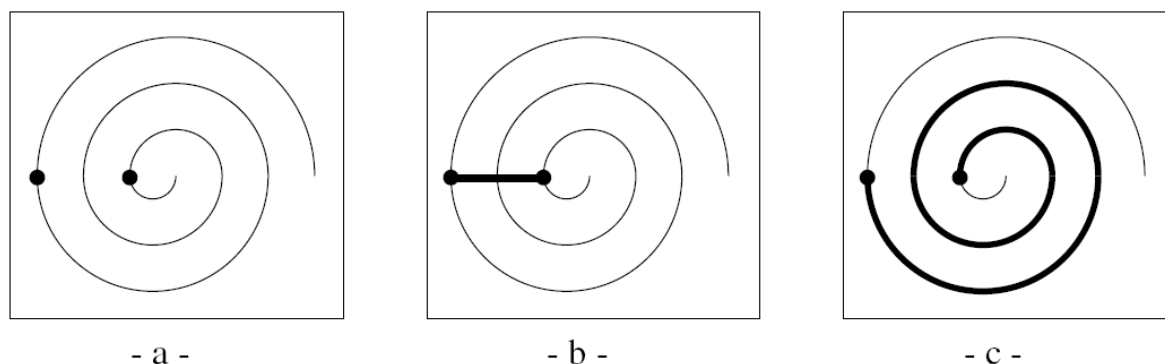


Figure III.16 : Différences entre distances Euclidienne et curviligne (tiré de (Lee et al. 2000)).

Dans les faits, les distances curvilignes sont calculées en deux temps. Premièrement, on construit un graphe de voisinage entre les points de la distribution, en utilisant le voisinage induit par les k -NN, l' ε -voisinage ou tout autre relation permettant d'en définir un. Les liens de ce graphe sont pondérés par la distance Euclidienne entre les points qui sont mis en relation. La distance curviligne entre deux vecteurs est alors calculée comme la distance minimale entre ces deux vecteurs dans le graphe, en utilisant par exemple l'algorithme de Dijkstra.

Finalement, on applique l'algorithme original de la CCA en utilisant les distances curvilignes. Cet algorithme permet de traiter des distributions présentant de grandes non-linéarités et est beaucoup plus robuste que CCA par rapport au choix des fonctions α et λ .

Sammon's mapping (MDS non linéaire) :

Cette technique, bien antérieure à la CCA, cherche également à minimiser les différences entre les distances entre points dans l'espace originel et l'espace projeté (Sammon 1969), d'où son nom de MDS non linéaire. Pour se faire un critère de « stress » est minimiser par une descente de gradient :

$$S = \frac{1}{2} \frac{\sum_i \sum_{j \neq i} \frac{(d_{ij}^n - d_{ij}^p)^2}{d_{ij}^n}}{\sum_i \sum_{j \neq i} d_{ij}^n} \quad (\text{III.21})$$

où d_{ij}^n (resp. d_{ij}^p) est la distance Euclidienne entre les vecteurs \vec{x}_i et \vec{x}_j de la distribution étudiée dans l'espace originel (resp. dans l'espace projeté).

On peut remarquer que ce critère s'apparente à celui utilisé par la CCA, avec la différence majeure que la fonction de pondération F , qui sert à favoriser le respect des courtes distances, dépend uniquement des distances d'entrée. Ceci limite considérablement les facultés de dépliage. D'autre part, cette méthode s'avère très lourde en calcul et peu fiable pour des structures de grande dimension, de l'aveu même de son inventeur.

ISOMAP :

Cette technique présentée dans (Tenenbaum, de Silva, and Langford 2000), propose d'utiliser la méthode MDS en utilisant comme mesure de dissemblance non pas la distance Euclidienne mais les distances géodésiques le long de la surface à déplier. Pour se faire, ces distances sont approximées de la même manière qu'expliqué plus haut (pour la méthode CDA).

Dans ce travail, nous nous sommes intéressés en priorité à la CCA car plusieurs papiers font état de sa supériorité sur les autres techniques exposées plus haut, en termes de qualité de projection et d'efficacité algorithmique.

Tout d'abord, l'ACP est supplantée par les techniques non-linéaires (dont la CCA), à cause des limitations déjà mentionnées (principalement son caractère linéaire). Ceci est démontré dans (de Backer 2002) et (Lennon et al. 2001) dans le cas d'applications réelles, respectivement de classification de textures et de clustering d'images satellites.

D'autre part, dans (Duchene 2003) et (Bouveyron 2003), différentes méthodes de réduction de dimension non-linéaires sont comparées pour une tâche donnée ; il s'agit ici de « déplier » la surface corticale issue d'images IRMf du cerveau humain, qui est caractérisée par une courbure intrinsèque variable et localement fort. Ces travaux montrent que la CCA est plus performante que les techniques de Sammon's Mapping, ISOMAP et LLE, en termes de distorsion surfacique et de temps de calcul.

III.2.4 Réduction de dimension et classification neuronale des défauts

Travailler avec des données de haute dimension est problématique, en particulier pour les méthodes neuronales. Il paraît donc clair d'un point de vue théorique, qu'une réduction de dimension est favorable pour les performances brutes d'un classifieur mettant en oeuvre de telles techniques (sous réserve bien entendu que cette réduction soit d'une qualité suffisante), en plus du gain en complexité et en temps de calcul qu'elle entraîne. En fait, il existe de nombreux exemples dans la littérature d'applications qui confirment qu'une telle réduction est bénéfique ou au moins efficace. Par exemple, dans (Grother 1992), (Pilato 1998), et (Hong et al. 2004), l'utilisation de l'ACP pour la réduction de dimension est validée (elle permet d'atteindre des performances correctes) respectivement pour une tâche de reconnaissance de caractères manuscrits et pour un système d'analyse de spectre de rayonnements γ , mettant en oeuvre des MLPs, ainsi que pour un système de détection de dérive dans un process industriel utilisant un réseau de neurones modulaire (voir paragraphe III.3.1, p.141). Une telle validation concernant l'utilisation de la CCA est donnée dans (Vigneron 1997) qui décrit une application de spectrométrie γ , et dans (Pilato 1998) où est développé un système de classification de mouvement de produits radioactifs ; ces deux exemples emploient des MLPs. Dans (Buchala et al. 2005), l'auteur démontre que l'utilisation d'une réduction de dimension par CCA améliore très sensiblement les performances de classifieurs de type MLP et Support Vector Machine (SVM) lors d'une tâche de classification de photos d'adulte selon leur sexe, par rapport à l'utilisation des données brutes. Enfin, on peut évoquer (Silvén, Niskanen, and Kauppinen 2000), où la réduction de dimension des données brutes par SOM est suffisamment efficace pour permettre la classification de planches de bois par une simple labellisation de la carte obtenue.

Il faut cependant souligner que le bénéfice tiré d'une réduction de dimensionnalité est largement dépendant de la qualité de celle-ci : la projection effectuée ne doit pas induire de graves distorsions de la topologie des données, et surtout la dimension finale des données doit être en accord avec leur dimension intrinsèque. Si ce n'est pas le cas, la perte d'information entraînée peut nuire à la qualité d'une classification ultérieure.

Afin de confronter l'intérêt des techniques SOM, CCA et CDA, nous avons traité les données déjà étudiées dans le paragraphe III.1.2.2.2, grâce à ces trois techniques de réduction de dimension. Outre le fait de pouvoir comparer la qualité des projections induites par ces différentes techniques, les expériences décrites dans ce chapitre nous ont permis d'étudier la structure de la distribution manipulée et d'évaluer l'impact possible de telles transformations sur une tâche de classification (Voiry et al. 2007).

III.2.4.1 Etude des techniques SOM, CCA et CDA, appliquées aux « imageries » de défauts

Protocole expérimental :

Tout d'abord, rappelons ici brièvement la nature des données manipulées. Un grand nombre d'imageries de défauts ont été collectées lors de trois expérimentations mettant en œuvre deux pièces optiques différentes ayant bénéficié d'un nettoyage préalable, ou pas. Le Tableau III.2 rappelle les paramètres de ces trois expériences et le lecteur pourra se reporter au paragraphe III.1.2.2.2 (p.91) pour plus de précisions.

| Expérience | Identifiant de la pièce optique | Nettoyage | Nombre de champs étudiés | Aire correspondante | Nombre d'imageries recueillies |
|------------|---------------------------------|-----------|--------------------------|----------------------|--------------------------------|
| A | 1 | NON | 1178 | 28 cm ² | 3865 |
| B | 2 | NON | 605 | 14 cm ² | 1910 |
| C | 2 | OUI | 529 | 12,5 cm ² | 1544 |

Table III.2 : Paramètres des trois expériences pendant lesquelles ont été acquises les imageries.

Ces imageries ont été ensuite codées sous forme d'un vecteur à 13 composantes en utilisant la transformation de Fourier-Mellin; les vecteurs obtenus ont finalement été centrés-réduits.

Nous avons alors mené trois expériences de réduction de dimension : les vecteurs codant les imageries issues de l'expérimentation B (qui seront désignés par le terme de base de

données B) ont été projetés dans un espace de dimension 2 à l'aide des techniques SOM, CCA et CDA.

Cependant, on ne peut pas confronter directement les différentes représentations obtenues à cause de la nature discrète de la projection effectuée par la SOM. Afin que les résultats de ces trois expériences soient comparables, les projections issues de CCA et CDA furent traitées par l'algorithme SOM en utilisant la même forme de grille que dans l'expérience mettant en œuvre une SOM (20 x 8). Il est important de noter que, dans ces deux cas, l'algorithme SOM agissant sur un espace de dimension 2, il est utilisé uniquement pour accomplir une quantification de l'espace et non pour une réduction de dimension. Ainsi, on peut comparer directement les capacités relatives à la réduction de dimension des trois techniques en confrontant les différentes cartes SOM obtenues (directement dans le cas de la projection par SOM; en deux étapes pour ce qui est de la CCA et de la CDA). Le schéma de la Figure III.17 résume le déroulement des expériences menées.

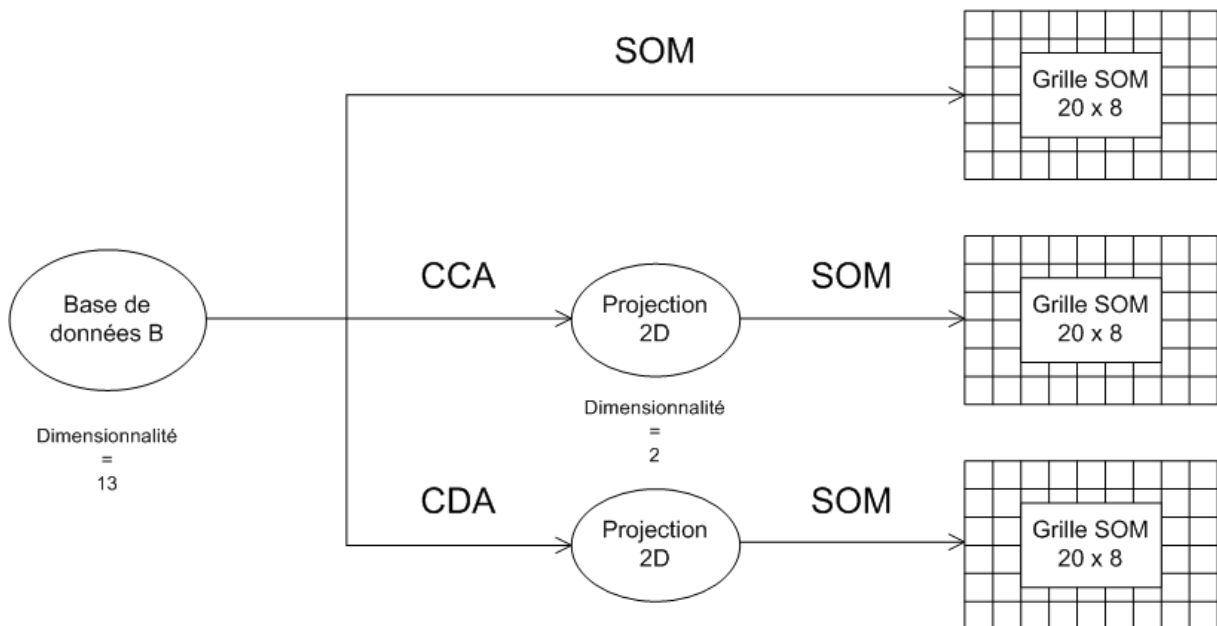


Figure III.17 : Représentation schématique des trois expériences menées.

Qualité de la projection :

Dans le but de comparer la qualité des projections obtenues, nous avons utilisé la représentation dy-dx (paragraphe III.1.2.2.3, p.97). Les graphiques correspondant aux trois différentes réductions de dimension sont représentés dans la Figure III.18.

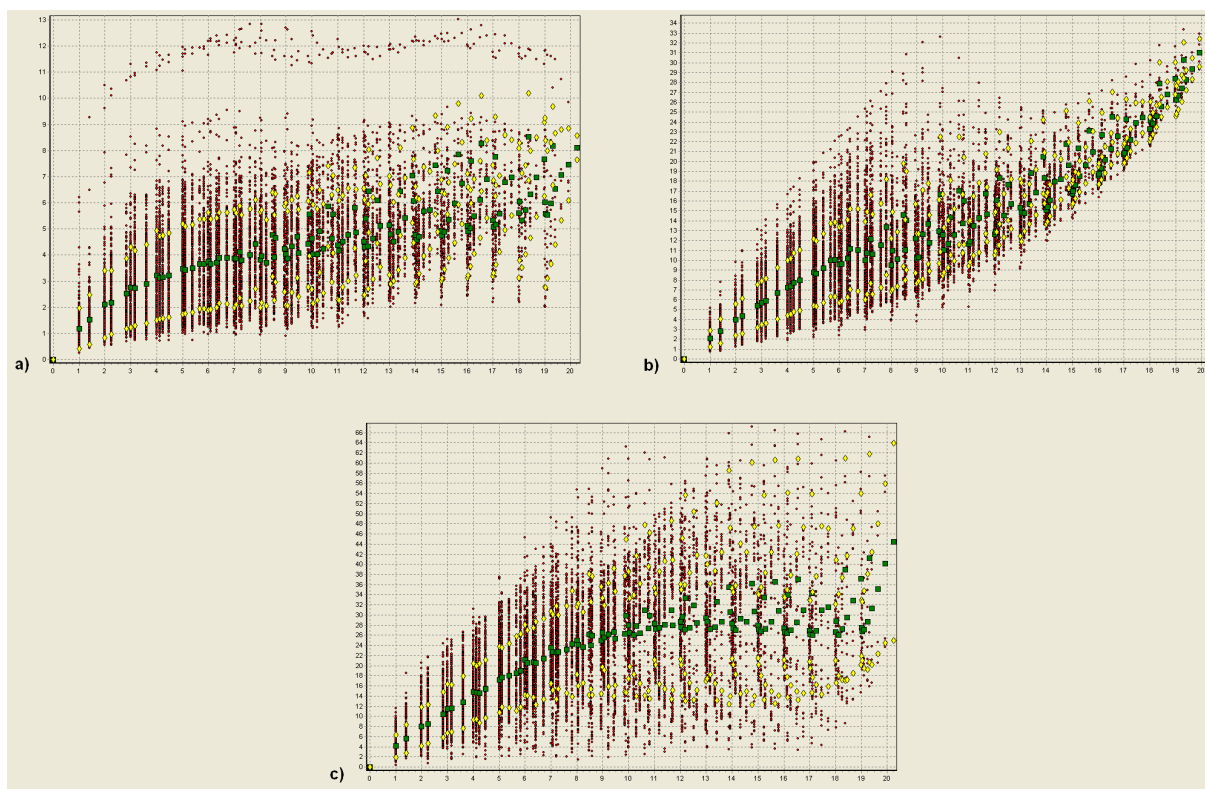


Figure III.18 : Représentation $dy-dx$ des trois SOMs obtenues (la moyenne \square et l'écart-type \diamond des dx sont également représentés). Gauche : SOM ; milieu : CCA ; droite : CDA.

Premièrement, le nuage de points correspondant à la projection SOM est beaucoup plus diffus que dans le cas de la CCA. De plus, la courbe constituée par la moyenne des dx pour chaque dy est moins uniformément croissante. Cela traduit le fait que les distances entre prototypes sont moins bien conservées : CCA est plus efficace que SOM, alors qu'approximativement la même quantité est minimisée par les deux algorithmes (on cherche dans les deux cas à maintenir les liens de voisinages topologiques lors de la projection).

Le nuage de points obtenu dans le cas de CDA est quant à lui très différent. Les distances dy sont en effet liées à la distance curviligne et non à la distance Euclidienne. La figure est cependant très ressemblante à celle associée à la CCA pour de petites valeurs dy , car, pour de petites distances, la distance Euclidienne constitue une bonne approximation de la distance curviligne. On peut donc considérer d'une part que la distribution de points étudiée est localement linéaire, et d'autre part, que l'utilisation de la distance curviligne est pertinente puisqu'à grande distance, elle est très différente de la distance Euclidienne.

Impact possible sur une tâche de catégorisation :

Si l'étude de la qualité de la projection nous apporte de précieux renseignements sur les capacités de chacune des techniques à conserver la topologie de l'espace lors de la réduction de dimension, elle ne nous éclaire pas sur l'impact que peut avoir une telle réduction sur une tâche de classification. Sur ce point, on présume cependant intuitivement qu'une « bonne » réduction de dimension facilite une classification ultérieure, en conservant le plus d'information possible sur le problème à traiter.

Afin d'étudier le rapport entre réduction de dimension et classification, on va évaluer l'impact de la réduction de dimension présentée plus haut sur une tâche de catégorisation réelle : discriminer les défauts « permanents » des « faux » défauts. Pour cela, on considère la base de données C (Tableau III.2), constituée d'une très grande majorité de défauts permanents puisque provenant de la mesure d'une pièce nettoyée, et l'on projette ses éléments dans les trois cartes SOM obtenues précédemment. A noter que cette projection consiste à trouver, pour chaque élément de la base, le prototype de la carte qui lui ressemble le plus (au sens de la norme Euclidienne) ; ce qui implique tout de même dans les cas de la CCA et de la CDA, une projection préalable des éléments vers un espace 2D à l'aide de ces deux techniques.

A dessein de comparaison, on projette les données brutes (de dimension 13) vers des prototypes issus de l'algorithme k -means, avec $k = 20 \times 8 = 160$. Etant donné que l'algorithme des k -means a un comportement équivalent à celui des SOMs, excepté en ce qui concerne les contraintes liées au voisinage, il a le même effet sur la distribution des éléments projetés, mais ne permet pas de représentation visuelle pertinente.

La répartition des projections des éléments de la base de données C dans les différentes cartes est représentée dans les Figures III.19, III.20 et III.21, correspondant respectivement à la réduction de dimension par SOM, CCA et CDA. Dans ces figures, le nombre d'éléments projetés dans chaque cellule est reporté en son centre ; pour faciliter la visualisation des résultats, ce nombre est également représenté en niveau de gris (plus la teinte est claire, plus le nombre d'éléments est important).

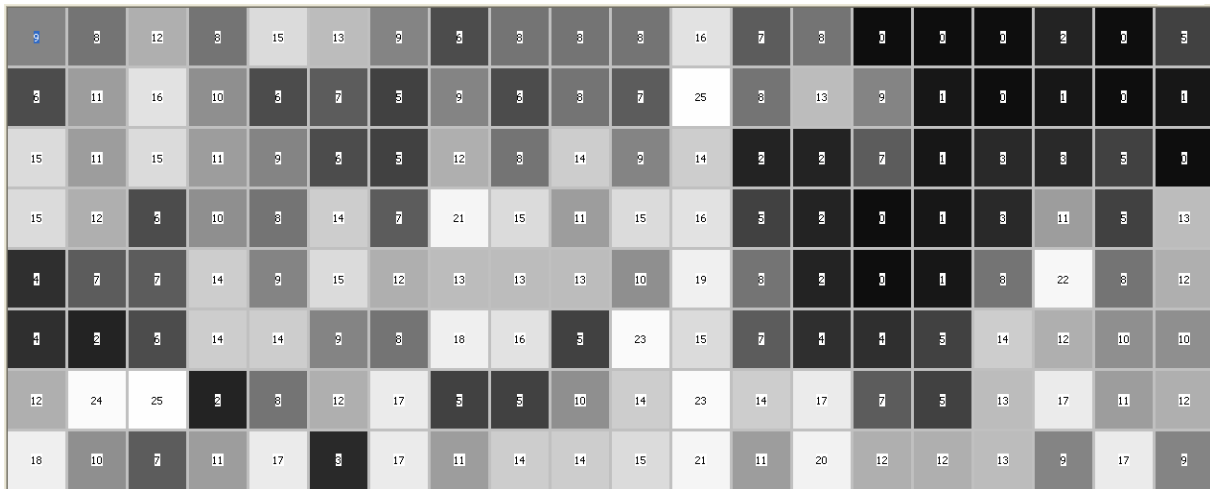


Figure III.19 : Répartition des éléments de la base de données C projetés dans la carte SOM.

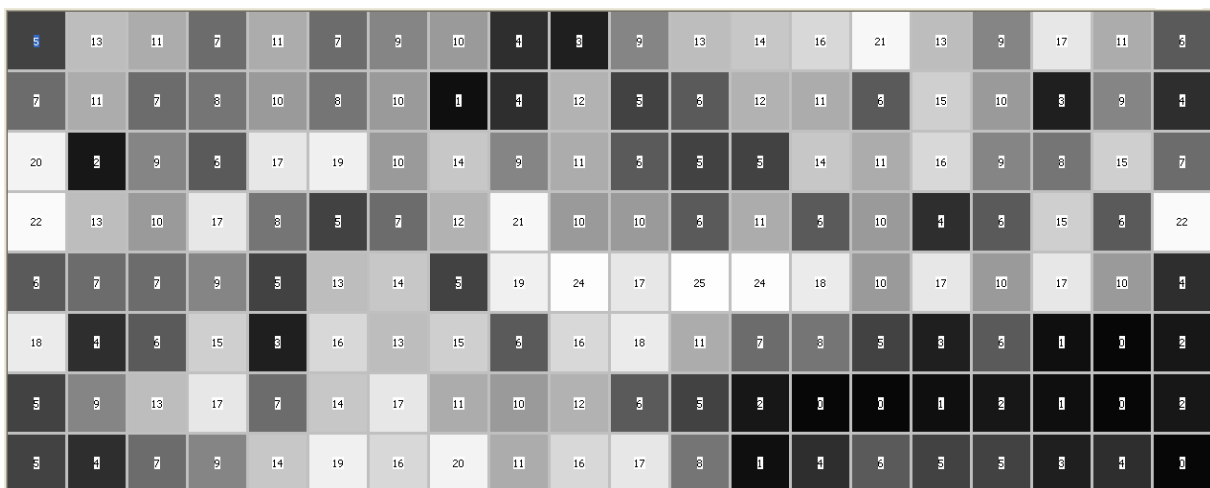


Figure III.20 : Répartition des éléments de la base de données C projetés dans la carte SOM, après réduction de dimension par CCA.

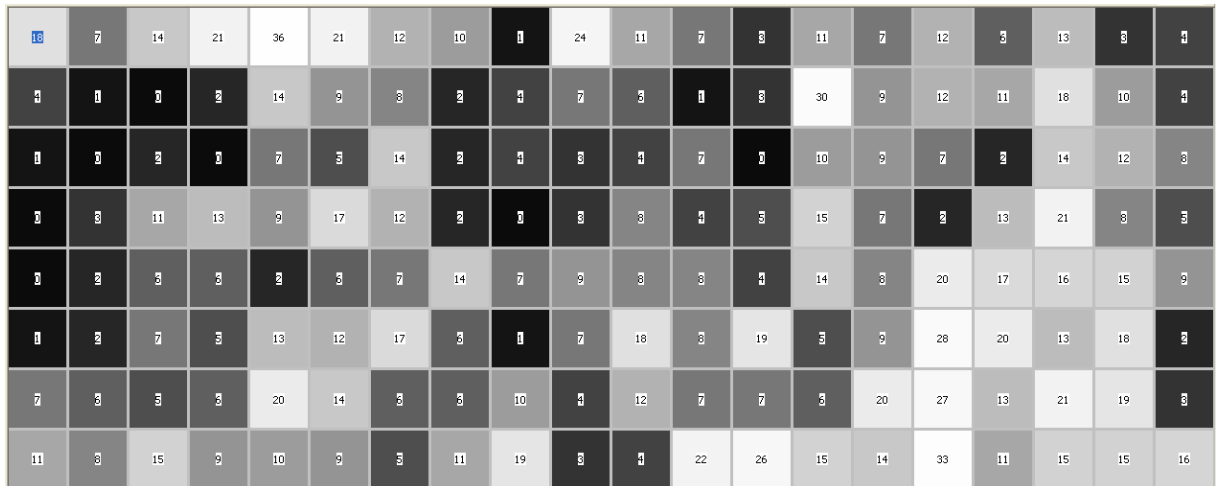


Figure III.21 : Répartition des éléments de la base de données C projetés dans la carte SOM, après réduction de dimension par CDA.

Afin d'étayer ces cartes dont on ne peut faire d'interprétation que qualitative, nous avons également calculé plusieurs valeurs caractéristiques évaluant l'« homogénéité » des répartitions de défauts permanents dans celles-ci. Dans le Tableau III.3 sont ainsi reportés l'entropie et l'écart-type du nombre d'éléments projetés dans chaque cellule ainsi que le nombre de cellules vides ou quasi-vides (moins de trois éléments).

| Techniques de réduction de dimension | Écart-type du nombre de défauts projetés par cellule | Entropie de la distribution des défauts projetés | Nombre de cellules vide | Nombre de cellules avec moins de 3 éléments |
|--------------------------------------|--|--|-------------------------|---|
| Aucune | 8.72 | 2.055 | 15 | 30 |
| SOM | 5.78 | 2.114 | 9 | 26 |
| CCA | 5.72 | 2.121 | 5 | 20 |
| CDA | 7.04 | 2.088 | 7 | 32 |

Table III.3 : Différentes mesures caractérisant la distribution des éléments de la base de données C (défauts permanents) projetés dans les SOMs obtenues précédemment.

L'idée sous-tendant ces différentes mesures est la suivante : le cas le plus défavorable pour une classification ultérieure, qui aurait pour but de discriminer les défauts permanents des autres défauts, est une répartition complètement homogène des éléments de la base de données C dans la carte SOM. En effet, par nature, SOM est une technique de quantification : un nombre à peu près constant d'éléments de la base de données B (qui a servi à l'apprentissage), contenant des représentations de défauts de tous types, est projeté dans chacune des cellules de la grille. On comprend alors que s'il en est de même pour les éléments de la base de données C (presque tous des défauts permanents), la classification est rendue plus difficile car elle nécessite la définition d'une frontière entre classes dans chacune des cellules de la grille. Une telle répartition homogène des défauts permanents dans la grille se traduit par une entropie élevée, un écart-type et un nombre de cellules vides faibles.

Les cartes et les mesures numériques obtenues pour SOM et CCA sont comparables et donc ces deux techniques semblent être équivalentes pour le problème considéré. Il faut cependant rappeler que SOM est plus difficile à mettre en œuvre, puisqu'elle nécessite le choix difficile et a priori de la forme de la grille ; d'autre part, CCA apporte plus d'informations car elle met en œuvre une projection continue.

CDA offre les mêmes avantages que CCA mais cette technique semble plus appropriée pour le prétraitement avant une hypothétique discrimination des défauts « permanents » (base de données C). La carte correspondante semble en effet exhiber des aires plus spécifiques aux éléments de la base de données C que les deux autres. Cela se traduit par la présence de moins de zones « grises », où les « faux » défauts et les défauts « permanents » sont mélangés, ce qui est évidemment positif pour une classification ultérieure. Cette impression est confirmée par les mesures numériques : l'entropie est plus basse que pour les expériences utilisant SOM et CCA, ce qui indique une meilleure organisation des éléments projetés ; l'écart-type est plus grand, car il existe un meilleur contraste entre les zones avec et sans défauts permanents ; enfin, le nombre de cellules qui sont quasi-vides est plus grand (rappelons ici que les défauts de la base de données ne sont malheureusement pas tous des défauts permanents).

Nous pensons que cette bonne organisation des défauts permanents dans l'espace 2-D a toutes les chances de faciliter une classification visant à discriminer ce genre de défauts. Enfin, on peut remarquer que les mesures correspondant à la CDA sont proches de celles obtenues avec les données brutes. Ceci montre qu'assez peu d'information permettant de discriminer les défauts « permanents » a été perdue lors de la réduction de dimension par cette technique.

Ces conclusions gagneraient bien entendu à être confirmées par une opération de classification qui suivrait la phase de réduction de dimension. Cette classification n'a pas été entreprise dans ce travail car d'une part, son analyse risque fort d'être faussée et d'autre part, car elle n'a pas d'intérêt applicatif direct. Une expertise a en effet révélé que, malgré le nettoyage subit par la pièce, un certain nombre de faux défauts « polluant » la base de données C qui ne représente donc pas purement la classe des défauts permanents. D'autre part, comme nous allons le voir, la dimension intrinsèque des données manipulées est bien plus grande que 2. Une réduction de dimension vers un espace bidimensionnel n'est donc pas adaptée ; par conséquent, il est probable qu'une telle classification donnerait des résultats loin d'être optimaux.

Cependant, nous souhaitons tout de même étudier les répercussions directes de la réduction de dimension sur des traitements ultérieurs. Nous avons donc analysé l'impact des trois techniques précédemment étudiées sur les performances d'un classifieur neuronal lors d'une tâche de catégorisation réelle. Cela implique ainsi une mise en œuvre plus « intégrée » de la réduction de dimension et permet de valider l'ensemble de la chaîne de traitement proposée.

III.2.4.2 Classification neuronale : discrimination des défauts de type « poussière »

Le problème considéré ici est celui de la discrimination des défauts de type « poussière » parmi tous les types de défauts rencontrés sur les pièces optiques. Il a tout d'abord l'avantage d'être représentatif de la complexité des classifications envisagées pour le système de contrôle automatique visé. Par ailleurs, les poussières constituant un type de défauts très répandu et nécessitant un traitement distinct, il présente un intérêt applicatif évident.

III.2.4.2.1 Prologue : dimension intrinsèque des données

Avant de s'intéresser à la réduction de dimension proprement dite, une étape préliminaire consiste à déterminer, au moins approximativement, la dimension intrinsèque des données considérées. En effet, un des paramètres des techniques CCA et CDA est la dimension de l'espace d'arrivée où l'on souhaite projeter les données de haute dimension. Or nous avons vu précédemment que, pour éviter une perte d'information notable, la dimension finale des données doit être en accord avec leur dimension intrinsèque. En ce qui concerne SOM, le problème ne se pose pas puisque nous sommes limités, pour des raisons de complexité algorithmique, à des espaces d'arrivée de dimension 2.

Pour estimer la dimension intrinsèque des données manipulées (la base de données B qui servira pour l'apprentissage du classifieur), nous avons utilisé la méthode de Grassberger-Procaccia (paragraphe III.2.1, p.104). Le graphe obtenu et la régression linéaire correspondante sont représentés dans la Figure III.22.

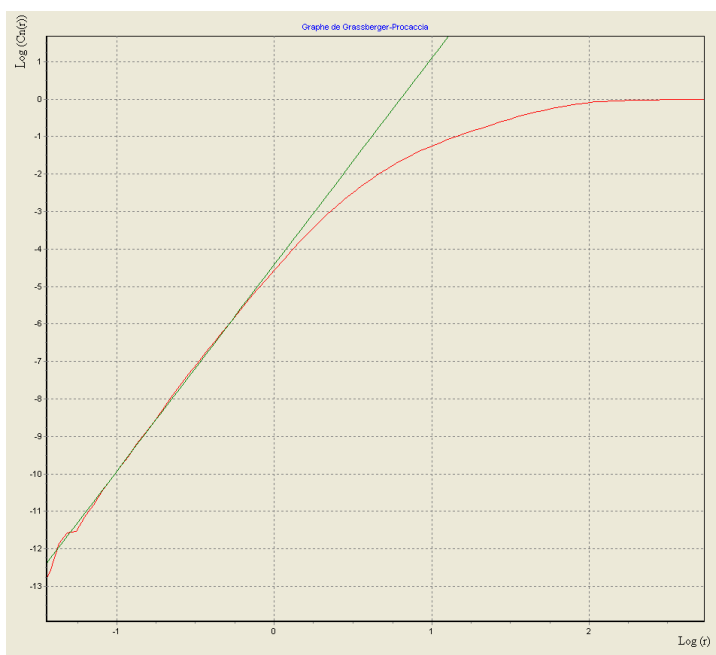


Figure III.22 : Estimation de la dimension intrinsèque de la distribution de données de la base B : $DI = 5,50$.

Afin de limiter la sous-évaluation de la dimension intrinsèque inhérente à cette méthode, nous avons utilisé la calibration proposée dans (Camastra and Vinciarelli 2002) (paragraphe III.2.1, p.104) ; la Figure III.23 montre les points mesurés et la courbe interpolée.

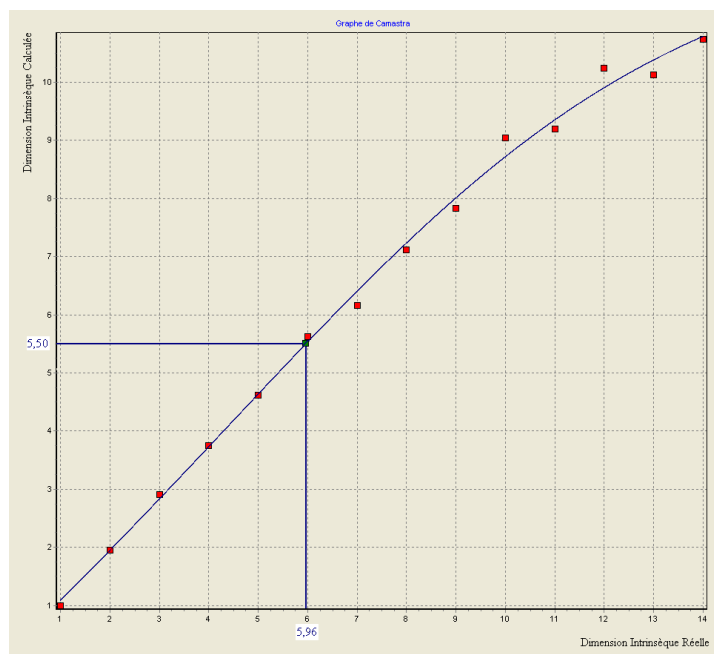


Figure III.23 : Estimation de la dimension intrinsèque de la Figure III.22 corrigée grâce à une calibration.

Finalement, la valeur corrigée obtenue est de 5,96. A titre de comparaison, une évaluation de la dimension intrinsèque des données a été effectuée à l'aide des méthodes MLE (Levina and Bickel 2007) et Geodesic Entropic Graphs (Costa and Hero 2004), ce qui a conduit respectivement aux valeurs 6,92 et 5,14. Rappelons ici que la technique exposée dans (Costa and Hero 2004) est de nature statistique ; le résultat présenté ici provient de 100 évaluations moyennées (on peut noter d'ailleurs l'écart-type assez important de ces évaluations, 1,61 ; nous reviendrons sur ce point dans le paragraphe III.3.3.1).

Ces différentes estimations ne sont donc pas tout à fait concordantes. Nous avons par conséquent choisi de tester plusieurs dimensions d'arrivée de la phase de réduction de dimension ; ceci rejoint par ailleurs les recommandations faites dans (Verleysen 2001), visant à optimiser la réduction de dimension en termes de respect de topologie des données.

III.2.4.2.2 Protocole expérimental

Tout d'abord, les imageries des bases de données A et B (Tableau III.2) ont été étiquetées par un expert avec les labels P : « défauts de type poussière », et NP : « défauts d'un autre type ». Cette tâche est délicate puisque certaines poussières peuvent être facilement confondues avec des défauts de type « piquêre ».

Ce fait, illustré par la Figure III.24, explique que la classification visée est assez ardue.

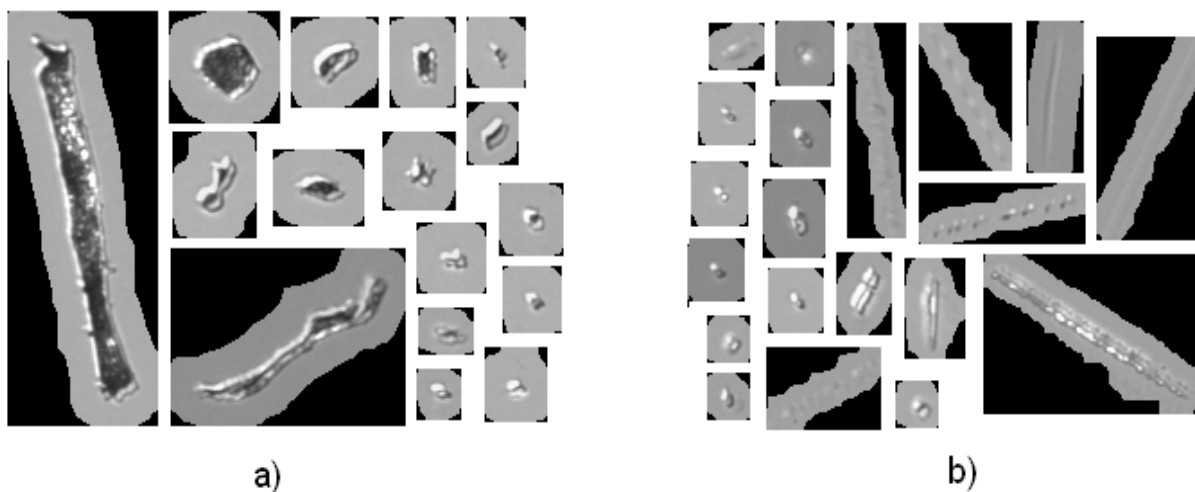


Figure III.24 : Exemples de défauts à classer : a) défaut de type « poussière » ; b) autre type de défauts .

La répartition des imagettes des deux bases de données entre les classes P et NP est donnée dans le Tableau III.4.

| Base de données | Issue de l'expérience | Nombre total d'éléments | Nombre d'éléments de la classe P | Nombre d'éléments de la classe NP |
|-----------------|-----------------------|-------------------------|----------------------------------|-----------------------------------|
| 1 | B | 1910 | 184 | 1726 |
| 2 | A | 3865 | 275 | 3590 |

Table III.4 : Description des bases de données utilisées pour l'expérience.

Toutes les imagettes ont ensuite été codées sous forme de vecteurs à 13 composantes en utilisant la transformation de Fourier-Mellin. Ces vecteurs ont finalement été utilisés pour une série d'expériences, dont le principe commun est expliqué par la Figure III.25. Pour ces expérimentations, la classification est réalisée à l'aide d'un réseau de neurones artificiels de type Multi Layer Perceptron (MLP).

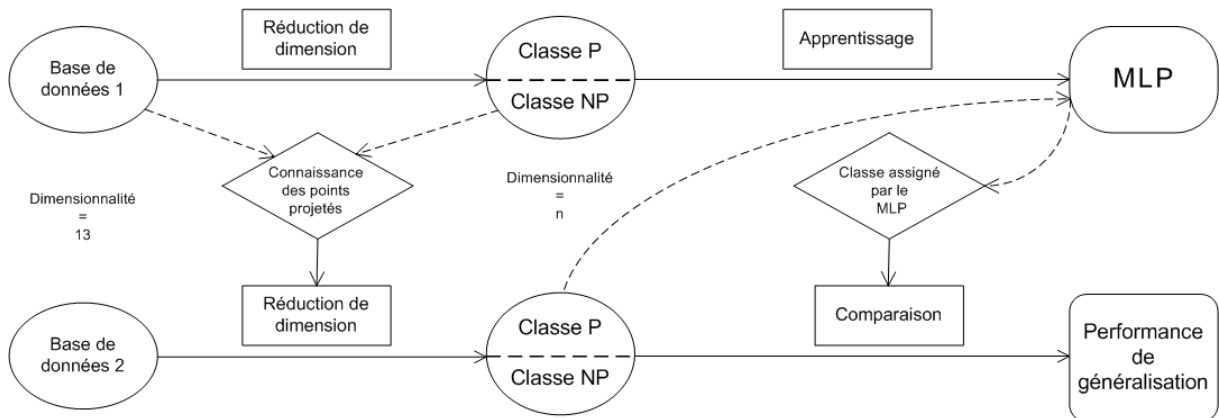


Figure III.25 : Principe des expériences menées. En haut, phase d'apprentissage, en bas, phase de test.

Chacune des expériences est composée de deux phases :

La phase d'apprentissage :

- Les éléments de la base de données 1 subissent une opération de réduction de dimensionnalité vers un espace de dimension n .
- Les représentations en dimension n servent à l'apprentissage d'un réseau de neurones de type MLP. Le but est d'apprendre le code correspondant à la classe associée au défaut présenté en entrée.

La phase de test :

- Les éléments de la base de données 2 subissent une opération de réduction de dimensionnalité vers un espace de dimension n . Pour cela on utilise la connaissance des éléments de la base de données 1 et de leurs projections associées (obtenues lors de l'apprentissage).
- Les représentations en dimension n sont présentées en entrée du MLP. Pour chaque exemple, on détermine la classe réponse du MLP en formant la différence des valeurs fournies par les deux neurones de sortie (si elle est négative le MLP a désigné la classe P).
- En comparant, pour chaque exemple, la classe réponse du MLP à la classe réelle, on détermine finalement un score de généralisation pour le MLP.

Ces expériences mettent en œuvre un réseau de neurones artificiels de type MLP avec n neurones d'entrée, 35 neurones dans une seule couche cachée et deux neurones de sortie. Les n neurones de la couche d'entrée correspondent aux n composantes des vecteurs obtenus après la phase de réduction de dimension. Les deux neurones de sortie sont quand à eux utilisés pour désigner la classe du défaut correspondant aux paramètres d'entrée du réseau selon le codage « grand-mère » : la sortie 0-1 désigne la classe P et la sortie 1-0 la classe NP. Le nombre de neurones de la couche cachée (35) a été fixé après une étude expérimentale comme un premier compromis entre performance et complexité. Enfin, tous les neurones utilisent une fonction d'activation de type sigmoïde (émulée par la fonction tangente hyperbolique). L'apprentissage de ce réseau de neurones est effectué grâce à l'algorithme BFGS (Broyden, Fletcher, Goldfarb, et Shanno) (Dennis and Schnabel 1983) auquel est associé une régularisation Bayésienne (Foresee and Hagen 1997)(MacKay 1992) afin de limiter les problèmes de surapprentissage. Pour ces expérimentations, l'implémentation des MLPs de la Neural Network Toolbox de Matlab®© a été mise en œuvre. Par ailleurs, on a utilisé les paramètres d'apprentissage proposés par défaut par le logiciel, ainsi que l'option de centrage-réduction préalable des données, qui permet d'éviter certains problèmes numériques.

En suivant ce principe commun, 24 expériences différentes ont été menées. Pour chacune de ces expériences, la procédure illustrée par la Figure III.25 a été appliquée 5 fois, afin d'obtenir des informations statistiques sur les performances de généralisation (une moyenne et un écart-type). Ces 24 expériences peuvent être classées en quatre groupes distincts :

- 11 expériences où la réduction de dimension est effectuée en utilisant la technique CCA ; la dimension n finale varie de 2 à 12 par pas de 1.
- 11 expériences où la réduction de dimension est effectuée en utilisant la technique CDA ; la dimension n finale varie de 2 à 12 par pas de 1.
- 1 expérience où la réduction de dimension est effectuée en utilisant la SOM ; la dimension n finale est 2.
- 1 expérience sans réduction de dimension ; on utilise directement les données brutes ($n = 13$).

Pour la réduction de dimension par SOM, nous avons utilisé la grille de la Figure III.2 (obtenue dans le paragraphe III.1.2.2.3) qui ne provient pas des mêmes données de base (qui avaient été centrées-réduites). Cependant, les expériences menées dans (Demartines and Blayo 1992) montrent que cette transformation n'a pas ou peu d'influence sur le comportement des SOMs pour une dimension assez grande (ce qui est notre cas) ; on peut donc légitimement considérer que les résultats obtenus ici sont équivalents à ceux qui auraient été obtenus en utilisant les données brutes.

III.2.4.2.3 Résultats et discussion

Avant d'aborder les résultats proprement dits, il faut relever quelques points importants. Premièrement, les données des bases A et B ne proviennent pas de mesures impliquant la même pièce ; les deux composants optiques étudiés n'étaient par ailleurs pas constitués du même matériau. Les performances de généralisation présentées sont donc significatives. Ceci est d'autant plus vrai que la base qui a servi à la phase de test contient à peu près deux fois plus d'exemples que la base d'apprentissage.

D'autre part, il est intéressant dans notre cas d'étudier les performances de classification pour les classes P et NP séparément, en plus des performances globales. En effet, la classe P est très sous-représentée par rapport à la classe NP (environ 10 fois moins d'exemples) dans les bases de données ; ceci donne un poids largement prépondérant à cette dernière classe dans les performances globales.

La Figure III.26 présente les résultats des différentes expériences menées en termes de performances globales de classification (c'est-à-dire le pourcentage d'exemples de la base de test correctement classés). Chacune de ces performances est obtenue en moyennant les résultats des 5 expériences correspondantes.

Les Figures III.27 et III.28 présentent les résultats en termes de performances de reconnaissance (c'est-à-dire le pourcentage d'exemples de la base de test d'une classe donnée reconnus) respectivement pour les classes NP et P. Sur ces deux figures sont également représentés les écarts-types associés, chacun d'eux étant calculé sur l'ensemble de 5 expériences mises en oeuvre.

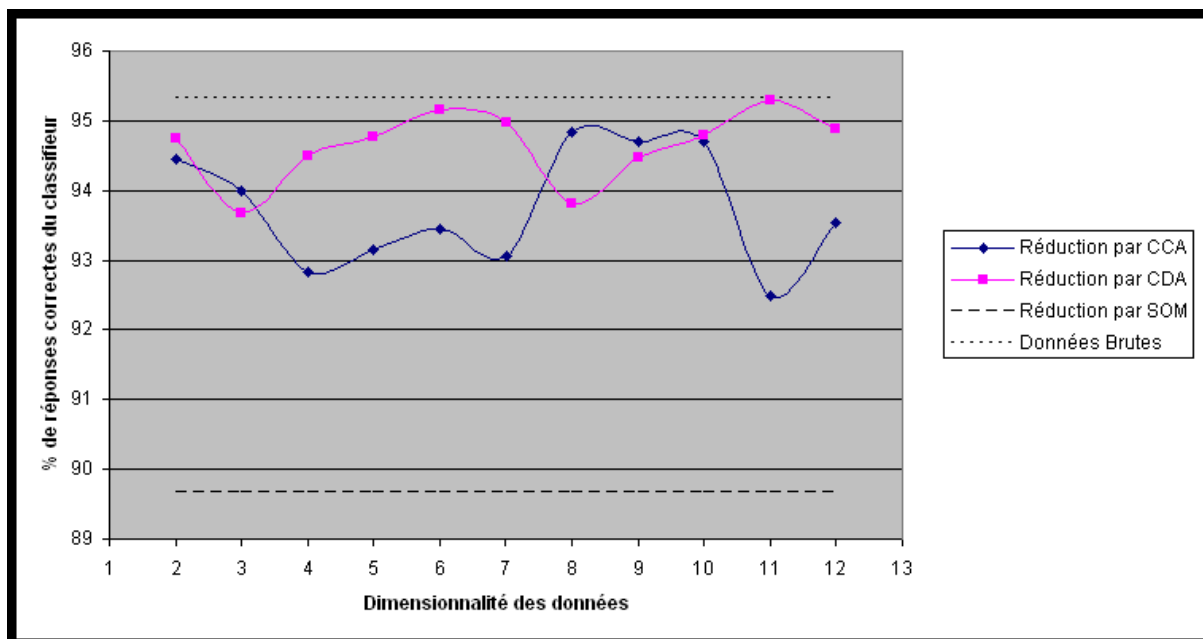


Figure III.26 : Performances globales de classification lors des différentes expériences.

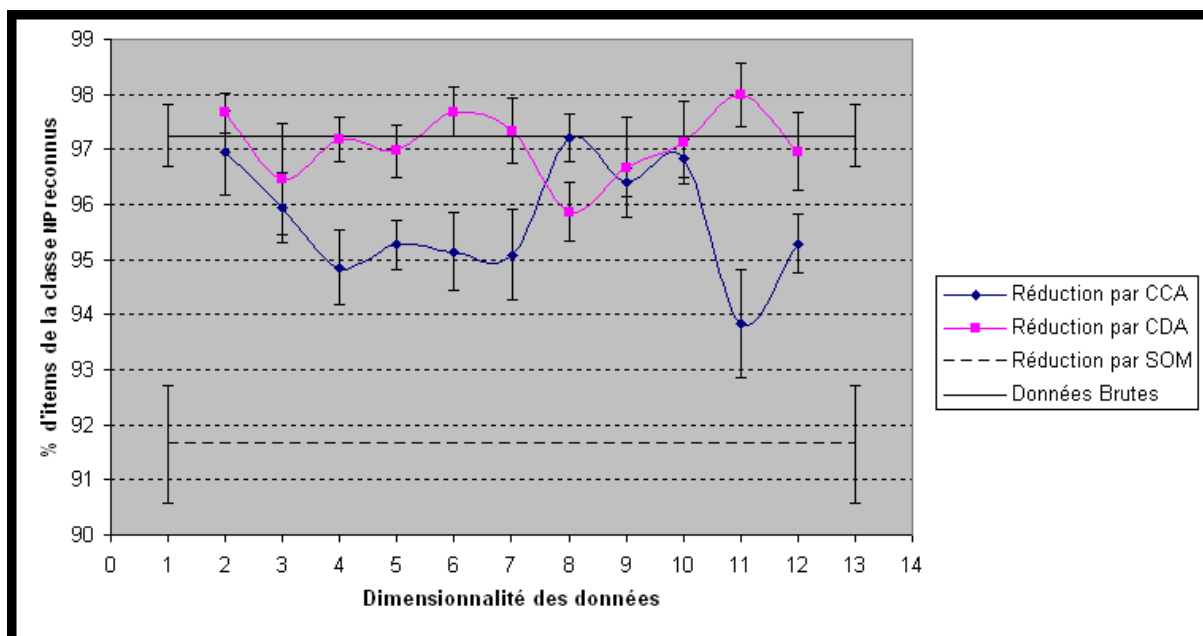


Figure III.27 : Performances de reconnaissance de la classe NP (« autres défauts »); l'écart-type relatif aux 5 expériences est également représenté.

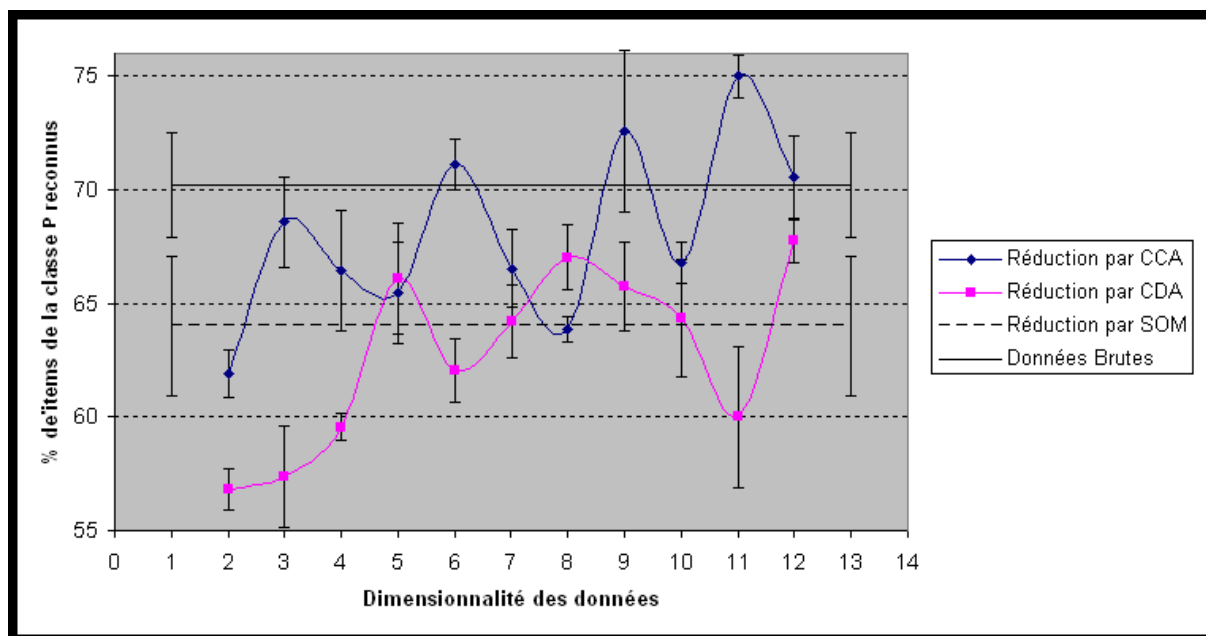


Figure III.28 : Performances de reconnaissance de la classe P (« défauts de type poussière ») ; l'écart-type relatif aux 5 expériences est également représenté.

Tout d'abord, il faut souligner, dans toutes les expériences, l'écart entre les performances de reconnaissance de la classe P (de l'ordre de 70 %) et de la classe NP (de l'ordre de 95 %). Cet écart est dû à la structure de la base de données utilisée dans la phase d'apprentissage, qui défavorise la classe P largement sous-représentée et donc moins bien « apprise ». Cette hypothèse se trouve confirmée par l'examen des écarts-types des taux de reconnaissance : ils sont très faibles (de l'ordre de 0.5 %) pour la classe NP, ce qui est le signe d'un apprentissage de qualité ; pour la classe P, ils sont largement plus élevés (environ 2 %) ce qui traduit un apprentissage plus délicat.

De plus, en étudiant les résultats associés à une réduction de dimension vers un espace de dimension 2, on peut comparer l'effet des trois techniques précédemment étudiées (CCA, CDA et SOM) sur les performances du classifieur. Ces résultats sont reportés dans le Tableau III.5. Si la réduction de dimension par SOM permet d'atteindre un score légèrement plus élevé pour la reconnaissance de la classe P que l'utilisation de la CCA, elle donne des résultats moyens pour la classe NP. La CDA conduit quant à elle aux meilleurs résultats pour la classe NP au détriment des scores de reconnaissance de la classe P. Au final, les performances atteintes en utilisant la CCA et la CDA sont à peu près équivalentes et surpassent les scores correspondant à l'utilisation d'une SOM. Cela est

sans doute dû à la meilleure qualité des projections obtenues à l'aide de ces techniques ainsi qu'à leur nature continue. Par ailleurs, ces résultats sont en accord avec les conclusions de la paragraphe précédente.

| Technique de réduction de dimension | SOM | CCA | CDA |
|---|---------------|---------------|---------------|
| Taux de reconnaissance de la classe P | 63.9 % | 61.9 % | 56.8 % |
| Taux de reconnaissance de la classe NP | 91.6 % | 96.9 % | 97.7 % |
| Performance globale | 89.7 % | 94.4 % | 94.8 % |

Table III.5 : Performances du classifieur pour les différentes techniques de réduction de dimensionnalité, pour une dimension finale de 2.

La comparaison des performances après réduction de dimension avec celles obtenues en utilisant les données brutes est également intéressante. On peut en effet atteindre des niveaux de performances comparables en utilisant une représentation des données plus compacte plutôt que la représentation exhaustive initiale. Ceci est illustré par les résultats relevés dans le Tableau III.6 qui met en avant certains résultats particulièrement intéressants.

| Technique de réduction de dimension | Aucune | CCA | CDA |
|---|---------------|---------------|---------------|
| Dimensionnalité finale des données | 13 | 9 | 5 |
| Taux de reconnaissance de la classe P | 70.2 % | 72.6 % | 66.1 % |
| Taux de reconnaissance de la classe NP | 97.2 % | 96.4 % | 97.0 % |
| Performance globale | 95.3 % | 94.7 % | 94.8 % |

Table III.6 : Performances « remarquables » atteintes par le classifieur pour les techniques de réduction de dimensionnalité CCA et CDA.

L'utilisation d'une phase de réduction de dimension permet ainsi de limiter la complexité du classifieur utilisé (puisque le nombre d'entrées du réseau de neurones est plus petit) et donc d'économiser du temps de calcul, tout en conservant le niveau de performance. Cette conclusion doit être cependant largement nuancée, comme nous le verrons dans la partie conclusive de ce paragraphe.

D'autre part, le taux de reconnaissance de la classe P (qui est celui qui subit de loin les plus fortes variations) est relativement bas lorsque l'on réduit la dimensionnalité des données à 2 : la représentation obtenue n'amène probablement pas assez d'information pour discriminer correctement les défauts de type poussière. Cependant, dans le cas de la réduction par CDA, ce taux augmente régulièrement jusqu'à atteindre une valeur proche de la valeur référence obtenue avec les données brutes, pour une dimension finale de 5. Ce résultat est en accord avec l'estimation de la dimension intrinsèque de l'ordre de 6, qui a été donnée plus haut. Ce niveau de performance n'est atteint de nouveau que pour des dimensions de l'espace d'arrivée (après réduction de dimension) de 8, 9, et 12 : la courbe d'évolution des scores présente de curieuses « ondulations ». Ce phénomène, qui sera analysé en détail et interprété dans le paragraphe III.3.3.1 est encore plus flagrant dans le cas d'une réduction par CCA.

Finalement, les performances obtenues, environ 70 % de défauts de type poussière et 97 % d'autres défauts correctement reconnus est plutôt satisfaisant pour l'application visée. Les 3% d'erreur reportée dans ce dernier cas peuvent cependant être problématique puisque tous les défauts permanents doivent être reportés. Il peut être également handicapant de classer un trop grand nombre de poussières comme autres défauts puisque l'on risque au final d'écarter une pièce en réalité conforme. Ces différents risques devraient être évalués de manière plus approfondie afin de fixer la dimension d'arrivée de la phase de réduction de dimensionnalité qui contrôle en partie le compromis entre les taux de reconnaissance des classes P et NP.

Cette série d'expériences a prouvé, conformément aux résultats de la paragraphe précédente, que les projections issues de la CCA ou de la CDA menaient à des représentations compactes plus adaptées à une classification que celles obtenues par le biais d'une SOM. Par ailleurs, elle a montré que l'on pouvait parvenir à des performances équivalentes en termes de taux de reconnaissance des classes mises en jeu en utilisant de telles représen-

tations plutôt que les données brutes.

Ces expérimentations revêtent également un intérêt applicatif direct puisque le problème de classification abordé, la discrimination des défauts de type poussière, est capital pour le système à concevoir. Dans cet optique, les performances obtenues en termes de généralisation sont satisfaisantes puisque environ 95 % des défauts de la base de test sont classés correctement. D'autre part, l'utilisation de techniques de réduction de dimension paraît séduisante puisqu'elle permet de simplifier l'architecture du classifieur et ainsi de réduire son temps de réponse. Dans le cadre qui est le notre, ceci peut être déterminant pour diagnostiquer les très nombreux défauts (permanents ou non) détectés sur une pièce optique dans un temps compatible avec les contraintes industrielles. Malheureusement, si le temps de réponse du classifieur est grandement amélioré, il faut prendre également en compte le temps de calcul nécessaire à la projection des données dans le nouvel espace de représentation. L'algorithme de projection associé à CCA et à CDA est itératif (paragraphe III.2.3.2.3) et nécessite le calcul de nombreuses distances euclidiennes. Même si cet algorithme peut être simplifié, les calculs associés compensent amplement le temps gagné sur la classification proprement dite...

L'autre intérêt majeur de la réduction de dimension consiste à limiter les effets néfastes du « fléau de la dimensionnalité ». Il semble que ces effets ne soient pas très importants pour la dimension modérée (13) des données que nous avons manipulées, puisque la réduction de dimension n'a pas permis d'améliorer sensiblement la qualité du système de classification.

III.3 Classification neuronale modulaire

Si l'utilisation de techniques de réduction de dimension a donné des résultats intéressants d'un point de vue théorique et dans une moindre mesure d'un point de vue pratique, elle n'a pas permis d'améliorer de manière notable la qualité d'un classifieur à base de MLP, que ce soit en termes de complexité calculatoire (donc de temps de calcul) ou de performances brutes (taux de classification correcte). Afin d'obtenir des avancées significatives sur l'un ou l'autre de ces points (idéalement les deux!), nous avons mis en œuvre un traitement modulaire de type Multi Neural Network (MNN).

Plus précisément, nous nous sommes intéressés à un algorithme appelé T-DTS (Treelike Divide To Simplify) utilisant le paradigme « diviser pour régner ». L'idée est de diviser, de façon arborescente, le problème original en plusieurs sous-problèmes que l'on espère plus faciles à résoudre. La phase de décomposition revêt donc pour cette technique une importance primordiale. Or, les méthodes de quantification vectorielle utilisées à cet effet dans l'algorithme original nous paraissent souffrir d'une limitation importante que nous détaillerons plus loin.

Nous proposons dans ce paragraphe un nouveau mécanisme de décomposition basé sur les dimensions intrinsèques locales et la topologie des données. Le concept est validé sur un problème jouet puis sur les données réelles issues de notre problème industriel (les « imageries » de défauts).

III.3.1 Classification modulaire : définition et rapide état de l'art

Le concept de modularité s'oppose à celui du traitement de données par une seule entité : il conduit à utiliser un certain nombre de modules capables de traiter l'information et de communiquer entre eux (de manière partielle ou non) au sein d'une structure (plus ou moins clairement définie). L'hypothèse sous-jacente est qu'il est possible de faire émerger une solution de meilleure qualité à partir d'un groupe d'« experts » qu'en considérant un « expert » unique. Sa pertinence a été montrée dans de nombreuses applications mettant en jeu une classification (Hong et al. 2004)(Hing et al. 2004)(Harb 2003). Notons ici que, de manière générale, les modules du système n'ont pas lieu d'être de nature homogène ; dans le

cas d'un problème de discrimination, ces « experts » peuvent implémenter indifféremment une des techniques de classification classiques (arbres de décision, réseaux de neurones...). Cette idée maîtresse a été déclinée de nombreuses manières, donnant naissance à un ensemble de méthodes couvrant les domaines des comités de machines (Volker 2001), des systèmes multi-agent (Ferber 1998) et de l'intelligence artificielle distribuée (Avouris and Gasser 1992). En première approche, ces différentes méthodes peuvent être classées en deux groupes fondamentalement distincts :

- Les structures dont les modules travaillent sur l'ensemble de l'espace des données : le paramétrage de chaque module (apprentissage dans le cas de réseaux de neurones) est effectué à partir de données provenant indifféremment de toutes les régions de l'espace d'entrée du problème, même si tous les modules ne traitent pas nécessairement les mêmes données. Dans ce cas, une prise de décision résulte d'une combinaison (linéaire ou non) de l'avis de tous les experts. Ce type de méthodes a pour objectif principal d'améliorer les performances de généralisation, en évitant des erreurs qui pourraient être commises par chaque expert individuellement. Cette amélioration se fait cependant au détriment du temps de calcul puisqu'à chaque décision du système, l'ensemble des modules est consulté.

- Les structures dont les modules travaillent d'une manière locale (seulement sur une partie de l'espace des données) :

Le problème est tout d'abord découpé en un certain nombre de sous-problèmes, selon le principe « diviser pour régner » (« divide et impera », Jules César). A chaque sous-problème, correspondant à une région donnée de l'espace d'entrée du problème, est associé un module. Ainsi chaque module est paramétré en n'utilisant qu'une partie des données disponibles, ce qui amène généralement à des modèles assez simples. Lors d'une décision du système, un nombre très limité de modules est interrogé. Les principaux avantages de ce type de méthodes sont, d'une part la compréhensibilité (on dispose au final de modèles simples et locaux) et d'autre part, la complexité algorithmique moindre (de part la simplicité des modules mis en œuvre), qui permet un apprentissage et une exécution plus rapide. L'augmentation des performances de généralisation n'est en revanche pas la motivation principale, et reste hypothétique (Volker 2001).

Dans notre cas, le taux de reconnaissance étant déjà globalement satisfaisant, nous avons choisi de mettre en place une méthode provenant du deuxième groupe, afin de tenter d'améliorer les temps d'apprentissage et de classification du système

III.3.2 La technique Treelike Divide To Simplify (T-DTS)

Dans ce travail, nous sommes particulièrement intéressés à une technique baptisée Treelike Divide To Simplify (T-DTS) (Rybnik 2006). Comme son nom l'indique, il s'agit d'une approche arborescente : l'idée est de diviser récursivement le problème original pour obtenir un ensemble de sous-problèmes de tailles plus réduites. Cette décomposition est assurée par des unités de décomposition (UD) situées au nœud de l'arbre. Chacun des sous-ensembles de données résultants est ensuite traité par une unité de traitement (UT), qui constitue une feuille de l'arbre. Finalement, la structure d'un classifieur de type T-DTS est conforme à l'organigramme donné dans la Figure III.29.

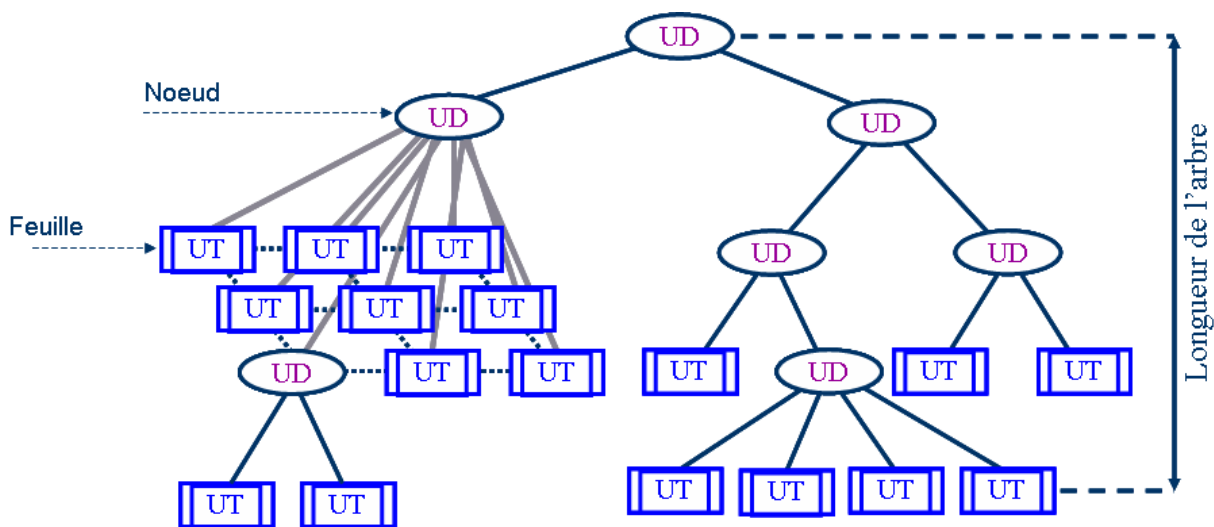


Figure III.29 : Structure générale d'un classifieur de type T-DTS (tiré de (Bouyoucef 2007)).

Décrivons maintenant succinctement les deux phases de fonctionnement de l'algorithme, l'apprentissage et l'exécution.

L'étape d'apprentissage débute par la génération automatique de la structure arborescente, à partir des données de la base d'apprentissage. On cherche en fait à décomposer cette base en plusieurs sous-ensembles de données : chacun de ces sous-ensembles définira un sous-problème de classification correspondant à une région particulière de l'espace

d'entrée du problème original. A chaque nœud de l'arbre en construction, une estimation de la complexité du sous-problème de classification local est calculée à l'aide d'une des nombreuses méthodes issues de la littérature (critère de Fischer, entropie collective, bornes de Battacharrya ...).

Si la complexité n'excède pas un seuil prédéfini, on considère que le problème est assez simple, on arrête donc la décomposition. On crée alors une unité de traitement dont le rôle est de modéliser ce sous-problème. Ces unités de traitements peuvent implémenter tous types d'algorithmes de classification supervisés ; dans ce travail nous ne considérerons cependant que des architectures mettant en œuvre des réseaux de neurones.

Dans le cas contraire, on crée une unité de décomposition afin de diviser l'ensemble de données disponibles au nœud en plusieurs sous-ensembles, associés ultérieurement à de nouveaux nœuds. Cette tâche est réalisée dans la version originale de T-DTS à l'aide de méthodes non-supervisées de quantification vectorielle (Competitive Network ou SOM (Kohonen 2001)).

Lors de la phase d'exécution, la donnée à classer est tout d'abord affectée à une des unités de traitement et une seule. Pour ce faire, on peut parcourir l'arbre de décomposition jusqu'à une de ses feuilles. Il est également possible de définir un prototype pour chaque sous-ensemble de données correspondant à une feuille. Il suffit alors d'attribuer la donnée sur la base d'une mesure de similitude (typiquement la distance euclidienne) avec ces prototypes : la donnée est assignée à l'UT dont le prototype lui « ressemble » le plus. Finalement on utilise le modèle implémenté par l'unité de traitement sélectionnée pour fournir une décision quant à la classe d'appartenance de la donnée considérée.

Dans (Madani, Chebira, and Rybnik 2003a) et (Madani, Chebira, and Rybnik 2003b), les auteurs démontrent que l'utilisation de cette structure se traduit par un gain de temps d'apprentissage, d'exécution et éventuellement de capacité de généralisation par rapport aux techniques neuronales conventionnelles. Dans (Bouyoucef 2007), les performances de généralisation de T-TDS sont comparées à celles obtenues par de nombreux algorithmes de classification classiques pour plusieurs problèmes réels issus de la base de données *UCI Repository of Machine Learning Databases and Domain Theorie*. Ces travaux montrent que la méthode T-DTS permet d'obtenir des classifieurs très compétitifs pour bon nombre de problèmes (taux de généralisation équivalents aux meilleurs classifieurs de la littérature).

On note cependant, pour certaines bases de données, une dégradation des performances par rapport à des techniques non modulaires.

On peut légitimement se demander pourquoi cette technique modulaire permet, pour certains problèmes, une amélioration notable des performances alors que pour d'autres, c'est l'inverse qui se produit. Il est alors utile de rappeler que les méthodes modulaires « locales » n'ont pas pour vocation première l'amélioration des performances de classification des systèmes mais leur simplification (même si ces deux aspects peuvent être liés). Pour nous, une éventuelle amélioration est entièrement conditionnée par la qualité de la phase de décomposition du problème. En d'autres termes, si la décomposition est satisfaisante, on obtient des sous-problèmes simples et l'on a de bonnes chances d'améliorer les performances en généralisation du système ; dans le cas contraire, on va les dégrader. On peut comprendre ce phénomène de la manière suivante. Par le mécanisme de décomposition, on découpe la base d'apprentissage en plusieurs sous-ensembles : on dispose donc de moins de données pour fixer les paramètres des experts locaux que lorsque l'on considère le problème global. Par conséquent, si le problème local n'est pas plus simple à résoudre, l'apprentissage sera de moins bonne qualité que pour le problème original.

Si l'on suit ce raisonnement, les baisses de performance constatées pour certains problèmes peuvent être expliquées par les faiblesses du mécanisme de décomposition. En effet, la version originale de T-DTS utilise au sein des unités de décomposition des techniques de quantification vectorielles. Ces unités ne modélisent donc que les densités de probabilité associées aux données sans prendre en compte la nature du problème de classification en soi. Il n'y a donc pas de raison a priori que ce type de décomposition génère des sous-problèmes simples.

Dans les faits, ce problème est limité dans T-DTS par le mécanisme de contrôle de la décomposition par l'estimation de la complexité des sous-problèmes. Il est probable en effet que certains des « découpages » de l'espace effectués soient profitables (même s'il n'y a pas de raison a priori). On peut en effet comprendre intuitivement qu'un nombre suffisant de décompositions finisse par isoler un sous-espace d'entrée où la classification est effectivement plus simple à effectuer. L'estimation de la complexité permet de détecter les découpages qui sont globalement bénéfiques (lorsque la complexité du problème baisse) et d'écartier les autres (dans ce cas on stoppe le processus de décomposition). Ce mécanisme

permet ainsi de favoriser la construction de sous-problèmes « simples ».

Cela dit, nous pensons que le mécanisme de décomposition est en l'état, non optimal. Nous proposons donc dans ce chapitre une nouvelle manière de décomposer les problèmes. Cette approche différente est fondée sur l'analyse de la structure topologique des données.

III.3.3 Technique de décomposition proposée

III.3.3.1 Fondement principal

La technique de décomposition que nous proposons s'appuie sur une hypothèse fondamentale : l'existence au sein des distributions de points issues de données réelles (un point correspondant à une donnée), de sous-distributions de dimensions intrinsèques différentes. En d'autres termes, on suppose qu'il existe des groupes de points présentant une dimension intrinsèque homogène, au sein d'une distribution dont cette caractéristique est globalement hétérogène.

Certaines observations semblent confirmer la véracité de cette hypothèse pour les données réelles manipulées dans les paragraphes III.1.2.2.2 et III.2.4.2.2. Premièrement, on peut comparer l'allure globale des graphes de Grassberger-Procaccia associés respectivement à notre distribution réelle (correspondant à la base de données B présentée dans le paragraphe III.1.2.2.2) et à une distribution de points de dimension intrinsèque fixée. Pour ce faire, nous avons construit un ensemble de 1910 points de dimension brute 13 (mêmes caractéristiques que la base de données B). Les 6 premières composantes des vecteurs associés aux points ont été fixées aléatoirement et les 7 autres ont été calculées comme des combinaisons linéaires et non-linéaires de ces composantes. On obtient par ce biais une distribution de points de dimension intrinsèque 6 (la dimension intrinsèque globale associée à la base de données) ; de plus, cette caractéristique est homogène pour l'ensemble des données. Les courbes obtenues par la méthode de Grassberger-Procaccia sont données dans la Figure III.30.

Bien que leurs portions linéaires présentent à peu près le même coefficient directeur (5,50 et 5,28), ces deux graphes sont dissemblables dans les parties correspondant à des distances relativement petites. On peut donc supposer que, bien qu'ayant des dimensions intrinsèques globales similaires, les deux distributions sont bien différentes localement, ce qui pourrait être expliqué par la présence d'un « mélange » de distributions de dimensions

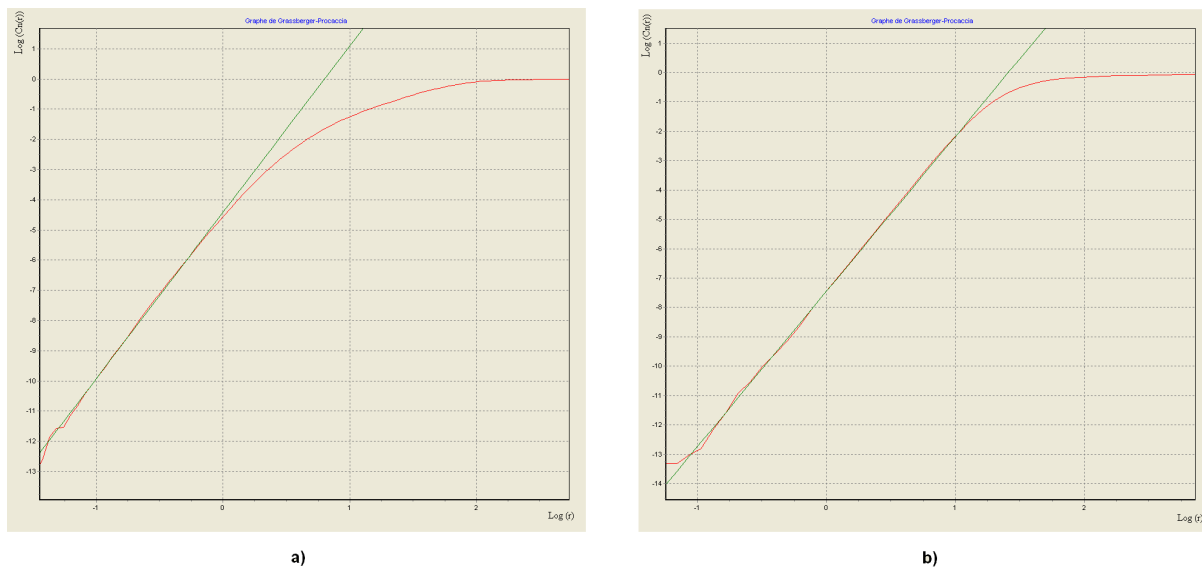


Figure III.30 : Les courbes de Grassberger-Proccacia obtenues pour a) les données réelles ; b) les données synthétiques.

intrinsèques distinctes au sein des données réelles. Ce phénomène pourrait également expliquer l'écart-type important des estimations statistiques de dimension par la méthode de (Costa and Hero 2004) (voir paragraphe III.2.1, p.130).

Examinons maintenant les résultats obtenus dans le paragraphe III.2.4.2.3, rappelés dans la Figure III.31, concernant les performances de reconnaissance pour la classe P du problème réel, en fonction de la dimension des données après la réduction de dimensionnalité par CCA et CDA.

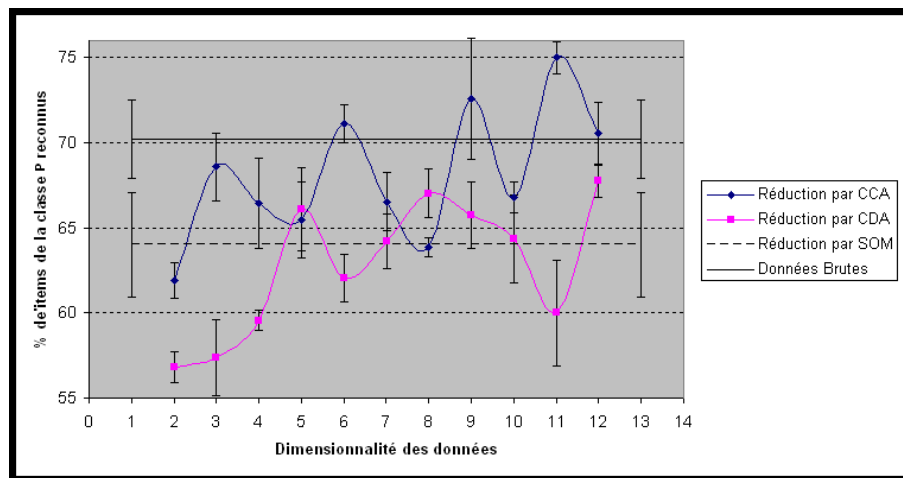


Figure III.31 : Performances de reconnaissance de la classe P ; l'écart-type relatif aux 5 expériences est également représenté.

Les taux de reconnaissance présentent d'importantes fluctuations, ce qui est normal puisque pour être efficace, une réduction de dimension doit se faire vers un espace d'une dimension compatible avec la dimension intrinsèque des données. Plus singulier, ces taux présentent des maximas (que les écart-types faibles nous permettent de penser significatifs) pour plusieurs dimensions d'arrivée différentes. Cela semble confirmer la présence du mélange évoqué plus haut.

Pour finir, l'existence de plusieurs sous-espaces de dimensions intrinsèques différentes au sein d'une même distribution issue d'un problème réel a déjà été relevée dans la littérature à plusieurs reprises (Raich, Costa, and Hero 2006) (Haro, Randall, and Sapiro 2006)(Barbarà and Chen 2007).

De plus, on peut fournir une interprétation de la dimension intrinsèque d'une distribution : elle correspond au nombre de degré de liberté du processus physique qui a engendré les données. L'existence de plusieurs sous-distributions de dimensions intrinsèques différentes n'est donc pas étonnante ; elles correspondent à des données issues de processus physiques distincts. Ceci nous amène à l'idée directrice sous-tendant le mécanisme de décomposition proposé dans ce travail : on va chercher à isoler des groupes de dimension intrinsèque homogène pour les traiter séparément.

D'une part, un tel processus de décomposition de problème est guidé par un critère précis. D'autre part, on peut légitimement penser qu'il est pertinent de traiter séparément des données engendrées par des processus physiques différents. Il est en effet probable que les sous-problèmes ainsi définis, « homogènes » d'un point de vue physique, soient plus faciles à traiter que le problème initial. On peut également espérer que certains des groupes obtenus soient composés exclusivement de points correspondant à des données d'une classe unique (ce qui simplifie considérablement la classification !); des items de classes différentes proviennent de processus physiques distincts, les groupes correspondant ne doivent donc pas présenter la même dimensionnalité. Cette affirmation doit toutefois être nuancée à cause de la phase de représentation du problème (extraction des paramètres observés) qui peut gommer une partie importante de la variabilité du système.

La technique de décomposition que nous proposons dans ce travail se compose de deux phases distinctes. La première consiste à estimer localement la dimension intrinsèque de la distribution en chacun de ses points. La deuxième regroupe, de proche en proche,

les points qui sont voisins d'un point de vue topologique et auxquels sont associées des dimensions intrinsèques locales comparables. Le résultat de cette dernière étape est un ensemble de groupes de points, dont les dimensions intrinsèques peuvent être considérées comme homogènes.

III.3.3.2 Estimation locale de la dimension intrinsèque

Nous souhaitons former des groupes de points présentant des dimensions intrinsèques homogènes, ce qui suppose implicitement que l'on connaisse localement cette caractéristique. Or, les méthodes présentées précédemment (paragraphe III.2.1) permettent quasiment toutes d'accéder seulement à une estimation globale de la dimension intrinsèque d'une distribution. Les rares techniques qui proposent une information locale ne sont pas très robustes et souvent fortement biaisées, ce qui explique que ces estimateurs locaux soient moyennés sur l'ensemble de la distribution de points afin de donner une estimation globale d'une fiabilité satisfaisante.

Dans la littérature, la seule technique locale d'une robustesse suffisante est présentée dans (Costa, Girotra, and Hero 2005). Malheureusement, cette technique, dérivée de la méthode présentée par le même auteur dans (Costa and Hero 2004) utilise une estimation statistique par bootstrap et donc n'est pas déterministe, ce qui interdit toute expérience reproductible. Pour s'affranchir de cette limitation, nous avons adapté la méthode globale MLE (Levina and Bickel 2004) en suivant l'idée présentée dans (Costa, Girotra, and Hero 2005).

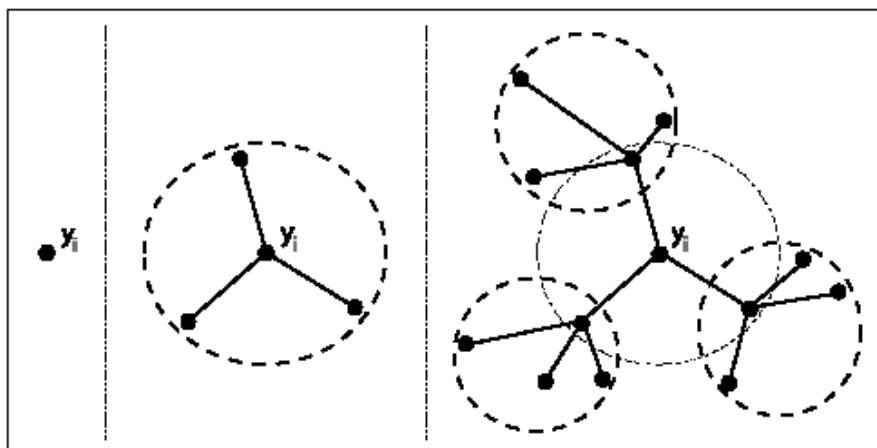


Figure III.32 : Construction du voisinage du point y_i , avec $k=3$

L'estimation de la dimension intrinsèque locale en un point donné d'une distribution se fait alors en deux temps. On construit tout d'abord un voisinage local du point en considérant itérativement ses k plus proches voisins ; ce mécanisme tiré de (Costa, Girotra, and Hero 2005) est illustré par la Figure III.32, tirée de (Costa, Girotra, and Hero 2005). On applique ensuite la méthode globale MLE (en lieu et place de la méthode de (Costa and Hero 2004)) sur l'ensemble de points obtenu.

Au final, l'algorithme de calcul des dimensions intrinsèques locales se présente comme ceci (Algorithme 1) :

Paramètres d'entrée :

- P : la distribution de points à traiter,
- $Iteration$: nombre d'itération (typiquement 3),
- NN : nombre de points considérés comme « plus proches voisins » (typiquement 5).

Paramètre de sortie :

- $DimensionIntrinsequeLocale$: Ensemble des dimensions intrinsèques locales calculées pour les points de P .

Algorithme III.1 : Algorithme de calcul des dimensions intrinsèques locales

```

pour chaque point  $x$  de  $P$  faire
    % Construire un voisinage du point considéré ;
     $V \leftarrow \{x\}$  ;
    pour  $i \leftarrow 1$  à  $Iteration$  faire
        pour chaque point  $y \in P$  faire
            | Calculer l'ensemble  $V_y$  des  $NN$  plus proches voisins de  $y$  ;
        fin
         $V \leftarrow V \cup V_y$ ;
    fin
    % Calculer la dimension intrinsèque locale au point considéré;
     $DimensionIntrinsequeLocale(x) \leftarrow MLE(V)$  ;
    % MLE est directement tirée de (Levina and Bickel 2004) ;
fin

```

III.3.3.3 Formation de groupes de dimensions intrinsèques homogènes

Par cette opération, nous souhaitons former des groupes de points, voisins d'un point de vue topologique et auxquels sont associées des dimensions intrinsèques locales comparables. Pour ce faire, on commence par construire un graphe « géodésique » recouvrant les données. Ce graphe, liant chaque point à ses k plus proches voisins (classiquement $k = 5$) permet de modéliser la topologie des données. On va donc chercher à regrouper des points connexes (au sens de ce graphe), respectant ainsi la topographie de la distribution.

Toutes les données sont attribuées à un sous-ensemble et un seul, lors de plusieurs phases de « déploiement » initiées à partir d'un point « graine » choisi aléatoirement. Un déploiement consiste à agréger itérativement les plus proches voisins des points d'un groupe, à la condition qu'ils présentent une dimension intrinsèque locale compatible avec celle du groupe. Cette notion de compatibilité fait intervenir un seuil de tolérance qui définit un intervalle où les dimensions intrinsèques sont considérées comme semblables (il est en effet illusoire de penser trouver des dimensions intrinsèques locales rigoureusement égales).

En fait, on effectue de cette manière des découpages de la distribution en sous-ensembles pour des tolérances de plus en plus grandes et en recalculant à chaque itération la dimension intrinsèque locale associée à un point comme la dimension intrinsèque moyenne du sous-ensemble auquel il appartient. Ce mécanisme permet d'éviter d'incorporer à un groupe des données de plus en plus éloignées en termes de dimension intrinsèque locale. Au final, l'algorithme de construction des sous-ensembles de points se présente comme ceci (Algorithme III.2 et III.3) :

Paramètres d'entrée :

- P : la distribution de points à traiter,
- $MaxTolerance$: Tolérance maximum sur la dimension intrinsèque,
- Pas : Pas de l'algorithme (augmentation de la tolérance sur la dimensionnalité locale),
- NN : Nombre de points considérés comme « plus proches voisins ».

Paramètre de sortie :

- $Labels$: Ensemble des labels définissant l'appartenance des points de P aux différents groupes obtenus.

Algorithme III.2 : Algorithme de construction des sous-ensembles de dimension intrinsèque homogène

```

% Initialisation ;
pour chaque point  $x$  de  $P$  faire
    Labels( $x$ )  $\leftarrow$  0;
    Dimensions( $x$ )  $\leftarrow$  DimensionIntrinsequeLocale( $x$ ) ;
    Calculer l'ensemble  $V_x$  des  $NN$  plus proches voisins de  $x$  ;
    Voisinages( $x$ )  $\leftarrow$   $V_x$  ;
fin
% Construire les groupes désirés ;
pour tolerance  $\leftarrow$  Pas à MaxTolerance par pas de Pas faire
    ID  $\leftarrow$  0 ;
    % Construction de groupes temporaires ;
    pour chaque point  $x$  de  $P$  faire
        si Labels( $x$ ) = 0 alors
            ID  $\leftarrow$  ID + 1 ;
            PropagerID ( $x$ , ID, Dimension( $x$ ), tolerance) ;
        fin
    fin
    % Calcul de la dimensionnalité moyenne de chaque groupe
    NbPoints = Tableau[1..ID] ; Dims = Tableau[1..ID] ;
    pour chaque point  $x$  de  $P$  faire
        NbPoints(Labels( $x$ ))  $\leftarrow$  NbPoints(Labels( $x$ )) + 1 ;
        Dims(Labels( $x$ ))  $\leftarrow$  Dims(Labels( $x$ )) + Dimensions( $x$ ) ;
    fin
    pour chaque point  $x$  de  $P$  faire
        Dimensions( $x$ )  $\leftarrow$  Dims(Labels( $x$ ))/NbPoints(Labels( $x$ )) ;
        si tolerance < MaxTolerance alors
            Labels( $x$ )  $\leftarrow$  0 ;
        fin
    fin
fin

```

Algorithme III.3 : Fonction itérative PropagerIDFonction PropagerID (*point*, *ID*, *dimension*, *tolerance*)

% Propage récursivement un ID si le critère de tolérance est satisfait

si (*Labels*(*point*) = 0)et (*dimension* – *tolerance* < *Dimensions*(*point*))et (*dimension* + *tolerance* > *Dimensions*(*point*)) **alors**| *Labels*(*point*) ← *ID* ;| **pour chaque** *point* *x* de *Voisinages*(*point*) **faire**| | PropagerID (*x*, *ID*, *dimension*, *tolerance*) ;| **fin****fin**

Remarquons tout d'abord que les sous-ensembles de points résultants de cet algorithme ne dépendent pas de l'ordre dans lequel sont arrangées les données au départ. D'autre part, il peut arriver que des points se trouvent « orphelins » (ils font partie d'un groupe ne contenant qu'eux même), ou que des sous-ensembles soient constitués de peu de données. Ceci est préjudiciable dans la perspective d'un traitement ultérieur. On attribue alors ces points au groupe de taille acceptable le plus proche au sens de la distance euclidienne.

III.3.4 Validation de la méthode sur un problème « jouet »

Afin de vérifier l'efficacité de l'algorithme proposé, nous avons étudié un problème synthétique simple inspiré de (Martinetz and Schulten 1991).

III.3.4.1 Présentation du problème

La distribution de points étudiée ici, a été utilisée dans plusieurs travaux (Demartines 1994)(Cselényi 2005)(Fritzke 1995) ; elle est représentée dans la Figure III.33. Sa principale caractéristique est qu'elle présente des dimensions intrinsèques différentes suivant l'endroit où l'on se place ; on peut en effet la décomposer en trois parties distinctes :

- 2000 points aléatoirement répartis dans un cube de côté 1 ; ces points forment un groupe de dimension intrinsèque 3 ;

- 800 points aléatoirement répartis sur une portion de plan carrée de côté 1 ; ces points forment un groupe de dimension intrinsèque 2 ;
- 200 points aléatoirement répartis sur un segment de droite (100 points) et un cercle (100 points) ; ces points forment un groupe de dimension intrinsèque 1.

Il est à noter que, les coordonnées étant exprimées dans un repère 3D, la distribution de données considérée est de dimension brute 3.

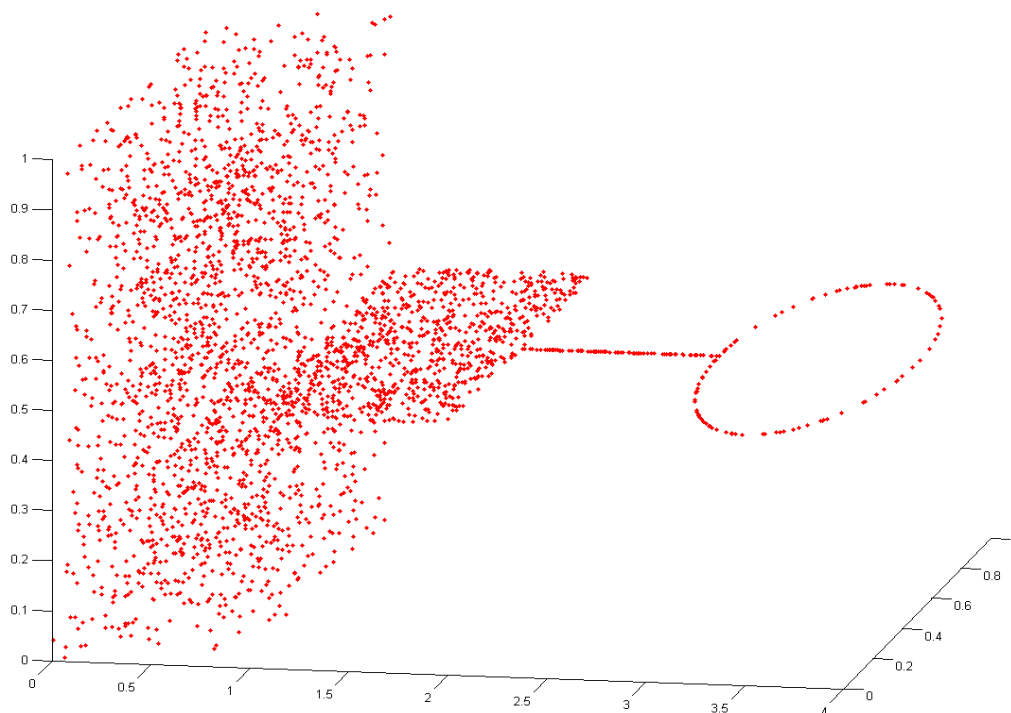


Figure III.33 : La distribution utilisée dans le problème jouet.

Afin de disposer d'un problème de classification, les points de cette distribution ont été ensuite séparés en deux groupes distincts, qui définissent deux classes d'appartenance (1 et -1). Pour les points du cube, la classe 1 est constituée des points situés à l'intérieur d'une boule de centre le centre du cube et de rayon fixé, et de 8 boules de même rayon et dont les centres sont les sommets du cube. Le plan est divisé en suivant le même principe (à la différence notable que, travaillant dans un plan 2D, on considère des disques et non pas des boules) : le résultat est représenté dans la Figure III.34. Enfin, le segment de droite et le cercle ont été divisés en 8 parties équivalentes le long de l'axe des abscisses.

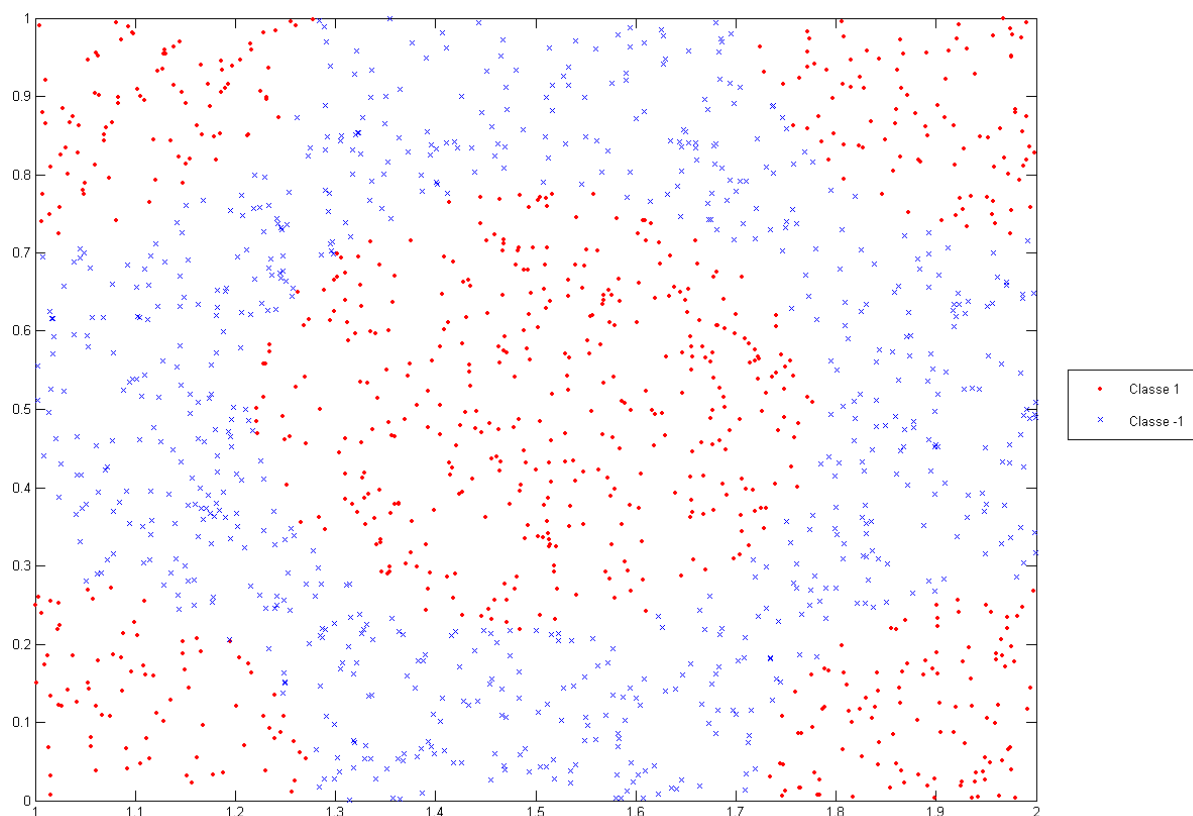


Figure III.34 : Répartition des classes 1 et -1 dans la partie plane de la distribution de points du problème jouet.

Il est à noter que les frontières entre les deux classes (déterminés directement par les rayons des disques et des boules) ont été choisies de manière à obtenir un nombre équivalent de données pour chacune des classes. Finalement l'ensemble des points a été scindé aléatoirement en deux bases de données de tailles égales décrites dans le Tableau III.7.

| Base de données | Nombre total de points | Nombre d'éléments de la classe 1 | Nombre d'éléments de la classe -1 |
|-----------------|------------------------|----------------------------------|-----------------------------------|
| 1 | 1500 | 750 | 750 |
| 2 | 1500 | 768 | 732 |

Table III.7 : Description des bases de données utilisées pour le problème jouet.

III.3.4.2 Protocole expérimental

Afin de valider la méthode de décomposition proposée, nous avons mené plusieurs expériences impliquant le problème jouet décrit ci-dessus. Ces expériences ont étudié deux aspects distincts de la décomposition : d'une part, sa capacité à extraire la structure topologique du problème, et d'autre part, la pertinence de l'exploitation de cette structure dans le cadre d'une classification modulaire.

Dans un premier temps, on a étudié l'influence des paramètres sensibles de notre méthode de décomposition sur les clusters obtenus par l'application de celle-ci sur les données du problème jouet.

Ensuite, on a testé les performances d'une structure modulaire de classification construite à partir de ces groupes de points. Pour ce faire, on a associé un module de traitement à chaque sous-problème défini par un cluster de points donné. De ce point de vue, les phases d'apprentissage et de généralisation se déroulent comme pour la technique T-DTS, décrite dans le paragraphe III.3.2. Les modules peuvent implémenter tout type d'algorithme de classification ; nous avons choisi d'utiliser le même type de réseaux de neurones artificiels que dans le paragraphe III.2.4.2.2, c'est-à-dire des MLPs. Par ailleurs, leurs caractéristiques (type de fonction de transfert des neurones, nombre de neurones dans la couche de sortie) et l'algorithme d'apprentissage utilisé sont identiques que pour ces expériences précédentes. La base de données 1 a été utilisée pour l'apprentissage du classifieur, et la base de données 2 dans la phase de généralisation ; afin d'obtenir des résultats statistiques, cinq apprentissages distincts ont été mis en œuvre. A titre de comparaison, le problème de classification a également été traité par un réseau de neurones MLP de même type mais cette fois ci de manière globale (sans décomposition).

III.3.4.3 Résultats et discussion

Le premier groupe d'expériences a permis une étude grossière des différents paramètres de la méthode. Dans ce but, les paramètres de voisinage ont été fixés à leur valeur classique ($k = 5$ et $Iteration = 5$). On a tout d'abord constaté que le paramètre correspondant au pas d'augmentation de la tolérance sur la dimension locale, s'il a une valeur assez petite, n'a pas d'impact sur les groupes de points obtenus. Concernant le paramètre de tolérance maximum, son influence est conforme à l'intuition. Plus il est grand, plus les points ont la

possibilité de s'agréger et moins le nombre de groupes générés et de points « orphelins » est important. Finalement, nous avons déterminé expérimentalement qu'un premier compromis satisfaisant était atteint pour une tolérance maximum de 0,3. Il faut cependant noter que cette valeur n'est pas critique puisque les résultats obtenus sont comparables pour une fourchette de tolérance maximum assez large (de 0,15 à 0,5).

En utilisant ces paramètres, notre méthode de décomposition a généré trois groupes de points associés à trois dimensions intrinsèques différentes. Le Tableau III.8 décrit la composition de ces trois clusters en termes d'appartenance au trois sous-distributions originales du problème jouet.

| Cluster | Dimension intrinsèque locale moyenne | Nombre de points du cube | Nombre de points du plan | Nombre de points de la ligne ou du cercle |
|----------------|---|---------------------------------|---------------------------------|--|
| 1 | 0.923 | 0 | 0 | 84 |
| 2 | 1.784 | 6 | 414 | 4 |
| 3 | 2.446 | 989 | 3 | 0 |

Table III.8 : Description des clusters générés par notre méthode de décomposition appliquée aux données du problème jouet.

Ces résultats montrent que la méthode proposée permet de mettre en évidence la structure topologique de la distribution étudiée. On peut remarquer que la dimension intrinsèque moyenne de chaque groupe est sous-évaluée. Ce phénomène de biais bien connu (paragraphe III.2.1) ne pose ici pas de problème particulier puisque les valeurs de dimension intrinsèque locale sont utilisées de façon exclusivement relative.

Nous avons ensuite mis en œuvre deux séries d'expériences de classification conformément au protocole expérimental présenté plus haut. La première série a considéré le problème jouet de manière global alors que la deuxième a mis en œuvre une structure modulaire (chaque module correspondant à un des clusters construits précédemment). Dans les deux cas, le nombre de neurones de la couche cachée des réseaux MLP utilisés a

été optimisé grossièrement de manière à maximiser les performances de généralisation du classifieur. Les résultats obtenus en termes de taux de reconnaissance (en généralisation) sont présentés dans le Tableau III.9.

| Classification | Nombre de neurones dans la couche cachée | Taux de reconnaissance de la classe -1 | Taux de reconnaissance de la classe 1 | Performance globale | Ecart-type de la performance globale |
|-----------------------|---|---|--|----------------------------|---|
| Globale | 40 | 87.9 | 85.5 | 86.7 | 1.83 |
| Modulaire | 20 | 92.8 | 88.5 | 90.7 | 1.25 |

Table III.9 : Nombre de neurones dans la couche cachée et performances de généralisation des classifieurs modulaires et globaux pour le problème jouet.

L'approche modulaire donne de meilleurs résultats que le classifieur global, en utilisant des modèles plus simples (moins de neurones dans la couche cachée). Une amélioration d'environ 4% du taux de généralisation a en effet été constatée, ce qui est significatif au vu des écarts-types associés (en dessous de 2%). Par ailleurs, ces résultats pourraient sûrement être encore améliorés en optimisant indépendamment le nombre de neurones cachés de chacun des trois réseaux de neurones (correspondant aux trois sous-problèmes). D'autre part, l'écart-type des performances est plus faible pour la structure modulaire : les sous-problèmes sont moins ambigus que le problème global, ce qui se traduit par des apprentissages plus stables.

D'une manière générale, ces premiers résultats montrent donc la pertinence de la technique de décomposition proposée.

III.3.5 Classification neuronale modulaire des défauts d'aspect

Dans le but de démontrer cette pertinence pour un problème de classification réel, les mêmes expériences ont été menées sur les données provenant de notre prototype.

Les données que nous avons étudié afin de valider la technique proposée sur des données réelles sont les mêmes que celles utilisées dans les paragraphes III.1.2.2.2 et III.2.4.2.2.

Nous avons également considéré le même problème de classification associé à ces données, consistant à discriminer les défauts de type « poussière » (classe P) des autres défauts (classe NP). Le Tableau III.10 rappelle brièvement la répartition des données entre les classes P et NP pour les bases de données 1 et 2.

| Base de données | Issue de l'expérience | Nombre total d'éléments | Nombre d'éléments de la classe P | Nombre d'éléments de la classe NP |
|-----------------|-----------------------|-------------------------|----------------------------------|-----------------------------------|
| 1 | B | 1910 | 184 | 1726 |
| 2 | A | 3865 | 275 | 3590 |

Table III.10 : Description des bases de données utilisées pour le problème réel.

En utilisant les mêmes paramètres que dans le paragraphe précédent, notre méthode de décomposition a permis d'isoler cinq groupes de données. La composition de ces cinq clusters en termes d'appartenance à la classe P ou NP est donnée dans le Tableau III.11.

| Cluster | 1 | 2 | 3 | 4 | 5 |
|-----------------------------------|-----|-----|-----|-----|----|
| Nombre d'éléments de la classe P | 175 | 5 | 3 | 1 | 0 |
| Nombre d'éléments de la classe NP | 983 | 228 | 264 | 159 | 92 |

Table III.11 : Description des clusters générés par notre méthode de décomposition appliquée aux données du problème réel.

Les défauts appartenant aux clusters 2,3,4 et 5 sont presque exclusivement des défauts autres que des poussières (entre 0 et 2% de poussières dans ces groupes). Cela valide notre hypothèse concernant le sens physique des sous-distributions de dimension intrinsèque homogène.

Dans un second temps, nous avons comparé les classifications modulaires et globales exactement de la même manière que pour le problème jouet. Les résultats obtenus sont présentés dans le Tableau III.12.

| Classification | Nombre de neurones dans la couche cachée | Taux de reconnaissance de la classe P | Taux de reconnaissance de la classe NP | Performance globale | Ecart-type de la performance globale |
|-----------------------|---|--|---|----------------------------|---|
| Globale | 35 | 70.2 | 97.2 | 95.3 | 0.46 |
| Modulaire | 20 | 72.8 | 98.5 | 96.7 | 0.11 |

Table III.12 : Nombre de neurones dans la couche cachée et performances de généralisation des classifieurs modulaires et globaux pour le problème réel.

Comme pour le problème jouet, l'approche modulaire donne de meilleurs résultats que le classifieur global. D'une manière générale, les conclusions que l'on peut tirer de ces expériences sont les mêmes que celles énoncées dans le paragraphe précédent : la technique de décomposition proposée est pertinente pour le traitement de notre problème réel.

Par ailleurs, on a mesuré le temps d'apprentissage et d'exécution (classification d'un défaut donné) pour la structure modulaire construite et le réseau de neurones unique. Les temps obtenus sont reportés dans le Tableau III.13.

| Classification | Phase d'apprentissage | Classification (de 500000 défauts) |
|-----------------------|------------------------------|---|
| Globale | 11180 s | 6.4 s |
| Modulaire | 644 s | 5.9 s |

Table III.13 : Temps caractéristiques associés aux classifieurs modulaires et globaux pour les phases d'apprentissage et de généralisation.

L'utilisation d'une structure modulaire permet d'économiser du temps en classification mais surtout pendant la phase d'apprentissage. Ceci est bénéfique pour le système de contrôle automatique à concevoir dans la mesure où il devient envisageable de le faire évoluer périodiquement par de nouveaux apprentissages.

III.4 Conclusions du chapitre

Dans ce chapitre nous nous sommes tout d'abord intéressés au problème du codage des données que l'on souhaite catégoriser. En effet, la taille importante des imagerie de défaut à traiter, le fait que cette taille ne soit pas constante et que les informations qu'elles contiennent soit très redondantes, rendent un traitement direct difficilement envisageable. Si un codage de ces images est nécessaire, il doit satisfaire à un certain nombre de critères visant à maintenir l'intégrité de l'information afin de garantir la représentativité des données tout en réduisant leur taille drastiquement. Dans ce sens le codage est fondamental puisqu'il conditionne la qualité de l'information traitée en aval par le système de classification et influe donc directement sur ses performances.

Notre choix s'est porté sur les invariants issus de la transformée de Fourier-Mellin. En effet, d'une part, ces invariants répondent aux contraintes fixées plus haut, et d'autre part, ils se sont révélés supérieurs à d'autres descripteurs, pour des tâches de classification et d'indexation dans plusieurs applications rapportées dans la littérature.

Afin de valider ce choix et de vérifier sa pertinence, une série d'expériences a été mise en œuvre. Ces expérimentations ont utilisé un grand nombre d'imagerie de défauts acquises sur des pièces optiques réelles et dans des conditions industrielles, grâce au prototype présenté au paragraphe II.2.4.1 (p. 61) Les imagerie obtenues ont été traitées, après codage, par une technique non supervisée, les Self Organizing Maps. On a pu mettre en évidence des clusters de défauts d'apparence semblable, ayant une signification physique adéquate pour la classification visée. Malheureusement, le fait que certaines cellules des SOMs, situées aux frontières entre classes, soient constituées de défauts de plusieurs classes différentes ne permet pas leur utilisation pour une classification directe. Les résultats obtenus sont néanmoins encourageants et nous paraissent suffisants pour valider le codage employé, d'autant plus qu'une bonne partie de l'information portée par les vecteurs descriptifs a été perdue lors de leur projection dans l'espace 2D des SOMs. Cet espace 2D est en effet d'une dimensionnalité trop petite pour traduire l'intégralité de la variabilité de ces données, dont la dimension intrinsèque est entre 6 et 7.

Nous avons étudié par la suite la problématique de la réduction de la dimension, dont le but est de donner une représentation compacte de l'information, en résolvant les rap-

ports linéaires et non-linéaires existants entre les composantes des vecteurs bruts codant les données. Les techniques de réduction de dimensionnalité permettent ainsi de limiter le phénomène du « fléau de la dimension », défavorable à l'utilisation de réseaux neuronaux. Nous avons montré, en utilisant les données réelles provenant des mesures de pièces optiques, que les techniques CCA et CDA sont supérieures à SOM en termes de qualité de projection. L'étude de cette projection a par ailleurs mis en lumière le fait que la distance curviligne est, pour ces données, significativement différente de la distance Euclidienne, ce qui justifie un emploi hypothétique de la CDA. L'impact possible de ces différentes techniques de réduction de dimensionnalité sur une tâche de classification réelle visant la discrimination des défauts permanents (ou tout du moins les éléments de la base de données C) a également été étudié. La CDA est la technique qui favorise le plus la localisation des défauts permanents dans la carte SOM et nous croyons donc qu'elle est la plus indiquée pour la discrimination de ces défauts. On peut de plus noter qu'elle semble avoir traduit la majeure partie de l'information présente dans les données brutes.

La comparaison de ces trois techniques s'est poursuivie avec l'étude d'une tâche de discrimination des défauts de type « poussière » par un réseau de neurones artificiels MLP, après une phase de réduction de dimension. Cette classification présente un grand intérêt pour notre projet et elle a permis de valider la chaîne de traitement complète, menant de l'imagette d'un défaut à sa nature. En effet, les performances de généralisation obtenues, environ 70% de poussières et 95% d'autres types de défauts correctement reconnus, sont en première approche satisfaisantes pour l'application visée.

Par ailleurs, les expériences menées ont montré que les réductions de dimension par CCA et CDA conduisaient à de meilleures performances de classification que l'utilisation de SOM, confirmant ainsi les conclusions énoncées plus haut. Nous avons également prouvé que l'on pouvait obtenir des résultats équivalents en utilisant indifféremment les données brutes ou des représentations de dimensions moindres (environ 2 fois plus compactes) obtenues par le biais de la CCA ou de la CDA. Tout d'abord, cela établit la validité du concept de réduction de dimension, puisque les projections « efficaces » correspondent à des dimensions d'arrivée compatibles avec la dimension intrinsèque estimée de la distribution de données. D'autre part, cela confirme la pertinence de cette approche d'un point de vue pratique puisqu'il permet l'utilisation de classifieurs plus simples.

Cependant, la réduction de dimension ne trouve dans notre cas pas d'intérêt applicatif puisqu'elle n'a pas permis d'améliorer notablement les performances du classifieur que ce soit en termes de taux de généralisation (la dimension du problème, 13, n'est sans doute pas assez grande pour que le « fléau de la dimension » soit très gênant) ou de temps de calcul (les algorithmes CCA et CDA, s'ils permettent d'utiliser des classifieurs plus simples, sont lourds algorithmiquement parlant).

Afin d'améliorer ces performances, nous avons mis en œuvre un traitement neuronal modulaire, qui emploie plusieurs « experts » locaux, mettant ainsi à profit le paradigme « diviser pour régner ». Ce type de structure vise en premier lieu l'utilisation de modules simples qui permettent un gain algorithmique important. Le classifieur proposé reprend les grandes lignes de la technique T-DTS mais utilise un nouveau mécanisme de décomposition des problèmes de classification.

Ce mécanisme tente d'isoler les différentes sous-distributions de dimensions intrinsèques homogènes qui existent au sein des données traitées. En effet, la dimension intrinsèque d'un ensemble de données réelles est liée au nombre de degrés de liberté du processus physique qui les a engendrées ; on cherche ainsi à traiter séparément des problèmes homogènes d'un point de vue physique, que l'on espère résoudre d'une manière plus efficace que le problème considéré dans sa globalité.

Pour appliquer cette idée, un algorithme de décomposition a été mis au point ; il utilise une estimation locale de la dimensionnalité et une phase de clustering itératif prenant en compte la topologie des données. Cet algorithme a été validé sur des données synthétiques dont il permet d'extraire la structure topologique particulière. D'autre part, des expériences de classification ont été menées sur ce problème jouet, impliquant des réseaux de neurones de type MLP. Elles ont montré que la structure modulaire obtenue conduisait à de meilleurs taux de généralisation (un gain d'environ 4%) que dans le cas d'un traitement global, en mettant en œuvre des classifieurs plus simples. Ces résultats montrent la pertinence de la technique développée.

Nous avons enfin utilisé cette technique sur nos données réelles. La phase de décomposition du problème a mis en évidence des groupes de données ayant un sens physique, ce qui indique que notre hypothèse de base (l'existence de sous-distributions de données de dimensions différentes issues de processus physiques distincts) était fondée. La structure

modulaire résultante permet d'améliorer les résultats obtenus précédemment avec des classifieurs globaux. Tout d'abord, les performances en généralisation sont supérieures d'environ 2%, ce qui est significatif au vu des écarts-types correspondants (environ 0,1 et 0,4%). Mais surtout, les modèles locaux utilisés sont très simples et permettent ainsi de limiter considérablement la durée de la phase d'apprentissage (qui est divisée par 20).

Conclusion générale

Dans un cadre industriel de production de grande série, on mise généralement sur la maîtrise de la chaîne de fabrication afin de réduire les contrôles qualité et les rebuts associés. Au contraire, le domaine des optiques hautes performances implique la fabrication de petites séries à forte valeur ajoutée et requiert donc un contrôle unitaire des composants produits. Afin de s'assurer que la pièce polie pourra remplir son rôle au sein du système optomécanique auquel elle est destinée, elle doit satisfaire à un certain nombre de critères, formalisés par l'utilisation de normes. Parmi ces critères, l'aspect de la pièce est incontournable ; les imperfections de surface tels que les « piqûres » et les « rayures », générées pendant les différentes phases de fabrication du composant, engendrent en effet des lumières parasites qui peuvent dégrader les performances optiques finales. Le nombre de ces imperfections (et leur gravité) doit donc être limité pour que la qualité de la pièce soit compatible avec l'application visée. Le rôle du contrôle d'aspect est de garantir cette qualité.

Traditionnellement, ce contrôle est réalisé à partir d'inspections visuelles qui s'avèrent insuffisantes en termes d'exhaustivité et de sensibilité et représentent une tâche pénible demandant une grande expertise et beaucoup de temps. Les marchés émergents, avec des critères de détection de défauts de plus en plus exigeants et des pièces optiques de plus en plus grandes, nécessitent une meilleure caractérisation des défauts d'aspect, d'autant plus que ces derniers peuvent mettre en péril l'intégrité physique des pièces.

Avant cette thèse, peu d'équipements automatisés de contrôle permettaient le diagnostic des défauts d'aspect, et les systèmes existants n'étaient pas, pour diverses raisons, adaptés aux conditions de production réelles. Ce travail constitue une première réponse à cette problématique industrielle par l'étude et la conception d'un système automatisé de contrôle d'aspect des pièces optiques hautes performances.

Afin de détecter les défauts présents sur la pièce optique contrôlée et d'en extraire des informations de bas niveau (position, taille et représentation image) dans un temps raisonnable, nous avons proposé de mettre en place un système exploitant deux capteurs distincts. Le premier capteur (dit de détection « grossière »), mettant à profit le principe de la strioscopie ou de l'éclairage par les tranches, est capable de détecter les défauts d'aspect recherchés en un temps restreint (typiquement de l'ordre de la minute). Le deuxième capteur (dit de détection « fine »), qui utilise la microscopie Nomarski, visite les sites

identifiés dans la première phase du contrôle pour apporter une caractérisation des défauts en termes de taille, position et nature. Cette structure multi-résolution permet de tirer avantages des caractéristiques propres aux deux types de capteurs : d'une part, la rapidité de la strioscopie (ou de la technique d'éclairage par les tranches), d'autre part, la sensibilité, la robustesse et la richesse en termes d'information de la microscopie Nomarski. Il est ainsi possible de satisfaire à la contrainte limitant le temps de contrôle, tout en assurant une caractérisation fiable et robuste des défauts d'aspect.

Nous nous sommes plus particulièrement concentrés sur l'étude du système mettant en œuvre le capteur de détection « fine », utilisant la microscopie Nomarski. De nouvelles techniques de traitement ont été proposées, permettant la détection des défauts d'aspect et la mesure de la rugosité à partir d'images issues de cette technologie. D'un point de vue métrologique, les mesures obtenues à partir de ces algorithmes sont satisfaisantes en termes de justesse et de répétabilité, ce qui garantit la robustesse et la fiabilité des contrôles les utilisant. Ces méthodes ont été implémentées au sein d'un prototype associant un microscope, une platine de translation et un environnement logiciel dédié. Au final, le prototype développé permet le contrôle de pièces optiques de taille limitée et une exploitation poussée des mesures obtenues (cartographies, recalage de cartes, etc). Il est utilisé en production au sein de la Société REOSC avec succès, ce qui démontre sa fiabilité et valide ainsi les techniques proposées.

Pour apporter une solution au problème de catégorisation des défauts d'aspect, prépondérant pour la pertinence du contrôle, nous avons choisi de mettre en œuvre des approches basées sur les réseaux de neurones artificiels. Cependant, les « imagerie » de défauts extraites des images Nomarski lors de la phase de détection ne peuvent être traitées directement. Leur codage nécessaire a été effectué par le biais des invariants associés à la transformée de Fourier-Mellin. Au final, la chaîne conduisant de l'image Nomarski brute à une décision quant à la nature d'un défaut se présente comme décrit dans la figure ci-dessous.

L'utilisation de cette technique de codage a été validée par une expérience de clustérisation à partir de SOM. Des résultats encourageants ont été obtenus, puisqu'on a mis en évidence des groupes de données correspondant à des défauts d'aspect ayant la même apparence (et donc probablement la même nature). Cela montre qu'une bonne partie de



Chaîne de traitement conduisant d'une image Nomarski à la nature d'un défaut.

l'information « visuelle » sur les défauts a été retranscrite dans un espace 2-D par la SOM, impliquant que cette information était disponible par le biais du codage utilisé, et donc que celui-ci est pertinent.

Cependant, cette opération de clustering ne permet pas d'obtenir une solution satisfaisante pour notre problème de catégorisation par un étiquetage des cellules de la SOM. On observe en effet que, près des frontières entre classes, les données projetées dans les cellules sont hétérogènes, ce qui interdit une classification directe performante. Ce phénomène n'est par ailleurs pas étonnant puisque la dimension intrinsèque de la distribution étudiée est plutôt aux environs de 6 : une SOM 2D met donc fatalement une bonne partie de l'information disponible de côté.

Nous avons étudié par la suite la problématique de réduction de la dimensionnalité du problème, ayant pour but de donner une représentation compacte de l'information contenue dans les vecteurs bruts représentant les données associées, en résolvant les rapports linéaires et non-linéaires existant entre leurs composantes. Les techniques de réduction de dimension ont en ce sens un intérêt applicatif puisqu'elles permettent de limiter le phénomène du « fléau » de la dimension, néfaste au fonctionnement correct des réseaux neuronaux. Nous avons montré en étudiant des données réelles issues de mesures de pièces optiques par notre prototype, que les techniques CCA et CDA sont supérieures à l'algorithme SOM, en termes de qualité de projection et de localisation des classes de défauts dans l'espace des données. Ceci permet de conjecturer que l'impact de ces deux techniques sur une classification postérieure doit être plus bénéfique que celui des SOMs.

La validité de cette hypothèse a été démontrée dans le cadre d'une tâche de discrimination des défauts de type « poussière » par un réseau de neurones artificiels de type MLP, ayant par ailleurs un intérêt applicatif évident pour notre projet. Les expériences menées ont également prouvé que l'on pouvait obtenir des résultats équivalents en utilisant indifféremment les données brutes ou des représentations de dimensions moindres

(environ 2 fois plus compactes) obtenues par le biais de la CCA ou de la CDA. De plus, les dimensions d'arrivée de ces projections étaient en accord avec la dimension intrinsèque estimée de la distribution de données. Ces deux résultats montrent la validité du fondement théorique de la réduction de dimension.

D'un point de vue plus pragmatique, les performances en généralisation des classifieurs mis en œuvre, environ 70% de poussières et 95% d'autres types de défauts correctement reconnus, sont en première approche satisfaisantes pour l'application visée. D'autre part, la réduction de dimension ne trouve dans l'immédiat pas d'intérêt applicatif puisqu'elle n'a pas permis d'améliorer notablement les performances du classifieur que ce soit en termes de taux de généralisation (la dimension du problème, 13, n'est sans doute pas assez grande pour que le « fléau » de la dimension soit très gênant) ou de temps de calcul (les algorithmes CCA et CDA, s'ils permettent d'utiliser des classifieurs plus simples, sont lourds algorithmiquement parlant).

Cependant, une des perspectives à court terme de ce travail est d'utiliser plus d'invariants de Fourier-Mellin pour la représentation de chaque défaut, amenant par ce biais une plus grande quantité d'information sur le problème envisagé afin d'atteindre de meilleures performances. Dans ce cas, la réduction de la dimension des données serait probablement une étape clé pour conserver un classifieur d'une complexité raisonnable et lui éviter de souffrir des effets de la haute dimensionnalité du problème.

Ces expériences ont également mis à jour de curieux phénomènes d'« ondulation » des performances de classification lors de modifications de la dimension d'arrivée de l'étape de réduction de dimensionnalité. Ceci semble indiquer qu'il existe plusieurs sous-distributions de dimensions intrinsèques différentes au sein de la distribution de données globale. D'un point de vue physique, cette hypothèse est pertinente puisque la dimension intrinsèque d'un ensemble de données réelles est liée au nombre de degrés de liberté du processus physique qui les a engendrées. Les sous-distributions de dimensions différentes correspondraient alors à des processus physiques distincts. Nous avons mis à profit ce phénomène dans une nouvelle approche de décomposition de problème pour les structures neuronales modulaires du type « diviser pour régner » (les modules travaillent sur des sous-problèmes locaux), dont un exemple est l'algorithme T-DTS. L'idée est de traiter séparément chaque sous-distribution de dimension homogène existante. Ainsi, la structure topologique des

données est respectée et au final les modules traitent des problèmes homogènes d'un point de vue physique, que l'on espère plus simples à résoudre que le problème considéré dans sa globalité. D'autre part, cela fournit une motivation claire et compréhensible pour guider le processus de décomposition de problèmes (contrairement aux techniques de quantification vectorielle utilisées conventionnellement).

Pour mettre en oeuvre cette idée, nous avons proposé un algorithme basé sur une estimation locale de la dimension intrinsèque et un clustering itératif. Cet algorithme a été validé sur un problème jouet puisqu'il a permis de recouvrer la structure topologique particulière de la distribution considérée. Une expérience de classification impliquant ces données a par ailleurs montré que les sous-problèmes isolés étaient plus faciles à traiter que le problème global. Enfin, l'utilisation de cette technique modulaire sur nos données industrielles a permis de dégager une solution plus performante que précédemment, en termes de taux de généralisation et de temps d'apprentissage (un temps d'apprentissage ramené à moins de 700 secondes et un gain de performances en généralisation d'environ 2%, significatif au vue des écarts-types obtenus).

Plusieurs perspectives concernant cette nouvelle approche sont envisagées. Tout d'abord, il serait souhaitable de comparer ses performances avec celles obtenues par d'autres méthodes de classification sur un panel plus important de problèmes réels. Il est également envisageable d'intégrer ce mode de décomposition à la structure T-DTS afin de tirer profit du mécanisme de contrôle de la décomposition par estimation de la complexité, propre à cette structure arborescente. Enfin, une hybridation avec des techniques de réduction de dimension paraît pertinente : le fait de travailler sur des sous-problèmes de dimension intrinsèque homogène est en effet probablement bénéfique pour la qualité des projections obtenues par de telles techniques.

Finalement, s'il a permis d'identifier la structure multi-résolution décrite plus haut et les capteurs correspondants, ce travail s'est concentré principalement sur le deuxième capteur et la caractérisation des défauts à partir des données qu'il fournit. Une perspective est donc de réaliser un prototype du système complet, répondant aux spécifications telles qu'elles ont été exprimées. La figure ci-dessous présente le schéma de principe de ce prototype capable de contrôler des pièces de taille importante et d'en donner un diagnostic de conformité suivant le formalisme d'une norme. Il met en oeuvre un capteur de détection

« grossière », le prototype présenté dans ce travail (dédié aux pièces de taille limitée) et un système expert codant les règles tirées des normes.

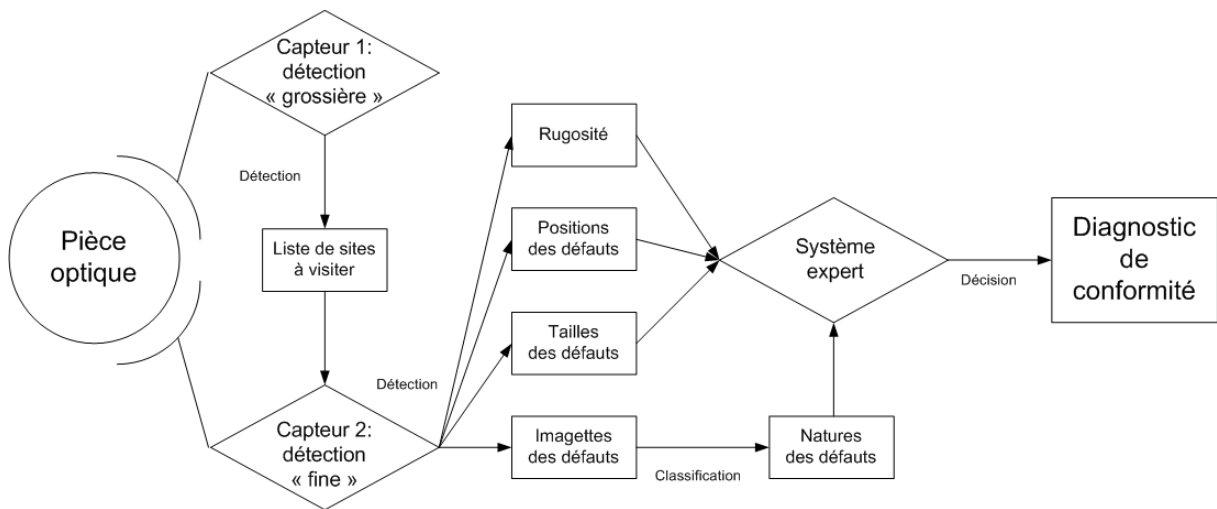


Schéma de principe du prototype complet.

Dans ce cadre, il est également nécessaire de traiter des problèmes de classification mettant en jeu plusieurs classes d'appartenance. La discrimination des seuls défauts de type « poussière » est en effet insuffisante pour le diagnostic recherché. Un travail conséquent sur la phase de catégorisation des défauts d'aspect est donc à envisager pour compléter les avancées exposées dans cette thèse.

Annexes

Normes concernant le contrôle d'aspect

Sommaire

| | |
|--|------------|
| A.1 Norme allemande - DIN 3140 P.10 - 1978 | 178 |
| A.1.1 Fonctionnement de la norme : | 178 |
| A.1.1.1 Défauts de surface | 178 |
| A.1.1.1.1 Notation rencontrée | 178 |
| A.1.1.1.2 Exemple | 179 |
| A.1.1.1.3 Signification | 179 |
| A.1.1.2 Lacunes des revêtements de surface | 179 |
| A.1.1.2.1 Définition | 179 |
| A.1.1.2.2 Notation rencontrée | 179 |
| A.1.1.2.3 Exemple | 179 |
| A.1.1.2.4 Signification | 180 |
| A.1.1.3 Applicable aux deux types de défauts (défauts de sur- face et lacunes de revêtement) | 180 |
| A.1.1.3.1 Fractionnement | 180 |
| A.1.1.3.2 Concentration | 180 |
| A.1.2 Avantages | 181 |
| A.1.3 Inconvénients | 181 |
| A.1.4 Document de référence | 182 |
| A.2 Norme Internationale - NF ISO 10110-7 Déc 1996 - Méthode I182 | |
| A.2.1 Fonctionnement de la norme | 182 |

| | | |
|------------|---|------------|
| A.2.1.1 | Notation rencontrée | 182 |
| A.2.1.2 | Exemple | 182 |
| A.2.1.3 | Signification | 183 |
| A.2.1.4 | Fractionnement | 183 |
| A.2.1.5 | Concentration | 184 |
| A.2.2 | Avantages | 185 |
| A.2.3 | Inconvénients | 185 |
| A.2.4 | Documents de référence | 185 |
| A.3 | Norme internationale - NF ISO 10110-7 Déc 1996 - Méthode | |
| II | | 185 |
| A.3.1 | Fonctionnement de la norme | 185 |
| A.3.1.1 | Notation rencontrée | 186 |
| A.3.1.2 | Exemple | 186 |
| A.3.1.3 | Signification | 186 |
| A.3.1.4 | Utilisation | 187 |
| A.3.2 | Avantages | 187 |
| A.3.3 | Inconvénients | 187 |
| A.3.4 | Documents de référence | 187 |
| A.4 | Norme américaine - MIL-PRF-13830B - Janv 1997 | 188 |
| A.4.1 | Fonctionnement de la norme | 188 |
| A.4.1.1 | Notation rencontrée | 188 |
| A.4.1.2 | Exemple | 188 |
| A.4.1.3 | Scratch | 189 |
| A.4.1.3.1 | Pour les éléments circulaires | 189 |
| A.4.1.3.2 | Pour les éléments non circulaires | 190 |
| A.4.1.3.3 | Dispersion des rayures | 190 |
| A.4.1.3.4 | Défaut de traitement | 190 |
| A.4.1.4 | Dig | 190 |
| A.4.1.4.1 | Dispersion des piqûres | 191 |
| A.4.1.4.2 | Bulles et inclusions | 191 |
| A.4.2 | Avantages | 191 |

A.4.3 Inconvénients 191

A.4.4 Documents de référence 191

A.5 Conclusion 192

La norme est incontournable dans le milieu industriel puisqu'elle permet un dialogue clair entre le fabricant et son client et qu'elle est une garantie de qualité pour ce dernier. Les défauts d'aspect étant des critères fonctionnels importants voire critiques dans le domaine de l'optique, ils font naturellement l'objet de normes. Cependant, comme nous le verrons, le lien entre fonctionnalité de la pièce et défauts d'aspect l'entachant est difficile à définir, ce qui explique peut-être l'existence de plusieurs normes les concernant. Nous étudierons dans ce document ces différentes normes en expliquant leur fonctionnement et en présentant leurs avantages et leurs inconvénients.

A.1 Norme allemande - DIN 3140 P.10 - 1978

Cette norme allemande, datant de 1978 n'est maintenant presque plus utilisée. Elle est cependant très intéressante par sa robustesse et car elle jette les bases de la future norme ISO décrite plus loin.

A.1.1 Fonctionnement de la norme :

Le principe de base de cette norme est de caractériser le défaut par sa nature et par ses dimensions (son aire pour un défaut ponctuel et sa largeur pour une rayure). Ces critères dimensionnels ont l'avantage d'être objectifs mais ils ne rendent pas directement compte de l'impact fonctionnel des défauts. Elle est applicable aux défauts de surface de la pièce avant ou après traitement.

A.1.1.1 Défauts de surface

A.1.1.1.1 Notation rencontrée

Cette notation donne le niveau d'acceptabilité :

| | |
|---|------------------------------------|
| | 5 / N × A ; K N' × A' ; R P |
| - des imperfections générales de surfaces | _____ |
| - des rayures longues | _____ |
| - des égrenures | _____ |

Nota : Ceci est l'indication complète d'imperfection de surface. Plus couramment, seules les indications sur les imperfections générales de surface sont données (**5 / N × A**).

A.1.1.1.2 Exemple

5 / 3 × 0,63 ; K 2 × 0,01 ; R 0,5

A.1.1.1.3 Signification

| | |
|--|---|
| 5 / N × A 5 N A | Numéro de code pour les imperfections de surface. Nombre des imperfections de surface autorisées de la dimension maximale admise. Racine carrée de l'aire maximale autorisée de la surface du défaut (en mm). |
| K N' × A' K N' A' | Lettre désignant les rayures (de longueur quelconque). Nombre de rayures admises. Largeur des rayures admises (en mm). |
| R P R P | Lettre désignant les égrenures. Profondeur des égrenures admises (en mm). |

A est appelé palier.

A.1.1.2 Lacunes des revêtements de surface

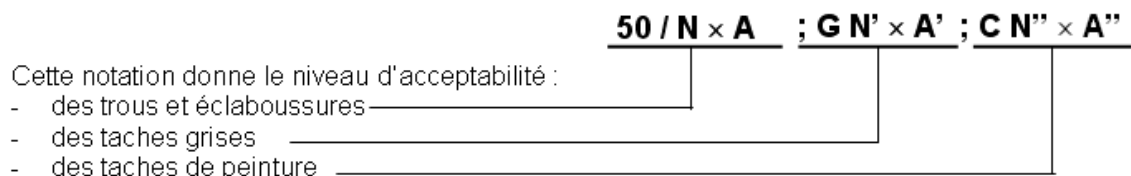
A.1.1.2.1 Définition

Les lacunes des revêtements de surface représentent toutes les défauts ou modifications localement délimitées des revêtements de surface.

On distingue notamment :

- les trous : emplacements non revêtus
- les éclaboussures : emplacements recouverts de couches
- les taches grises : emplacements dispersant le rayonnement incident
- les taches de peinture : emplacements réfléchissant ou transmettant le rayonnement incident dans des longueurs d'onde dans une autre plage que celle prévue.

A.1.1.2.2 Notation rencontrée



A.1.1.2.3 Exemple

50 / 3 × 0,25 ; G 2 × 10 ; C 1 × 16

A.1.1.2.4 Signification

| | |
|--|--|
| 50 / N × A 50 N A | Numéro de code pour les trous et éclaboussures. Nombre des trous et éclaboussures autorisés de la dimension maximale admise. Racine carrée de l'aire maximale autorisée de la surface d'un trou ou d'une éclaboussure (en mm). |
| G N' × A' G N' A' | Lettre désignant les taches grises. Nombre de taches grises admises. Racine carrée de l'aire maximale autorisée de la surface d'une tache grise (en mm). |
| C N'' × A'' C N'' A'' | Lettre désignant les taches de peinture. Nombre de taches de peinture autorisées de la dimension maximale admise. Racine carrée de l'aire maximale autorisée de la surface d'une tache grise (en mm). |

A est appelé palier.

A.1.1.3 Applicable aux deux types de défauts (défauts de surface et lacunes de revêtement)

A.1.1.3.1 Fractionnement

Une lacune d'un palier donné peut être remplacée par un plus grand nombre de lacunes d'un palier inférieur à condition que la somme de leurs surfaces ne dépasse pas celle de la lacune du palier le plus grand prévu. Le tableau de la page suivante indique comment fractionner les défauts.

Nota : Le fractionnement d'une lacune en plus de 40 plus petites lacunes et le fractionnement au-dessous d'un palier de 0,0025 ne sont pas pris en considération.

Nota : Si le fractionnement n'est pas autorisé la valeur codique est donnée entre parenthèses.

| |
|---------------------------------------|
| 5 / (N × A) ⇔ Fractionnement interdit |
|---------------------------------------|

A.1.1.3.2 Concentration

Les concentrations de lacunes ne sont pas admises.

On parle de concentration lorsque plus de deux lacunes ou 1/5 du nombre des lacunes se trouvent sur 4% de la surface de la zone contrôlée.

| |
|--|
| ds 4% de la zone contrôlée : + de 2 lacunes ou + 1/5 des lacunes ⇒ Concentration ⇒ NON CONFORME |
|--|

| Facteurs de multiplication du nbre d'imperfections | 1 | 2,5 | 6,3 | 16 | 40 |
|--|--------------|--------|--------|--------|--------|
| A en mm | 0,006 | 0,0040 | 0,0025 | | |
| | 0,010 | 0,0063 | 0,0040 | 0,0025 | |
| | 0,016 | 0,010 | 0,0063 | 0,0040 | 0,0025 |
| | 0,025 | 0,016 | 0,010 | 0,0063 | 0,0040 |
| | 0,040 | 0,025 | 0,016 | 0,010 | 0,0063 |
| | 0,063 | 0,040 | 0,025 | 0,016 | 0,010 |
| | 0,10 | 0,063 | 0,040 | 0,025 | 0,016 |
| | 0,16 | 0,10 | 0,063 | 0,040 | 0,025 |
| | 0,25 | 0,16 | 0,10 | 0,063 | 0,040 |
| | 0,40 | 0,25 | 0,16 | 0,10 | 0,063 |
| | 0,63 | 0,40 | 0,25 | 0,16 | 0,10 |
| | 1,0 | 0,63 | 0,40 | 0,25 | 0,16 |
| | 1,6 | 1,0 | 0,63 | 0,40 | 0,25 |
| | 2,5 | 1,6 | 1,0 | 0,63 | 0,40 |
| | 4,0 | 2,5 | 1,6 | 1,0 | 0,63 |
| 6,3 | 4,0 | 2,5 | 1,6 | 1,0 | |
| 10 | 6,3 | 4,0 | 2,5 | 1,6 | |

A.1.2 Avantages

- Clarté de la norme

A.1.3 Inconvénients

- Norme peu utilisée
- Norme fastidieuse (calculs d'aires)
- Ambiguïté de la norme (on peut choisir de compter un défaut comme une rayure ou une imperfection de surface)
- Rapport avec la fonctionnalité de la pièce peu clair

A.1.4 Document de référence

- Norme DIN 3140 Partie 7 (1978) :
Indication des cotes et tolérances des pièces optiques

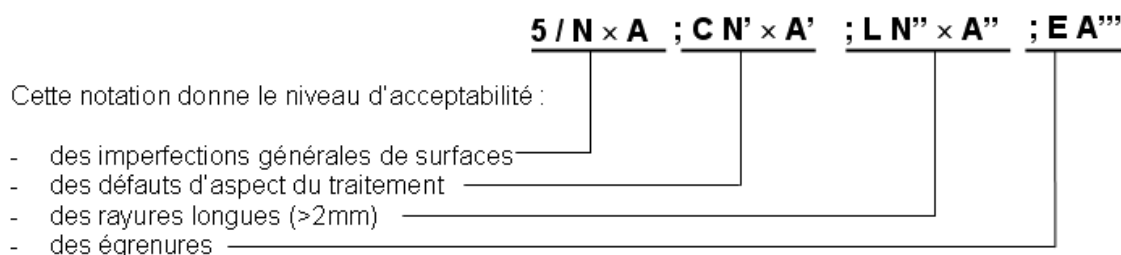
A.2 Norme Internationale - NF ISO 10110-7 Déc 1996 - Méthode I

Cette norme, proposée en décembre 1996, est massivement utilisée; elle dérive de la norme DIN 3140 présentée ci-dessus mais en diffère par quelques points importants.

A.2.1 Fonctionnement de la norme

Le principe de base de cette norme est de caractériser le défaut par sa nature et par ses dimensions (son aire pour un défaut ponctuel et sa largeur pour une rayure).

A.2.1.1 Notation rencontrée



Nota : Ceci est l'indication complète d'imperfection de surface. Plus couramment, seules les indications sur les imperfections générales de surface sont données ($5 / N \times A$).

A.2.1.2 Exemple

$5 / 3 \times 0,63 ; C 2 \times 1,6 ; L 2 \times 0,01 ; E 1,0$

A.2.1.3 Signification

| | |
|---|---|
| 5 / $N \times A$ 5 N A | Numéro de code pour les imperfections de surface. Nombre des imperfections de surface autorisées de la dimension maximale admise. Racine carrée de l'aire maximale autorisée de la surface du défaut (en mm). |
| C $N' \times A'$ C N' A' | Lettre désignant les défauts d'aspect du traitement. Nombre des défauts d'aspect du traitement autorisé de la dimension maximale admise. Racine carrée de l'aire maximale autorisée de la surface du défaut (en mm). |
| L $N'' \times A''$ L N'' A'' | Lettre désignant les rayures longues (>2mm). Nombre de rayures longues autorisées. Largeur maximale admise des rayures en millimètre (<i>Nota</i> : pas de limitation en longueur). |
| E A''' E A''' | Lettre désignant les égrenures. Etendue maximale admissible d'une égrenure depuis le bord physique de la surface de l'élément, mesuré parallèlement à la surface (en millimètre). <i>Nota</i> : Quelles que soient leurs quantités, les égrenures sont admises tant que leur étendue à partir du bord ne dépasse pas A''' . |

A et A' sont appelés références de classe

A.2.1.4 Fractionnement

- Pour les imperfections générales de surface et pour les défauts d'aspect du traitement :

Un nombre plus important d'imperfections avec une référence de classe inférieure est autorisé si la somme de leurs surfaces ne dépasse pas la surface totale maximale c'est à dire $N \times A^2$ (ou $N' \times A'^2$).

Le tableau de la page suivante permet de retrouver plus facilement le nombre d'imperfections autorisées et les surfaces correspondantes :

Nota : Les imperfections de surface, ayant une référence de classe 0,16 A ou moins ne doivent pas être comptées.

- Pour les rayures longues :

Un plus grand nombre de rayures longues est autorisé, à condition que la somme de leurs largeurs ne dépasse pas $N'' \times A''$.

Nota : Les rayures dont la largeur est inférieure à $0,3 \times A''$ ne doivent pas être comptées.

EXEMPLE : Si l'indication donnée est **L 2 × 0,01** ($2 \times 0,01 = 0,02$), sont autorisées :

2 rayures de largeur 0.01mm

ou 4 rayures de 0.005mm ($4 \times 0,005 \leq 0,02$)

ou 1 rayure de 0,01mm et 2 rayures de 0,05mm ($0,01 + 2 \times 0,05 \leq 0,02$) ...

| Facteurs de multiplication du nombre d'imperfections | 1 | 2,5 | 6,3 | 16 |
|--|-------|-------|-------|-------|
| A en mm | 0,006 | | | |
| | 0,010 | 0,006 | | |
| | 0,016 | 0,010 | 0,006 | |
| | 0,025 | 0,016 | 0,010 | 0,006 |
| | 0,040 | 0,025 | 0,016 | 0,010 |
| | 0,063 | 0,040 | 0,025 | 0,016 |
| | 0,10 | 0,063 | 0,040 | 0,025 |
| | 0,16 | 0,10 | 0,063 | 0,040 |
| | 0,25 | 0,16 | 0,10 | 0,063 |
| | 0,40 | 0,25 | 0,16 | 0,10 |
| | 0,63 | 0,40 | 0,25 | 0,16 |
| | 1,0 | 0,63 | 0,40 | 0,25 |
| | 1,6 | 1,0 | 0,63 | 0,40 |
| | 2,5 | 1,6 | 1,0 | 0,63 |
| | 4,0 | 2,5 | 1,6 | 1,0 |
| | 6,3 | 4,0 | 2,5 | 1,6 |
| 10 | 6,3 | 4,0 | 2,5 | |
| 16 | 10 | 6,3 | 4,0 | |
| 25 | 16 | 10 | 6,3 | |
| 40 | 25 | 16 | 10 | |
| 63 | 40 | 25 | 16 | |

EXEMPLE :

Si l'indication donnée est :

$$5/2 \times 0,25$$

sont autorisées :

2 imperfections de surfaces respectives de $0,25^2 = 0,0625 \text{ mm}^2$ (soit une surface totale maximale de $2 \times 0,0625 = 0,125 \text{ mm}^2$)

ou $2 \times 2,5 = 5$ imperfections de surfaces respectives de $0,16^2 = 0,0256 \text{ mm}^2$

ou $2 \times 6,3 \approx 12$ imperfections de surfaces respectives de $0,10^2 = 0,01 \text{ mm}^2$

ou $2 \times 16 = 32$ imperfections de surfaces respectives de $0,063^2 = 0,004 \text{ mm}^2$

Remarque : Pour simplifier, il est possible d'assimiler la valeur de A (en mm) au diamètre du défaut dans le cas où celui-ci est circulaire.

A.2.1.5 Concentration

Les concentrations d'imperfections de surfaces ne sont pas admises.

- On parle de concentration lorsque plus de 20% du nombre des défauts autorisés se trouvent dans 5% de la zone d'essai (ZE).

+ de 20% des défauts ds 5% de la ZE \Rightarrow NON-CONFORME (NC).

- Si le nombre de défauts est inférieur à 10, deux défauts ou plus se trouvant dans une zone de 5% constituent une concentration.

Si – de 10 défauts sur la ZE et + de 2 défauts ds 5% de la ZE \Rightarrow NC.

A.2.2 Avantages

- Clarté et objectivité de la norme
- Norme utilisée massivement

A.2.3 Inconvénients

- Norme fastidieuse (calculs d'aires)
- Cas particuliers non pris en compte (cas des rayures longues mais inférieures à 2mm qui doivent être comptées en tant qu'imperfections de surface et donc dont l'aire seule importe)
- Rapport avec la fonctionnalité de la pièce peu clair

A.2.4 Documents de référence

- Norme ISO 10110 Partie 7 (1996)
Optique et instruments d'optique - Indication sur les dessins pour éléments et systèmes optiques (Partie 7 : Tolérances d'imperfection de surface)

A.3 Norme internationale - NF ISO 10110-7 Déc 1996 - Méthode II

Cette norme, datant de décembre 1996, dérive de la norme AFNOR S 10-006, fruit des travaux menés conjointement par l'Ecole Supérieure d'optique et le GIFO (Groupement des Industries Françaises de l'Optique). Elle n'est que très peu utilisée car elle nécessite l'utilisation d'un matériel assez coûteux (que la REOSC ne possède pas).

A.3.1 Fonctionnement de la norme

Le principe de base de cette norme est de caractériser le défaut en mesurant (grâce à l'évaluation du flux reçu par la pupille de l'œil) la quantité de lumière qu'il diffracte. La mesure de la lumière parasite engendrée prend en compte l'ensemble des grandeurs

du défaut : non seulement ses dimensions d'affleurement à la surface mais aussi la forme du défaut, ses flancs polis ou non. Naturellement elle tient compte aussi de l'indice de réfraction du support sur lequel le défaut est gravé.

Cette méthode nécessite d'examiner l'élément en utilisant un poste de contrôle spécifique qui met en œuvre le principe de la strioscopie. La lumière diffractée par les imperfections de surface de l'échantillon soumis à essai est comparée à un fond lumineux de référence. La classe de visibilité est déterminée par le niveau d'éclairage de l'échantillon auquel les défauts de surface deviennent visibles ; chaque classe de visibilité correspond à un éclairage d'échantillon particulier. Afin de garantir un niveau uniforme de sensibilité, indépendant de l'observateur, le fond lumineux de référence doit être réglé en utilisant un échantillon étalon.

A.3.1.1 Notation rencontrée

5 / T V ou 5 / R V ; E A''

A.3.1.2 Exemple

5 / T 3 ; E 1,0

A.3.1.3 Signification

| | |
|---------------|---|
| T ou R | Contrôle en Transmission ou Réflexion des imperfections de surface. |
| V | Entier de 1 à 5 par ordre de sévérité décroissante (1 ⇔ sévérité maxi). |
| E | Lettre désignant les égrenures. |
| l.2.1. | Etendue maximale admissible d'une égrenure depuis le bord physique de la surface de l'élément, mesuré parallèlement à la surface (en mm). |
| A'' | <i>Nota:</i> Quelles que soient leurs quantités, les égrenures sont admises tant que leur étendue à partir du bord ne dépasse pas A''. |

V est appelé classe de visibilité.

A.3.1.4 Utilisation

| Sens de Qualité croissante ↓ | Classes de visibilité | Eclairage | Fond de référence |
|---------------------------------|-----------------------|-------------------|-------------------|
| | T5 ou R5 | 310 lux \pm 5% | Réglé |
| | T4 ou R4 | 625 lux \pm 5% | Réglé |
| | T3 ou R3 | 1250 lux \pm 5% | Réglé |
| | T2 ou R2 | 2500 lux \pm 5% | Réglé |
| T1 ou R1 | 2500 lux \pm 5% | Noir | |

A.3.2 Avantages

- Clarté de la norme
- Rapidité de mise en œuvre de la norme
- Possibilité d'effectuer un contrôle de conformité en transmission ou en réflexion
- Approche originale basée sur la visibilité des défauts

A.3.3 Inconvénients

- Norme peu utilisée
- Norme nécessitant l'acquisition d'un matériel coûteux
- Manque de sensibilité
- Pas de différenciation des défauts par rapport à leur nature
- Taille des pièces mesurables limitée

A.3.4 Documents de référence

- Norme ISO 10110 Partie 7 (1996)
Optique et instruments d'optique - Indication sur les dessins pour éléments et systèmes optiques (Partie 7 : Tolérances d'imperfection de surface)
- Huard, A. *Visibility method for classifying microscopic surface defects for both reflection and transmission systems*, Proc. SPIE. 525 pp. 37-42 (1985)
- Marioge, J.P., Huard, A., Munier, M., Hautcolas, J.L. *Validation of a local defect classification procedure*, Proc. SPIE. 1009 p.218 (1988)
- *Optical inspection station*, GIAT Industries, F-78013 Versailles, France (Leaflet)

Remarque :

Cette méthode nécessite l'acquisition d'un poste de contrôle spécifique peu adapté à la taille des pièces optiques réalisées au centre de St Pierre du Perray.

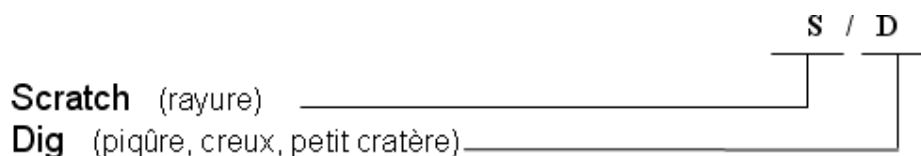
A.4 Norme américaine - MIL-PRF-13830B - Janvier 1997

Cette norme, datant de janvier 1997, est à la base une norme militaire américaine ; elle est pourtant aujourd'hui utilisée massivement dans le monde entier mais sujette à discussion.

A.4.1 Fonctionnement de la norme

Le principe de base de cette norme est de traiter séparément les défauts ponctuels (Dig), évalués de manière géométrique, et les défauts de type « rayure » (Scratch), évalués en utilisant leur visibilité. Employant le concept de visibilité, elle normalise également les conditions d'inspection des pièces (une des deux méthodes proposées préconise l'utilisation d'une lampe à incandescence de 40W, d'un fond noir et une observation en transmission de la lumière diffusée à 90 ° de la direction du rayon incident) .

A.4.1.1 Notation rencontrée



A.4.1.2 Exemple

80 / 50 : standard d'aspect communément acceptable

60 / 40 : standard acceptable pour la plupart des applications de recherche scientifique

10 / 5 : standard précis pour les applications laser très contraignantes

A.4.1.3 Scratch

S est appelé le **Scratch number**.

Cette notion pose de gros problèmes d'interprétation. En effet, la norme définit ce nombre comme une référence à un défaut étalon dont les caractéristiques sont indiquées dans le document C7641866. Un défaut a un scratch number S lorsqu'il a la même apparence que le défaut étalon de référence S. Cependant la définition de ces étalons a évolué (il en existe 11 versions!); on a généralement essayé de ramener le scratch number à une grandeur métrique (ce qui est un non-sens, le rapport entre visibilité et grandeurs géométriques étant très flou) mais il n'y a pas de véritable consensus à ce sujet d'autant plus que différentes versions donnent des valeurs différentes ($1 = 0,1\mu\text{m}$; $1 = 1\mu\text{m}$). La dernière version du document (L) propose une valeur de $0,1\mu\text{m}$ mais précise que l'étalon obtenu doit être confronté au défaut étalon. Le service qualité de la REOSC préconise de ramener ce nombre à une taille en microns comme le font la plupart des clients.

$$\mathbf{S}=80 \Leftrightarrow \text{rayure standard de } 80\mu\text{m de large}$$

A.4.1.3.1 Pour les éléments circulaires

La somme des longueurs des rayures correspondant à la taille maximale ne doit pas dépasser le quart du diamètre de la pièce contrôlée :

$$\Sigma L > 1/4 \varnothing \Rightarrow \text{NON CONFORME}$$

Nota : Si la pièce présente des rayures visibles correspondant à la taille maximale, la somme des surfaces de chaque défaut (Longueur \times largeur) divisée par le diamètre de l'élément (ou de la zone étudiée) ne doit pas dépasser la moitié du scratch number.

$$\Sigma (L \times l) / \varnothing > \mathbf{S}/2 \Rightarrow \text{NC}$$

Par contre, quand il n'y a pas de rayure de taille maximale, la somme des surfaces de chaque défaut divisée par le diamètre de l'élément de ne doit pas dépasser le scratch number.

$$\Sigma (L \times l) / \varnothing > \mathbf{S} \Rightarrow \text{NC}$$

A.4.1.3.2 Pour les éléments non circulaires

La méthode est la même que pour les éléments circulaires en estimant le diamètre qu'aurait une surface circulaire de même aire.

A.4.1.3.3 Dispersion des rayures

- Dans la zone utile (ZU), pour des scratch numbers ≥ 20 , il ne doit pas y avoir plus de quatre rayures dans une même zone de diamètre de 6 mm.

| |
|--|
| Si $10 \geq S \geq 20$ et si + 4 rayures dans un cercle $\varnothing 6\text{mm} \Rightarrow \text{NC}$ |
|--|

- Dans la zone non utile, l'élément est considéré avoir une spécification de 80-50 quelle que soit celle donnée.

| |
|----------------------------|
| Hors de la ZU spec : 80/50 |
|----------------------------|

A.4.1.3.4 Défaut de traitement

Les rayures apparues suite au traitement sont analysées comme décrits précédemment, mais considérées séparément.

A.4.1.4 Dig

D : **Diamètre** maximal accepté pour les défauts en **1/100 mm**.

| |
|--|
| $D=30 \Leftrightarrow \varnothing \text{ défaut max} = 0,3 \text{ mm}$ $\varnothing \text{ défaut} > D \Rightarrow \text{NC}$ |
|--|

Nota : Dans le cas de forme irrégulière, au lieu du diamètre, on utilise la moyenne de la longueur maximum et de la largeur correspondante.

Le nombre autorisé de piqûre correspondant à la taille maximale autorisée est de une par élément de surface de 20mm de diamètre. Dans cette zone, la somme de tous les diamètres des défauts ne doit pas dépasser $2 \times D$ (les défauts dont la taille est inférieure à $2,5 \mu\text{m}$ ne sont pas pris en compte).

| |
|--|
| Par zone de 20 mm de diamètre : 1 seul défaut de diamètre $\varnothing \approx D$ est autorisé $\Sigma \varnothing \text{ des piqûres} > 2 \times D \Rightarrow \text{NC}$ |
|--|

A.4.1.4.1 Dispersion des piqûres

Si $D \leq 10$, les piqûres doivent être séparées d'au moins un millimètre.

Distance entre 2 piqûres $< 1 \text{ mm} \Rightarrow \text{NC}$

La dispersion entre les piqûres n'est pas spécifiée pour les surfaces optiques ayant un dig number > 10 .

A.4.1.4.2 Bulles et inclusions

Les bulles et les inclusions sont classées comme les piqûres. La tolérance de taille est imposée par le dig number. En ce qui concerne la tolérance de dispersion, la règle est la même que pour les piqûres, les bulles devant être comptabilisées avec les piqûres.

A.4.2 Avantages

- Norme massivement utilisée

A.4.3 Inconvénients

- Ambiguïté de la norme (pas de précision quant à la limite basse de prise en compte des scratches) et débat sur la signification du scratch number
- Norme fastidieuse (interprétation pénible)
- Rapport avec la fonctionnalité de la pièce peu clair (cf. Référence)

A.4.4 Documents de référence

- Norme MIL-PRF-13830B (9 janvier 1997).
Performance specification – Optical components for fire control instruments ; general specification governing the manufacture, assembly, and inspection of.
- Young, M. *The scratch standard is only a cosmetic standard*
Damage in Laser Materials :1985, Nat. Bur. Stand. (U.S.) Spec. Publ. 746 (1987)

A.5 Conclusion

Les caractéristiques principales, les avantages et les inconvénients des normes présentées ici sont reportés dans le tableau ci-dessous.

| Norme | DIN3140 - 7 | ISO 10110 - 7 (Méthode 1) | ISO 10110 - 7 (Méthode 2) | MIL - PRF - 13830B |
|-------------------------|---|--|---|---|
| Caractéristiques | précurseur de la Norme ISO spécification Points/Rayures calcul d'aires et de largeur de défauts | très utilisée spécification Points/Rayures calcul d'aires et de largeur de défauts | basé sur la visibilité des défauts contrôle en transmission ou réflexion poste de contrôle spécifique | très utilisée spécification Points/Rayures basé sur la visibilité des défauts |
| Avantages | clarté de la norme | clarté de la norme massivement utilisée | clarté de la norme rapidité de mise en œuvre signification fonctionnelle forte | massivement utilisée |
| Inconvénients | peu utilisée fastidieuse (calcul d'aires) signification fonctionnelle floue | fastidieuse (calcul d'aires) signification fonctionnelle floue | nécessite un matériel coûteux peu utilisée manque de sensibilité | ambiguïté de la norme |

En ce qui concerne les normes ISO et MIL, qui sont utilisées dans l'immense majorité des cas, plusieurs remarques importantes doivent être rappelées ici. La norme ISO a l'avantage d'être sans ambiguïté mais n'est pas très représentative de l'impact fonctionnel des défauts mesurés. Au contraire, la norme MIL, en utilisant le concept de visibilité, cherche à se rapprocher de considérations fonctionnelles. Ceci se fait malheureusement au détriment de la clarté et de l'intégrité de la norme.

Finalement, les spécifications selon les normes citées ici s'avèrent plus une exigence de qualité découlant d'un consensus entre client et fournisseur, plutôt qu'un cadre formel rigide. Les normes, pour garder un sens se doivent donc d'être interprétées et adaptées afin d'assurer que leur raison d'être, c'est-à-dire la qualité du produit fourni au client, soit préservée.

Spécification du système à concevoir

Sommaire

| | | |
|------------|--|------------|
| B.1 | Identification de l'environnement | 195 |
| B.1.1 | Besoin | 195 |
| B.1.2 | Utilisateurs | 195 |
| B.1.3 | Produit | 195 |
| B.1.4 | Phases de vie du produit | 195 |
| B.2 | Spécification du besoin | 195 |
| B.2.1 | Répondre au besoin spécifique | 195 |
| B.2.2 | Permettre une généralisation du procédé | 196 |
| B.3 | Eléments de l'environnement | 197 |
| B.4 | Définitions des fonctions de service à satisfaire | 198 |
| B.4.1 | Fonction F1 | 198 |
| B.4.1.1 | Détecter les défauts | 198 |
| B.4.1.2 | Dimensionner les défauts | 198 |
| B.4.1.3 | Positionner les défauts | 199 |
| B.4.2 | Fonction F2 | 199 |
| B.4.2.1 | Etre compatible avec les cadences | 199 |
| B.4.2.2 | Être compatible avec les matériaux | 199 |
| B.4.2.3 | Être compatible avec les niveaux de finition | 199 |
| B.4.2.4 | Être compatible avec les géometries de pièces | 200 |
| B.4.3 | Fonction F3 | 200 |

| | | |
|---------|---|-----|
| B.4.3.1 | Fournir une interface de paramétrage | 200 |
| B.4.3.2 | Restituer les résultats du contrôle | 201 |
| B.4.4 | Fonction F4 | 201 |
| B.4.4.1 | Au cas par cas, représentation immédiate pour avis de l'expert | 201 |
| B.4.4.2 | Mensuellement, présentation d'un ensemble de cas (cas litigieux et cas limite) | 201 |
| B.4.5 | Fonction F5 | 202 |
| B.4.6 | Fonction F6 | 203 |
| B.4.7 | Fonction F7 | 203 |
| B.4.8 | Fonction F8 | 204 |
| B.4.9 | Fonction F9 | 204 |

B.1 Identification de l'environnement

B.1.1 Besoin

- Réaliser le contrôle d'aspect de pièces optiques.
- Permettre une généralisation du procédé mis en œuvre.

(voir ci dessous pour une spécification du besoin plus détaillée)

B.1.2 Utilisateurs

Opérateurs et experts du service Contrôle du centre SAGEM REOSC.

B.1.3 Produit

Système intelligent de contrôle d'aspect des pièces optiques.

B.1.4 Phases de vie du produit

- Mise au point
- Utilisation
- Maintenance

B.2 Spécification du besoin

B.2.1 Répondre au besoin spécifique

Réaliser le contrôle d'aspect de pièces optiques, ce avec les propriétés suivantes :

- *Exhaustivité* : le système doit permettre par ce biais une utilisation pertinente du formalisme des normes et un suivi des pièces via un archivage des données.
- *Répétabilité* : cette propriété, fondamentale pour un outil de contrôle, associée au caractère automatique du traitement, assure l'impartialité du contrôle.
- *Autonomie* : on vise ainsi une réduction maximale de l'intervention des experts.

- Respect des cadences : le système doit s'intégrer à la chaîne de production avec un coût minimum en termes de temps et de disponibilité.
- Robustesse vis-à-vis de l'environnement de travail : le système doit être capable de fonctionner de manière correcte dans les conditions correspondant à son lieu d'intégration.

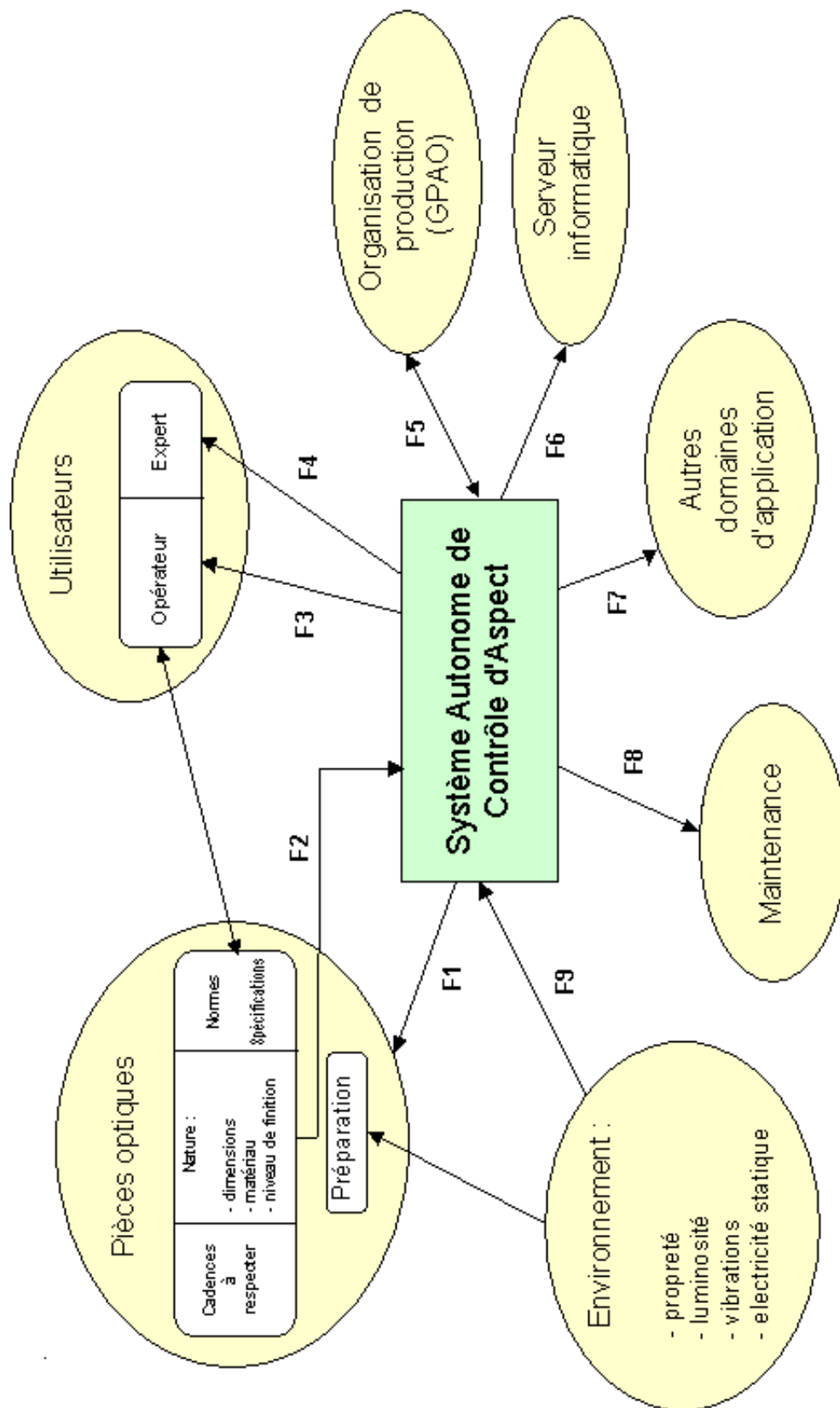
B.2.2 Permettre une généralisation du procédé

Développer l'outil de contrôle en conservant une bonne capacité de généralisation :

- Implantation d'un soft modulaire, permettant un portage de la solution (ou d'une de ses parties) vers d'autres domaines.
- Acquisition d'un savoir-faire innovant.
- Processus de certification dans le but de valider de façon formelle l'approche utilisée.

On vise ainsi le développement potentiel d'autres applications à moindre coût, dans un premier temps pour le contrôle de composants optiques plus petits, dans un deuxième temps pour le contrôle d'aspect de pièces hors du domaine de l'optique.

B.3 Eléments de l'environnement



B.4 Définitions des fonctions de service à satisfaire

B.4.1 Fonction F1

Détecter, dimensionner, et positionner *tous* les défauts d'aspects présents sur la pièce de façon *répétable*.

B.4.1.1 Détecter les défauts

On considère qu'un défaut correspond à une discontinuité de surface. Pour les pièces destinées à être utilisées en transmission, les bulles et les inclusions dans la matière sont également des défauts. Enfin les traces de nettoyage sont considérées comme des défauts (pour les optiques traitées) mais pas les poussières. On doit dans tous les cas préciser la nature du défaut.

On doit pouvoir détecter les défauts dont une des dimensions dépasse $2.5 \mu\text{m}$.

Il est cependant important de noter qu'il s'agit d'un cas limite : pour certains types de pièces, on se limitera aux défauts d'une taille supérieure ; pour mémoire, on se reportera au tableau qui suit :

| Application | Very Extreme Ultraviolet (VEUV) | Extreme Ultraviolet (EUV) | Laser | Optique Conventiennelle | Astronomie |
|--|---------------------------------|---------------------------|--------------------|-------------------------|-------------------|
| Taille des pièces (ordre de grandeur) | 100 cm ² | 100 cm ² | 1 m ² | 0,1 m ² | 10 m ² |
| Taille minimale des défauts à détecter | 50 nm | 2 μm | 5-30 μm | 50 μm | 100 μm |

B.4.1.2 Dimensionner les défauts

La largeur doit être connue avec une précision égale à 1% de la longueur du défaut avec une limite inférieure de $0.6 \mu\text{m}$.

La longueur, si elle est inférieure à $50 \mu\text{m}$ doit être connue avec une précision de $0.6 \mu\text{m}$; si elle est comprise entre $50 \mu\text{m}$ et 2 mm , elle doit être connue avec une précision de $5 \mu\text{m}$; si elle est supérieure à 2 mm , elle doit être connue avec une précision de 0.1 mm .

On doit connaître la nature du défaut : « creux » ou « bosse » ainsi que fournir une cartographie 2D de celui-ci. ; la connaissance approximative de sa profondeur est souhaitable.

B.4.1.3 Positionner les défauts

On doit positionner les défauts sur la pièce à contrôler avec une précision de l'ordre de 0,1 mm.

De plus, on doit avoir accès à un positionnement relatif entre les défauts de l'ordre de $10\mu\text{m}$ qui permette de séparer des défauts très proches l'un de l'autre.

B.4.2 Fonction F2

Être compatible avec les *cadences* et les différents types de *matériaux*, de niveaux de *finition*, et de *géométrie* de pièce.

B.4.2.1 Être compatible avec les cadences

Le temps de préparation, de mise en place au sein du système et de contrôle de la pièce doit être strictement inférieur au temps que mettrait un opérateur humain pour effectuer le même contrôle.

B.4.2.2 Être compatible avec les matériaux

Le système de contrôle doit fonctionner pour des pièces produites à partir de matériaux différents :

- Pièce à l'aspect vitreux transparent : ULE, silice, BK7 ...
- Pièce à l'aspect vitreux coloré : Zéro dur, Clearceram, Néodyme...
- Pièce à l'aspect métallique : SiC, verres traités métalliques...
- Pièce à la face arrière dépolie, amincie ou allégée.

B.4.2.3 Être compatible avec les niveaux de finition

Le système de contrôle doit fonctionner pour des pièces ayant des niveaux de finition différents :

- Pièces brutes

- Pièces ébauchées
- Pièces doucies
- Pièces polies
- Pièces traitées

A noter que la classe de défaut à détecter est différente suivant le niveau de finition.

B.4.2.4 Être compatible avec les géométries de pièces

Le système de contrôle doit fonctionner pour des pièces aux géométries différentes :

- Pièces planes
- Pièces convexes
- Pièces concaves
- Pièces à géométrie complexe (prisme...)
- Pièces de gros volume
- Pièces particulières dont l'arrangement des faces pourrait poser problème (lentille...)

B.4.3 Fonction F3

Fournir une *interface de paramétrage* et **restituer** les résultats du contrôle de façon *ergonomique*.

B.4.3.1 Fournir une interface de paramétrage

L'interface doit fournir de façon ergonomique la possibilité de :

- Fixer les paramètres descriptifs de la pièce (forme, dimensions, matière, rayon de courbure)
- Fixer une norme ou des spécifications auxquelles devront être confrontés les résultats du contrôle
- Enregistrer ces paramètres en vue de futurs contrôles
- Préciser le niveau de finition de la pièce
- Choisir la précision du contrôle (et donc sa durée)

B.4.3.2 Restituer les résultats du contrôle

L'interface doit fournir une restitution des résultats sous différentes formes :

- Listing brut des défauts détectés avec tous les renseignements disponibles correspondants (nature, localisation, taille)
- Carte de la pièce avec matérialisation des défauts détectés (avec un code couleur)
- Mise en relation des résultats du contrôle vis-à-vis de la norme ou des spécifications (avis de conformité ou non-conformité)
- Rédaction d'un PV de contrôle
- Mise en relation des résultats du contrôle avec ceux des contrôles précédents

B.4.4 Fonction F4

Fournir une *représentation détaillée* des défauts non-classifiés :

B.4.4.1 Au cas par cas, représentation immédiate pour avis de l'expert

Lorsque le système de contrôle est incapable de prendre une décision quant à la nature d'un défaut de manière sûre, un expert doit trancher ; pour cela le système doit lui présenter toutes les informations dont il dispose :

- Position du défaut
- Taille du défaut
- Informations de type morphologique (représentation du défaut si possible)
- Informations sur la pièce (nature, niveau de finition, spécifications...)
- Informations complémentaires (résultats du contrôle sur la pièce entière, évaluation du défaut par le système, état du système)

Le système doit également fournir une interface qui permette à l'expert de faire part de sa décision quant au défaut.

B.4.4.2 Mensuellement, présentation d'un ensemble de cas (cas litigieux et cas limite)

Tous les mois, le système doit proposer un certain nombre de « cas d'école », c'est-à-dire les cas qui ont nécessité l'intervention d'un expert et les cas qu'il a tranché le moins

sûrement, ce pour s'assurer que ses décisions sont toujours judicieuses.

Pour cela, le système doit présenter ces exemples avec toutes les informations dont il dispose :

- Position du défaut
- Taille du défaut
- Informations de type morphologique (représentation du défaut si possible)
- Informations sur la pièce (nature, niveau de finition, spécifications. . .)
- Informations complémentaires (résultats du contrôle sur la pièce entière, évaluation du défaut par le système, état du système)
- Décision prise par rapport au défaut considéré

Le système doit également fournir une interface qui permette au collègue d'experts de faire part de sa décision quant au défaut en vue de corriger son modèle interne.

B.4.5 Fonction F5

S'intégrer à une *organisation de production*.

Le système doit s'intégrer à l'organisation de production et ce en termes de :

- *Temps* :

Le temps de préparation et de contrôle des pièces doit être strictement inférieur au temps que mettrait un opérateur humain pour effectuer le même contrôle.

- *Disponibilité* :

Le système de contrôle doit avoir une disponibilité maximale pour le flux de production (robustesse, optimisation du temps machine).

- *Manipulation* :

Le système de contrôle doit nécessiter une intervention de l'opérateur minimale et les manipulations de la pièce à contrôler doivent être compatibles avec sa nature. De plus l'utilisation du système ne doit pas compromettre la sécurité de l'utilisateur.

- *Gestion de l'information :*

Le système de contrôle doit fournir des sorties structurées qui soient utilisables facilement par les autres éléments de la chaîne de production et pouvoir utiliser des informations provenant de ces éléments.

B.4.6 Fonction F6

Stocker et **regrouper** de façon pertinente des *résultats de contrôle*.

Le système doit archiver les résultats de contrôle en créant pour chaque pièce produite un dossier particulier, où seront enregistrés tous les résultats des contrôles la concernant .

Pour cela, il doit assurer une liaison physique et logicielle avec un serveur informatique où seront enregistrés ces résultats.

Pour chaque résultat de contrôle, on doit avoir accès aux informations suivantes :

- Identification de la pièce contrôlée
- Etape de fabrication
- Date du contrôle
- Rapport de contrôle (on stocke une représentation compacte des résultats du contrôle)
- Mention d'une éventuelle intervention de l'expert

B.4.7 Fonction F7

Rester compatible avec une *généralisation* potentielle.

L'outil de contrôle doit être développé en conservant une bonne capacité de généralisation ; pour cela, on cherche à :

- *Utiliser une architecture modulaire :*

La partie soft du système doit être composée de modules bien séparés, pouvant être potentiellement réutilisés dans d'autres applications.

- *Capitaliser un savoir-faire innovant :*

La réalisation du projet doit permettre de formuler et de tester des concepts innovants, pouvant être utilisés pour répondre à d'autres problématiques.

- *Viser la certification :*

Le système de contrôle doit, dans l'idéal, bénéficier d'une certification des autorités, ce qui permettrait d'étendre son utilisation.

B.4.8 Fonction F8

Fournir des informations sur son *état de fonctionnement* utiles à sa maintenance.

L'outil doit être capable de contrôler certains de ses paramètres internes et d'en tirer une information utile à sa maintenance et à son utilisation.

Il doit mettre cette information à disposition de l'utilisateur selon sa qualité :

- *Pour les utilisateurs normaux :*

Le système doit proposer une fonction simple qui permet de vérifier qu'il a les capacités de réaliser un contrôle correct (c'est-à-dire un contrôle qui satisfasse la fonction F1).

- *Pour les utilisateur experts :*

Le système doit proposer, en plus de la fonction précédente, une fonction fournissant un diagnostic précis de chacun de ses composants et permettant de changer certains paramètres les concernant.

- *Pour les utilisateurs administrateurs :*

Le système doit proposer ,en plus de la fonction précédente, une fonction fournissant un accès à tous ses paramètres internes connus et permettant un paramétrage total de ses composants.

B.4.9 Fonction F9

Etre robuste à l'*environnement* (propreté, luminosité, vibrations, électricité statique, température, humidité).

Le système de contrôle doit être capable d'avoir un comportement stable et pertinent dans un environnement dont des facteurs physiques sont changeants :

- *Propreté* :

Le système doit pouvoir fonctionner dans un lieu de type classe 100 000.

D'autre part, il ne doit pas générer trop de pollution pour être compatible avec un lieu de type classe 100.

- *Luminosité* :

Le système doit fonctionner dans un endroit d'une luminosité assez importante (80 lux environ) et inconstante (possibilité de changement brutal de cette valeur).

- *Température* :

Le système doit fonctionner pour des températures variant de 15 à 30 c°.

- *Vibrations* :

Le système doit fonctionner dans un milieu de type atelier, soumis à des vibrations assez importantes.

- *Electricité statique* :

Le système doit fonctionner dans un milieu de type atelier, non-exempt d'électricité statique.

- *Humidité* :

Le système doit fonctionner dans un milieu dont l'humidité de l'air ambiant est comprise entre 40% et 60%.

Capteurs pour la détection et la caractérisation des défauts d'aspect

Sommaire

| | |
|---|------------|
| C.1 Capteurs permettant la détection et la caractérisation des défauts | 210 |
| C.1.1 Capteurs à acquisition séquentielle | 211 |
| C.1.1.1 La méthode de focalisation dynamique classique . . . | 211 |
| C.1.1.1.1 Principe de la mesure | 211 |
| C.1.1.1.2 Réalisation technologique | 211 |
| C.1.1.1.3 Avantages et inconvénients | 213 |
| C.1.1.2 La méthode de focalisation dynamique par codage chromatique | 214 |
| C.1.1.2.1 Principe de la mesure | 214 |
| C.1.1.2.2 Réalisation technologique | 214 |
| C.1.1.2.3 Avantages et inconvénients | 215 |
| C.1.1.3 Conclusion | 215 |
| C.1.2 Capteurs à acquisition parallèle | 216 |
| C.1.2.1 La microscopie optique | 216 |
| C.1.2.1.1 Eclairage Bright Field | 216 |
| C.1.2.1.2 Eclairage Dark Field | 218 |
| C.1.2.1.3 Mode Nomarski | 220 |

| | | |
|------------|--|------------|
| C.1.2.1.4 | Photogrammétrie | 222 |
| C.1.2.2 | La microscopie interférométrique | 223 |
| C.1.2.2.1 | Principe de la mesure | 223 |
| C.1.2.2.2 | Réalisation technologique | 224 |
| C.1.2.2.3 | Avantages et inconvénients | 226 |
| C.1.2.3 | Conclusion | 226 |
| C.2 | Capteurs permettant seulement la détection des défauts . . . | 227 |
| C.2.1 | Capteurs sensibles à la dynamique | 227 |
| C.2.1.1 | L'interférométrie holographique par double exposition | 228 |
| C.2.1.1.1 | Principe de la mesure | 228 |
| C.2.1.1.2 | Réalisation technologique | 228 |
| C.2.1.1.3 | Avantages et inconvénients | 229 |
| C.2.1.2 | Autres techniques | 230 |
| C.2.1.3 | Conclusion | 230 |
| C.2.2 | Capteurs sensibles aux caractéristiques statiques | 231 |
| C.2.2.1 | La strioscopie | 231 |
| C.2.2.1.1 | Principe de la mesure | 231 |
| C.2.2.1.2 | Réalisation technologique | 231 |
| C.2.2.1.3 | Avantages et inconvénients | 233 |
| C.2.2.2 | La méthode d'illumination avec pose prolongée | 233 |
| C.2.2.2.1 | Principe de la mesure | 233 |
| C.2.2.2.2 | Réalisation technologique | 233 |
| C.2.2.2.3 | Avantages et inconvénients | 235 |
| C.2.2.3 | Les techniques de diffusométrie | 235 |
| C.2.2.3.1 | Principe de la mesure | 235 |
| C.2.2.3.2 | Réalisation technologique | 235 |
| C.2.2.3.3 | Avantages et inconvénients | 236 |
| C.2.2.4 | Conclusion | 237 |
| C.3 | Tableau récapitulatif des principales caractéristiques des méthodes | 237 |
| C.4 | Annexes : Techniques non retenues | 239 |

| | | |
|------------|---|------------|
| C.4.1 | La microscopie acoustique | 239 |
| C.4.1.1 | Principe de la mesure | 239 |
| C.4.1.2 | Réalisation technologique | 239 |
| C.4.1.3 | Avantages et inconvénients | 240 |
| C.4.1.4 | Conclusion | 241 |
| C.4.2 | La mesure par réflexion | 241 |
| C.4.2.1 | Principe de la mesure | 241 |
| C.4.2.2 | Réalisation technologique | 241 |
| C.4.2.3 | Avantages et inconvénients | 243 |
| C.4.2.4 | Conclusion | 243 |
| C.4.3 | La technique IIM (Imaging Interferometric Microscopy) | 243 |
| C.4.3.1 | Principe de la mesure | 243 |
| C.4.3.2 | Réalisation technologique | 244 |
| C.4.3.3 | Avantages et inconvénients | 244 |
| C.4.3.4 | Conclusion | 245 |
| C.5 | Bibliographie | 246 |

Ce document a pour but de recenser les techniques qui paraissent les plus adaptées à la détection et à la caractérisation des défauts d'aspect sur les pièces optiques, et d'en présenter les principales caractéristiques avec pour but ultime de permettre le choix du ou des capteurs constituant la base du système de contrôle automatique d'aspect des pièces optiques.

Elles se répartissent en deux catégories : les méthodes ne pouvant servir qu'à la détection seule (qui n'apportent pas suffisamment d'information sur le défaut considéré pour sa caractérisation) d'une part, les méthodes permettant la caractérisation des défauts (et par conséquent également leur détection) d'autre part. On suivra cette dichotomie dans la suite de ce document, où seront présentés pour chaque technique abordée :

- le principe de la mesure
- la réalisation technique de celle-ci
- les spécifications d'un produit commercial dérivé de cette technique
- les contraintes de mesure
- ses avantages et ses inconvénients
- une application à un contrôle qualité

On ne fournira cependant pas ici les descriptions des montages optiques les plus classiques et on se limitera aux réalisations techniques les plus performantes

Il est également à noter que certaines de ces techniques ne sont sûrement pas adaptées à l'application que nous souhaitons développer (contrôle d'aspect sur des pièces optiques de grande taille). Elles sont néanmoins décrites ici en Annexes car elles illustrent une idée ou une réalisation technique potentiellement intéressante.

C.1 Capteurs permettant la détection et la caractérisation des défauts

Il s'agit de capteurs fournissant une information assez précise pour permettre la caractérisation du défaut considéré ; ces capteurs permettent à fortiori également la détection des défauts. On peut distinguer dans cette catégorie deux types de capteurs : les capteurs à acquisition séquentielle, où le système mesure point par point la surface, et les capteurs

à acquisition parallèle, où de nombreux points sont acquis simultanément. Le Tableau C.1 ci-dessous, inspiré de (Rioufreyt 1990) permet une rapide comparaison de ces méthodes.

| Acquisition séquentielle | Acquisition parallèle |
|---|--|
| Acquisition lente | Acquisition rapide |
| Éléments mécaniques mobiles fondamentaux | Peu d'éléments mécaniques |
| Grande dynamique de mesure | Dynamique de mesure restreinte |
| Peu de travail algorithmique de restitution | Complexité algorithmique de la restitution |

Table C.1 : Comparaison entre capteurs à acquisition séquentielle et capteurs à acquisition parallèle.

C.1.1 Capteurs à acquisition séquentielle

Ces capteurs sont capables de fournir les coordonnées 3D d'un point de l'objet étudié. En répétant la mesure selon un tramage donné, on reconstruit un profil en trois dimensions de l'objet considéré.

C.1.1.1 La méthode de focalisation dynamique classique

C.1.1.1.1 Principe de la mesure

Il est dérivé du principe de fonctionnement d'un microscope confocal. On cherche à mesurer les écarts de mise au point de l'objectif de focalisation sur la surface de l'objet étudié : un photodétecteur placé derrière une optique confocale acquiert un signal qui est maximum lorsque cette surface est dans le plan focal du système. Ce système optique agit donc comme un « palpeur » donnant ainsi les coordonnées $(x; y; z)$ d'un point de l'objet considéré.

C.1.1.1.2 Réalisation technologique

Le schéma optique d'une telle réalisation est décrit dans la Figure C.1.

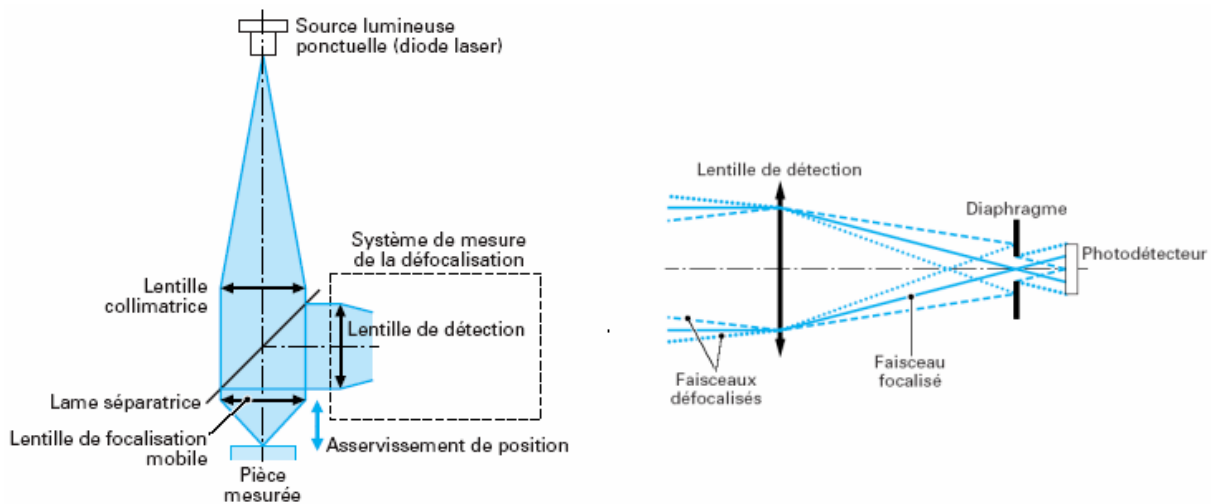


Figure C.1 : Principe de la microscopie confocale, extrait de (Charron 2004b)

Le système de mesure (Charron 2004b) est composé d'une source lumineuse ponctuelle, d'une lentille de collimation qui permet de réaliser un faisceau parallèle, d'une lentille de focalisation (objectif à grande ouverture numérique) qui permet d'éclairer la pièce à mesurer avec un point lumineux très petit (sa taille est voisine de la résolution de l'objectif). La lentille de focalisation est aussi utilisée pour voir la lumière renvoyée par la pièce à mesurer. Une lame séparatrice permet de diriger la lumière réfléchie par la pièce vers un dispositif qui permet de mesurer les écarts de mise au point de l'objectif de focalisation. La lentille de détection focalise la lumière réfléchie par la pièce sur un trou dont le diamètre correspond au diamètre minimal de l'image de la lumière réfléchie lorsque la pièce mesurée est au foyer de la lentille de focalisation. Un photodétecteur est placé derrière le trou, la lumière reçue est maximale lorsque la pièce mesurée est au foyer de la lentille de focalisation. Ainsi, le signal acquis par le photodétecteur est maximum lorsque la surface de l'objet étudié est dans le plan focal. On translate le capteur en Z grâce à un mécanisme piézoélectrique pour détecter ce maximum. En effet, l'intensité I de la lumière détectée ne dépend, pour un point de mesure donné, que de l'altitude z du capteur : $I = f(z)$. En traduisant le capteur de façon à ce que le gradient $\frac{dI}{dz}$ soit positif, on trouve le maximum (lorsque $\frac{dI}{dz} = 0$) : on a alors l'altitude z du point $(x; y)$ étudié. On associe à ce capteur une table motorisée en X et Y qui permet le maillage. La reconstruction tridimensionnelle finale de la surface est effectuée par un ordinateur. Il est à noter que la détection de la focalisation peut également être réalisée par une

méthode utilisant l'astigmatisme d'une lentille ou encore un prisme (Charron 2004b).

Spécifications techniques :

- Résolution en Z : dépend de la qualité de l'optique confocale utilisée et de l'élément piézoélectrique utilisé pour le déplacement en Z. Les meilleures réalisations atteignent une précision de l'ordre du nm.
- Résolution en X et Y : dépend directement de la précision de la table utilisée (peut être améliorée par un traitement informatique et une bonne calibration). Les meilleures réalisations atteignent une précision de l'ordre du micron.
- Vitesse d'acquisition : de l'ordre de 1kHz.

Contraintes de mesure :

- Mesure sur une pièce relativement propre
- Distance de travail très petite (de l'ordre du mm)

Les spécifications techniques fournies ici correspondent à l'appareil Altisurf 500 commercialisé par la société COTEC France.

C.1.1.1.3 Avantages et inconvénients

Avantages :

- Fournit une information en 3D (riche en vue de l'interprétation)
- Bonne résolution en X, Y, excellente en Z
- Relative insensibilité aux vibrations, aux écarts de température
- Méthode utilisable sur tout type de pièces

Inconvénients :

- Grande lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m^2 avec un maillage de $1\mu\text{m}$: temps de mesure de l'ordre du mois!)
- Coût élevé
- Distance de travail très courte donc bonne localisation de la pièce nécessaire

C.1.1.2 La méthode de focalisation dynamique par codage chromatique

C.1.1.2.1 Principe de la mesure

Une lentille entachée de chromatisme projette sur la surface à étudier un faisceau issu d'une lampe à large bande spectrale. Le point de focalisation dépendant de la longueur d'onde, on peut déduire une information sur l'altitude du point mesuré en étudiant le spectre de la lumière rétro-diffusée.

C.1.1.2.2 Réalisation technologique

Une lentille à aberration chromatique programmée transforme la lumière blanche (polychromatique) émise en une suite d'images monochromatiques représentant le champ de mesure (Figure C.2). A chaque image monochromatique correspond une distance depuis la référence du système optique : $d = f(\lambda)$.

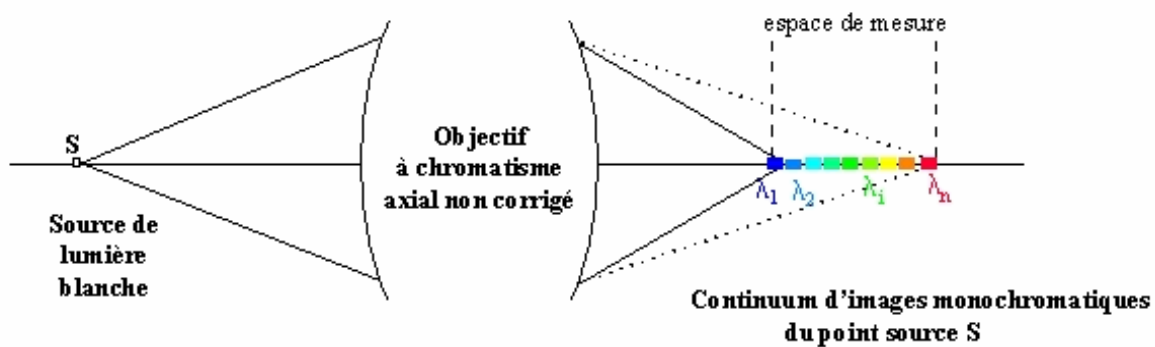


Figure C.2 : Principe du codage chromatique, extrait de (STIL 2005)

On connaît la fonction f grâce à une calibration précise de l'appareil. La surface à mesurer réfléchit une longueur d'onde spécifique correspondant à sa distance. La lumière réfléchie est alors guidée à travers une structure optique et un trou d'aiguille vers un spectrographe qui identifie la longueur d'onde réfléchie et par là même la distance à la surface de l'objet. Un montage complet est présenté dans (Cha et al. 2000).

On associe à ce système un mécanisme de scan basé sur des micromiroirs (Roblin 1999b) classiquement employé dans les montages de microscopie laser à balayage. Il permet d'acquérir, sans mouvements d'ensemble de l'appareil un champ de mesure de surface variable en un temps court. La reconstruction finale du profil de la pièce étudiée est effectuée par un ordinateur.

Spécifications techniques :

- Résolution en Z : de l'ordre de la dizaine de nm.
- Résolution en X et Y : résolution limite de $0.3\mu\text{m}$; dépend du mode d'acquisition de l'image et de la surface étudiée
- Vitesse d'acquisition : de l'ordre de 100kHz. ; dépend du mécanisme de scan

Contraintes de mesure :

- Mesure sur une pièce relativement propre
- Distance de travail très petite (de l'ordre du mm)

Les spécifications techniques fournies ici correspondent à l'appareil optoNCDT 2400 commercialisé par la société μ Epislon (cette appareil ne permet pas de scan en X, Y).

C.1.1.2.3 Avantages et inconvénientsAvantages :

- Fournit une information en 3D (riche en vue de l'interprétation)
- Très bonne résolution en X, Y, excellente en Z
- Relative insensibilité aux vibrations, aux écarts de température
- Méthode utilisable sur tout type de pièces

Inconvénients :

- Lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m^2 avec un maillage de $1\mu\text{m}$: temps de mesure de l'ordre de la journée!)
- Coût élevé
- Distance de travail très courte donc bonne localisation de la pièce nécessaire

C.1.1.3 Conclusion

Le principal avantage de ces méthodes est le fait qu'elles fournissent une information en trois dimensions, particulièrement intéressante du point de vue de l'interprétation des données. D'autre part, ces capteurs sont en général très précis. Cependant, la précision en X et Y est directement dépendante de la mécanique qui donne le maillage des mesures sur la pièce. Cette technique de maillage limite également le traitement de petits défauts

(par rapport à la taille caractéristique de ce maillage) présentant de grandes pentes. Mais l'inconvénient majeur est qu'il est impossible de mener de telles mesures sur de grandes surfaces car les temps de mesure sont prohibitifs.

Ces capteurs sont donc potentiellement intéressants pour l'application que nous cherchons à développer dans la mesure où on les utiliserait sur des champs très petits, ce qui permettrait de jouir de la précision et de la richesse de l'information fournie, sans pâtir de la lenteur de la mesure.

C.1.2 Capteurs à acquisition parallèle

Contrairement aux capteurs présentés précédemment, ces capteurs acquièrent en une mesure, une image ou une pseudo-image correspondant à un champ équivalent à un grand nombre de mesures ponctuelles (typiquement de l'ordre de 10^6).

C.1.2.1 La microscopie optique

On ne détaillera pas ici le principe bien connu du microscope (Flewitt and Wild 1994) mais on va plutôt s'attacher aux différents modes d'éclairage qu'il permet. Dans le cadre d'une application industrielle, on recueille les signaux fournis par ce type d'appareillage grâce à des capteurs de type CCD ou CMOS. Ces capteurs d'imagerie fixent la résolution de la mesure (limitée par la taille et le nombre de cellules du capteur) alors que les caractéristiques du montage optique influent sur la taille du champ étudié mais également sur la résolution de la mesure (limitée par la diffraction).

Les données fournies en ce qui concerne la vitesse d'acquisition d'un champ et l'acquisition sur une pièce d'un m^2 sont tirées d'expériences menées sur un microscope Olympus BX-52 équipé d'une platine Corvus.

C.1.2.1.1 Eclairage Bright Field

1. Principe de la mesure

On étudie premièrement l'éclairage en fond clair où l'objet est éclairé directement. Les rayons directs, non diffractés par l'objet car traversant ses parties vides ou sans structure, pénètrent dans l'objectif. Ils produisent dans le champ image un fond

clair uniforme auquel se superposent, souvent avec un faible contraste, les images des fins détails formées par les ondes diffractées dans la pupille.

2. Réalisation technologique

Le montage le plus classique est l'éclairage de Köhler (Roblin 1999a), qui permet la reproductibilité des conditions d'utilisation, une répartition homogène de la lumière dans le champ d'observation, une exploitation maximale de l'ouverture numérique de l'objectif employé et une élimination des rayons marginaux et de la lumière diffusée (qui entraîneraient une réduction de la qualité de l'image) (Figure C.3).

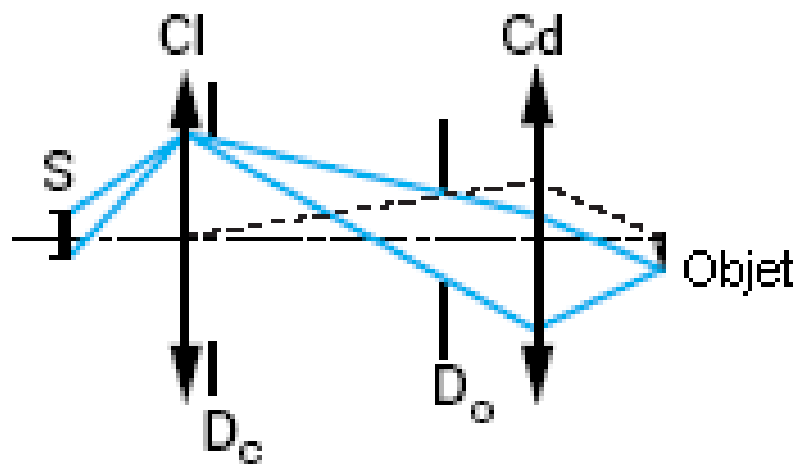


Figure C.3 : Schéma optique de l'éclairage de Köhler, inspiré de (Roblin 1999a)

Spécifications techniques :

- Résolution en X et Y : dépend du grossissement et du moyen d'acquisition (de l'ordre du μm pour un grossissement x50)
- Sensibilité : dépend du grossissement (typiquement la taille minimale des défauts détectés est de $2\mu\text{m}$ en grossissement x50)
- Vitesse d'acquisition : supérieure à 10kHz.
- Champs étudié : dépend du grossissement (typiquement 1mm^2 pour un grossissement x50)

Contraintes de mesure :

- Distance de travail limitée par l'objectif utilisé (de l'ordre du centimètre)

3. *Avantages et inconvénients*

Avantages :

- Rapidité de la mesure d'un champ
- Coût peu élevé
- Relative insensibilité aux vibrations et aux gradients de températures
- Méthode utilisable sur tout type de pièces

Inconvénients :

- Lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m^2 en grossissement $\times 100$: temps de mesure de l'ordre de la semaine !)
- Sensibilité limitée
- Données pas toujours faciles à interpréter
- Nécessité d'un réglage du focus d'autant plus difficile que la profondeur de champ est petite (c'est-à-dire que le grossissement est élevé)

On trouvera un exemple d'application de cette technique au contrôle de semi-conducteurs dans (Mäge et al. 2004).

C.1.2.1.2 **Eclairage Dark Field**

1. *Principe de la mesure*

Dans l'éclairage fond noir, la lumière directe est supprimée et l'on ne recueille que la lumière diffusée par l'objet dont les détails apparaissent sur un fond uniformément sombre. L'objectif ne reçoit que la lumière diffusée par l'objet sous des angles bien supérieurs à ceux qu'il pourrait admettre en usage normal. Les images ne sont pas vraiment conformes, mais des détails plus fins que la limite de résolution sont détectés.

2. *Réalisation technologique*

De nombreux montages existent pour accéder à ce genre d'éclairage (Roblin 1999a). Pour un montage très performant et des indications théoriques sur ce genre de détection se reporter à Tezuka et al. 2003).

Spécifications techniques :

- Résolution en X et Y : dépend du grossissement et du moyen d'acquisition (de l'ordre du μm pour un grossissement x50)
- Sensibilité : dépend du grossissement (typiquement la taille minimale des défauts détectés est de $0.2\mu\text{m}$ en grossissement x50)
- Vitesse d'acquisition : de l'ordre du Hz.
- Champs étudié : dépend du grossissement (typiquement 1mm^2 pour un grossissement x50)

Contraintes de mesure :

- Mesure sur une pièce relativement peu diffusante
- Distance de travail limitée par l'objectif utilisé (de l'ordre du centimètre)

3. *Avantages et inconvénients*Avantages :

- Coût relativement peu élevé
- Très bonne sensibilité
- Données faciles à traiter (blanc sur noir)
- Relative insensibilité aux vibrations et aux gradients de température

Inconvénients :

- Grande lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m^2 en grossissement x50 : temps de mesure de l'ordre de la semaine!)
- Manque de richesse des données (taille difficile à connaître, déformation. . .)
- Sensibilité liée à la nature de l'objet mesuré
- Nécessité d'un réglage du focus d'autant plus difficile que la profondeur de champ est petite (c'est-à-dire que le grossissement est élevé)

On trouvera un exemple d'application de cette technique au contrôle de Blank EUVL dans (Jcong et al. 2000), et au contrôle de céramique dans (Bakolias and Forrest 1997).

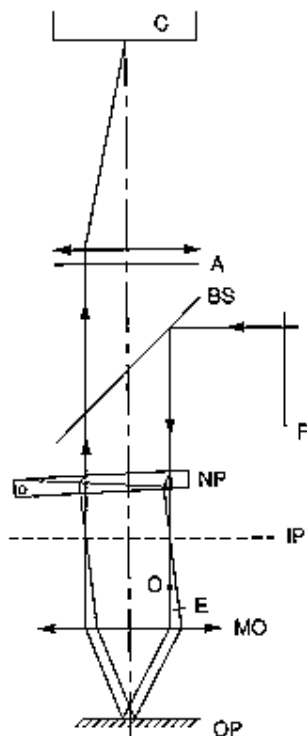
C.1.2.1.3 Mode Nomarski

1. *Principe de la mesure*

L'imagerie en mode Nomarski est une technique d'interférométrie différentielle (Bouchareine 1999) : après dédoublement latéral de l'onde incidente en deux ondes subissant l'influence de l'objet, leur recombinaison fournit un interférogramme mettant en évidence la différence de chemin optique (on visualise en fait un gradient de longueur de chemin optique).

2. *Réalisation technologique*

On produit deux images juxtaposées d'une source lumineuse ponctuelle en jouant avec deux directions orthogonales de polarisation de l'onde électromagnétique (en utilisant un prisme de Wollaston par exemple). Les interférences sont possibles en repolarisant les deux ondes dans le même plan grâce à un analyseur. Le schéma d'un tel montage optique est donné dans la Figure C.4.



Où

- OP désigne le plan objet
- MO, l'objectif du microscope
- IP, le plan d'intersection des rayons
- NP, le prisme de Nomarski
- BS, le cube séparateur
- P, le polariseur
- A, l'analyseur,
- et C, le plan image

Figure C.4 : Schéma optique d'un microscope DIC, extrait de (Chatterjee 2003)

Finalement, on obtient une image en pseudo 3D, une différence de teinte ou de niveau de gris correspondant à une différence de pente. En effet, si le dédoublement d est effectué dans la direction de x , en tout point introduisant une variation de chemin optique $\delta(x)$, la différence de marche $\Delta(x)$ entre les deux ondes est (Roblin 1999a) :

$$\Delta(x) = \delta(x + d/2) - \delta(x - d/2) \approx d\delta'(x)$$

avec $\delta'(x)$, le gradient en x de $\delta(x)$.

Or, à l'aide de certaines approximations, on montre que (Holmes and Levy 1987) :

$$I(x, y) = 2a^2[1 + \Delta(x)]$$

où $I(x, y)$ désigne l'intensité lumineuse observée à travers le microscope DIC, et a est une constante dépendant du microscope. Ainsi, l'intensité lumineuse détectée dépend linéairement du gradient en x du chemin optique au point considéré, c'est à dire de la pente en x du relief de la pièce au point considéré.

Pour des indications théoriques plus complètes, se rapporter à (Lessor et al. 1979).

Spécifications techniques :

- Résolution en X et Y : dépend du grossissement et du moyen d'acquisition (de l'ordre du μm pour un grossissement x50)
- Sensibilité : dépend du grossissement ; dans (Holmes and Levy 1987), on précise que l'on a accès à une information dépassant la limite de résolution optique classique ; en Z, de l'ordre de la dizaine de nm.
- Vitesse d'acquisition : de l'ordre de 5 kHz.
- Champs étudié : dépend du grossissement (typiquement 1mm^2 pour un grossissement x50)

Contraintes de mesure :

- Distance de travail limitée par l'objectif utilisé (de l'ordre du centimètre)

3. *Avantages et inconvénients*

Avantages :

- Coût relativement peu élevé
- Rapidité de la mesure d'un champ
- Données riches en information (pseudo 3D)
- Relative insensibilité aux vibrations et aux gradients de température

Inconvénients :

- Lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m² en grossissement x100 : temps de mesure de l'ordre de la semaine!)
- Données pas toujours aisées à interpréter
- Nécessité d'un réglage du focus d'autant plus difficile que la profondeur de champ est petite (c'est-à-dire que le grossissement est élevé)

Les images Nomarski ne fournissent que des informations qualitatives sur la profondeur. Plusieurs techniques (Li et al. 2001), (Arnison et al. 2004) ont été mises au point dans le but de reconstruire des cartes 3D à partir de telles images. Ces techniques, basées sur le concept de phase-shift nécessitent peu d'aménagement du montage de base et peuvent s'avérer intéressantes pour obtenir des mesures quantitatives de profondeur.

C.1.2.1.4 Photogrammétrie

On cite rapidement ici la méthode de photogrammétrie qui consiste à se servir de la prise d'images sous différents angles de vue pour produire une reconstruction tridimensionnelle de l'objet (Charron 2004b). On peut appliquer cette méthode à des images provenant de prises de vues microscopiques. A partir de deux vues stéréoscopiques de l'objet étudié, on obtient une information d'altitude Z par triangulation (Figure C.5).

Les coordonnées x, y, z du point mesuré sont données par les équations suivantes (Charron 2004b) :

$$\left. \begin{aligned} x &= \frac{d \tan(\Psi + \gamma)}{\tan(\Phi + \alpha) - \tan(\Psi + \gamma)} \\ y &= \frac{d}{\tan\left(\Phi + \alpha + \frac{\pi}{2}\right) - \tan\left(\Psi + \gamma + \frac{\pi}{2}\right)} \\ z &= \sqrt{x^2 + y^2} \tan \beta \end{aligned} \right\}$$

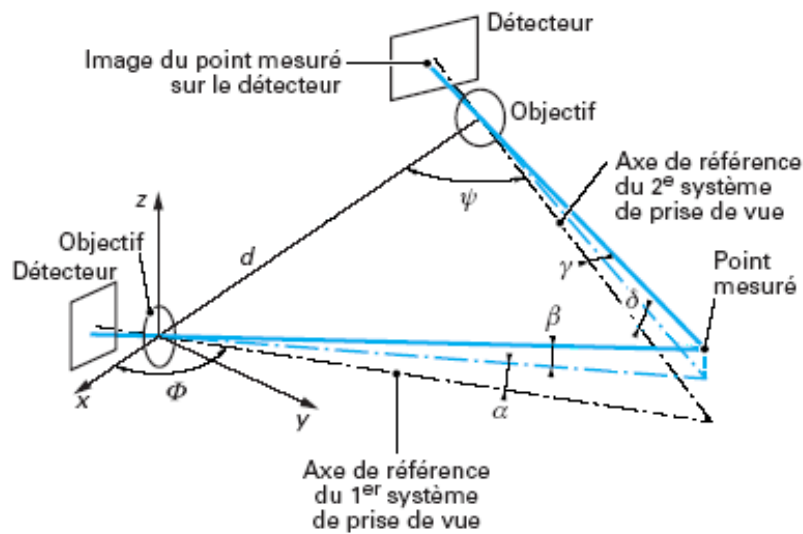


Figure C.5 : Principe de la photogrammétrie. inspiré de (Charron 2004b)

Les montages de photogrammétrie microscopique donne une précision relative (par rapport à la plus grande dimension de l'objet) de l'ordre de 10^{-3} . On peut donc obtenir avec ce genre de montage, sur des champs microscopiques, une information en Z de l'ordre de la centaine de nanomètres.

La principale difficulté de cette méthode est de localiser précisément des points singuliers de l'objet sur les différentes vues. Malheureusement cette mise en correspondance des points homologues (qui doit être extrêmement sûre) constitue un problème difficile ; certains algorithmes existent (Szeliski and Zabih 1999) mais ils sont lourds et pas toujours fiables. Le brevet (Spoto et al. 2002) propose une architecture physique implémentant la reconstruction rapide d'une carte d'altitude à partir de deux images stéréoscopiques.

C.1.2.2 La microscopie interférométrique

C.1.2.2.1 Principe de la mesure

Le principe est le même que celui de l'interférométrie classique : on compare le front d'onde engendré par la pièce à étudier à celui produit par un miroir de référence en les faisant interférer. En utilisant une méthode de décalage de phase (Phase Shifting) ou de recherche de frange (Peak Fringe Scanning), on reconstruit la cartographie 3D de la pièce étudiée.

C.1.2.2.2 Réalisation technologique

Pour ce genre de montage microscopique, on utilise généralement un interféromètre de Mirau (Figure C.6) ou de Michelson (Figure C.7) (Charron 2004b) :

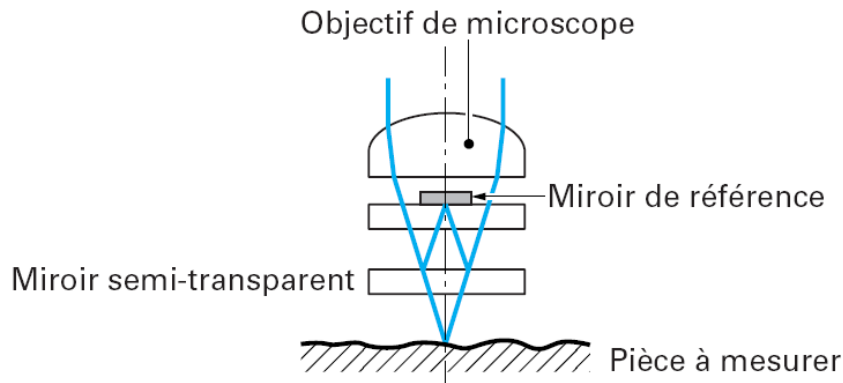


Figure C.6 : Interféromètre de Mirau, extrait de (Charron 2004b)

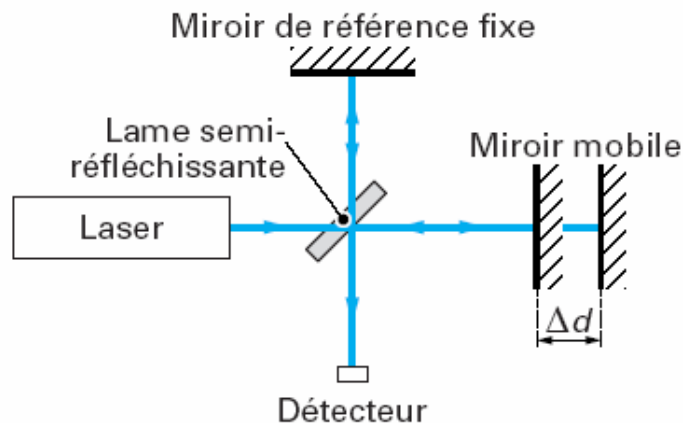


Figure C.7 : Interféromètre de Michelson, extrait de (Charron 2004b)

Le déplacement en Z qui permet le décalage de phase ou la recherche des franges en lumière blanche est donné par un mécanisme piézoélectrique. Les images des franges sont acquises par un capteur CCD puis transmises à un ordinateur qui effectue les calculs de phase (la résolution de la mesure est donc limitée par la taille et le nombre de cellules du capteur). On associe à ce montage une table motorisée en X et Y qui permet d'effectuer des mesures sur de grandes surfaces.

La méthode de décalage de phase utilise une source de lumière monochromatique et cohérente, et permet l'analyse tridimensionnelle de défauts peu profonds ($< 2 \mu\text{m}$ de profon-

deur). L'idée principale de cette méthode (Benatmane 2002) est de déterminer, grâce aux équations de l'interférométrie, l'altitude du relief d'un échantillon par l'intermédiaire de la phase d'un signal lumineux. Ceci peut être réalisé en introduisant des décalages de phase connus entre les deux faisceaux provenant respectivement de l'échantillon et du miroir de référence. Le support piézoélectrique permet de déplacer verticalement l'échantillon dans des positions correspondant à des sous-multiples de la longueur d'onde, et trois images ou plus sont ainsi acquises.

L'intensité en un point de coordonnées $(x; y)$ dans une figure d'interférence pouvant être exprimée par :

$$I(x, y) = I_0(x, y)[1 + \gamma_0 \cos(\Phi(x, y) + \alpha)]$$

avec α le décalage de phase introduit, I_0 le carré de l'amplitude incidente, γ_0 le facteur de visibilité des franges, et $\phi(x, y)$ la phase du signal qu'on cherche à calculer, l'acquisition d'au moins trois interférogrammes avec des déphasages contrôlés permet de déterminer les trois inconnues $I_0(x, y)$, γ_0 , et $\Phi(x, y)$.

La méthode de recherche de franges, quant à elle, utilise une source de lumière blanche (incohérente) et permet la mesure de structures de quelques microns voire d'une dizaine de microns de hauteur. L'idée maîtresse de cette technique est de détecter la frange centrale d'ordre zéro (Montgomery and Fillard 1992). La cohérence temporelle de la source étant limitée, on observe en effet une « enveloppe » d'interférence confinée à la surface de la pièce étudiée. La mesure du déplacement en Z nécessaire à la visualisation de cette frange centrale donne une mesure absolue de l'altitude du point $(x; y)$ de la surface correspondant.

Spécifications techniques :

- Résolution en Z : de l'ordre de l'Angström.
- Résolution en X et Y : résolution limite de l'optique « conventionnelle » : environ $0,5 \mu\text{m}$; dépend du grossissement et du moyen d'acquisition
- Vitesse d'acquisition : de l'ordre de 2s pour imager un champ pour les meilleures réalisations.
- Champs étudié : dépend du grossissement (typiquement 1mm^2 pour un grossissement x50)

Contraintes de mesure :

- La pièce contrôlée doit être réfléchissante (une réflectivité de 0.5% suffit cependant pour les meilleures réalisations)
- Méthode sensible aux vibrations et dans une moindre mesure aux gradients de température
- Distance de travail limitée par l'objectif utilisé (de l'ordre du centimètre)

Les spécifications techniques fournies ici correspondent à l'appareil Talysurf CCI 3000 commercialisé par la société Taylor Hobson, à l'appareil ZoomSurf 3D commercialisé par la société Fogale Nanotech France et à l'appareil Wyko NT 8000 commercialisé par la société Veeco.

C.1.2.2.3 Avantages et inconvénients

Avantages :

- Fournit une information en 3D (plus riche en vue de l'interprétation)
- Très bonne résolution en X, Y, excellente en Z

Inconvénients :

- Lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m^2 en grossissement $\times 25$: temps de mesure de l'ordre de la semaine!)
- Coût assez élevé
- Réglages difficiles à automatiser (focus, tilts...)
- Sensibilité aux vibrations

On trouvera un exemple d'application de cette technique au contrôle de MEMS dans (Serry et al. 2004).

C.1.2.3 Conclusion

Le principal avantage de ces méthodes est le fait qu'elles fournissent en une mesure rapide, une information sur un champ de mesure étendu (typiquement de l'ordre du mm^2 pour des grossissements moyens). D'autre part, ces capteurs sont en général précis. Cependant, la définition en X et Y est directement dépendante du grossissement employé

qui fixe la grandeur du champ mesuré. Ces méthodes nécessitent également des réglages qui peuvent être délicats à mettre en œuvre automatiquement. Mais l'inconvénient majeur est qu'il semble difficile de mener de telles mesures sur de grandes surfaces, ce qui nécessiterait des temps de mesure conséquents.

Ces capteurs sont donc potentiellement intéressants pour l'application que nous cherchons à développer à deux titres : premièrement, ils sont tout à fait adaptés à l'étude de champs de petite taille puisqu'ils peuvent en donner rapidement une image précise ; d'autre part, on pourrait mener ce genre de mesure sur des pièces entières en augmentant significativement la vitesse du scan (le temps le plus limitant n'étant pas le temps de mesure mais le temps de déplacement du capteur par rapport à la pièce) .

Enfin, il est à noter que chaque méthode présente un avantage qui lui est propre, le plus intéressant étant sans doute la possibilité pour certaines d'avoir accès à une information tridimensionnelle.

C.2 Capteurs permettant seulement la détection des défauts

Il s'agit de capteurs permettant la détection des défauts recherchés mais qui ne fournissent pas assez d'information pour permettre leur caractérisation. On peut distinguer dans cette catégorie deux types de capteurs : les capteurs sensibles à la dynamique, où l'on mesure une variation de la surface de l'objet, et les capteurs sensibles aux caractéristiques statiques, où l'objet est mesuré directement.

C.2.1 Capteurs sensibles à la dynamique

Contrairement aux techniques présentées précédemment, ces capteurs n'effectuent pas de mesures caractérisant l'état statique de l'objet à étudier mais sont capables de détecter des variations très fines de son état entre deux instants.

C.2.1.1 L'interférométrie holographique par double exposition

C.2.1.1.1 Principe de la mesure

L'idée est de comparer deux hologrammes par interférométrie. Ces deux hologrammes représentant le même objet à deux instants différents, on peut détecter et mesurer les variations de phase survenues entre les deux expositions.

C.2.1.1.2 Réalisation technologique

La technique employée est semblable à celle utilisée pour réaliser un hologramme conventionnel (Smigielski 2001a). On effectue une première exposition ; on applique alors une contrainte sur la pièce à tester, sans toucher à rien d'autre et on effectue alors une seconde pose (Figure C.8). Après développement photographique, on dispose d'une plaque contenant la somme de deux hologrammes, incohérents entre eux puisque réalisés à des instants différents. Cependant, à la restitution (Figure C.9), on obtient deux images cohérentes entre elles, puisque restituées à l'aide d'une même source de lumière cohérente (laser par exemple). Ces images peuvent donc interférer. Les franges d'interférence observées caractérisent la modification subie par l'objet entre les deux poses, c'est-à-dire la déformation due à la contrainte : ce sont les lignes d'isoamplitude de déplacement. Quand on passe d'une frange à l'autre, le déplacement varie de $\lambda/2$. Il existe de nombreuses techniques permettant de traiter de tels interférogrammes.

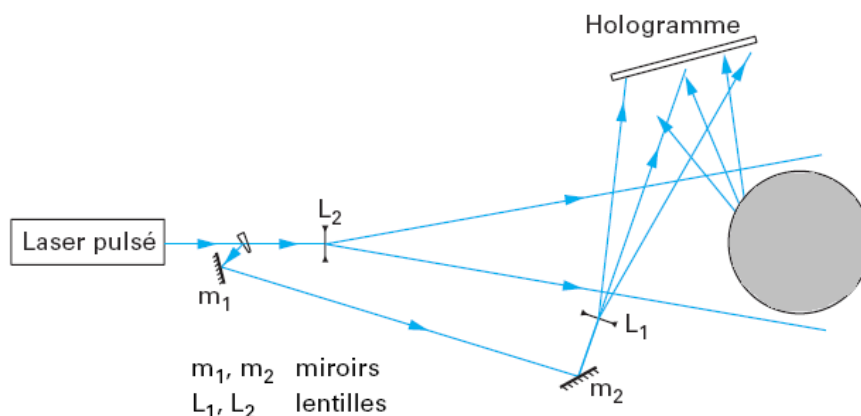


Figure C.8 : Montage d'enregistrement (deux poses successives), extrait de (Smigielski 2001a)

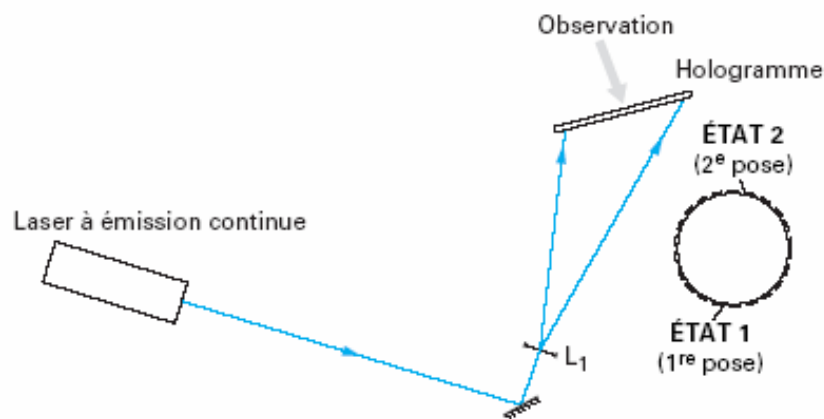


Figure C.9 : Montage de restitution (les deux images interfèrent), extrait de (Smigielski 2001a)

D'une façon générale, la méthode permet de détecter et de mesurer les variations de phase survenues entre les deux expositions. Ces variations de phase peuvent être dues à des variations de longueur, d'indice de réfraction ou de longueur d'onde causées par des contraintes diverses (thermiques, pneumatiques, mécaniques...).

Spécifications techniques :

- Résolution en Z : de l'ordre du nm.
- Résolution en X et Y : dépend du mode d'acquisition de la figure d'interférence et de la surface étudiée
- Vitesse d'acquisition : de l'ordre de la minute pour acquérir et traiter l'interférogramme.

Contraintes de mesure :

- Méthode sensible aux vibrations
- La pièce étudiée doit être suffisamment diffusante

C.2.1.1.3 Avantages et inconvénients

Avantages :

- Fournit une information précise sur un champ étendu
- Coût relativement peu élevé

- Rapidité de la mesure

Inconvénients :

- Méthode peu facile à automatiser
- Ne mesure qu'une réaction à une contrainte
- Méthode sensible aux vibrations
- Méthode applicable à des pièces suffisamment diffusantes

On trouvera des exemples d'application de cette technique au contrôle de matériaux et à l'étude de contraintes dans (Smigielski 2001a).

C.2.1.2 Autres techniques

Il est à noter que cette méthode est proche d'autres méthodes qu'on ne détaillera pas ici : shearographie, interférométrie de speckle, TV-holographie. Ces méthodes mettent en jeu la granularité laser, qui se manifeste sur des objets diffusant et présentant un relief microscopique important ; elles ne sont donc sans doute pas utilisables dans notre cas. Pour plus de renseignements sur ces techniques se reporter à (Honlet 1997), (Plaut-Aubry 2000), (Taillade 2003), et (Smigielski 2001b).

C.2.1.3 Conclusion

Le principal avantage de ces méthodes est le fait qu'elles fournissent en une mesure rapide, une information sur un champ de mesure très étendu. D'autre part, ces mesures sont très précises pour ce qui est du déplacement en Z. Cependant, la précision en X et Y est directement liée au mode d'acquisition et, au vue des larges surfaces à analyser, limitée. La principale difficulté pour le problème de détection est donc de trouver une contrainte adéquate qui mette en relief dynamiquement la présence de défaut, de manière à ce qu'il devienne détectable malgré une faible définition en X, Y. Un inconvénient majeur est également la difficulté d'automatisation de certaines de ces techniques.

Ces capteurs sont cependant potentiellement intéressants pour l'application que nous cherchons à développer. En effet, ils permettraient, sous réserve de trouver une contrainte adéquate, de détecter les défauts sur de grandes surfaces en un temps restreint. Il faut noter

qu'ils peuvent également fournir une information, certes limitée, utile à la caractérisation du défaut.

C.2.2 Capteurs sensibles aux caractéristiques statiques

Les capteurs abordés dans ce chapitre ont la particularité d'effectuer des mesures sur de très grands champs. Leur principe de base est donc d'amplifier le défaut pour qu'il soit détectable macroscopiquement.

C.2.2.1 La strioscopie

C.2.2.1.1 Principe de la mesure

On forme l'image de l'objet à étudier dans le plan de focalisation d'une lentille : cette image est en fait la transformée de Fourier de l'objet (Puech 2002). Ainsi, en disposant un cache au centre de cette image, on effectue un filtrage passe-haut en fréquence : les défauts apparaissent en surbrillance et peuvent facilement être détectés.

C.2.2.1.2 Réalisation technologique

Le montage optique correspondant est donné dans la Figure C.10.

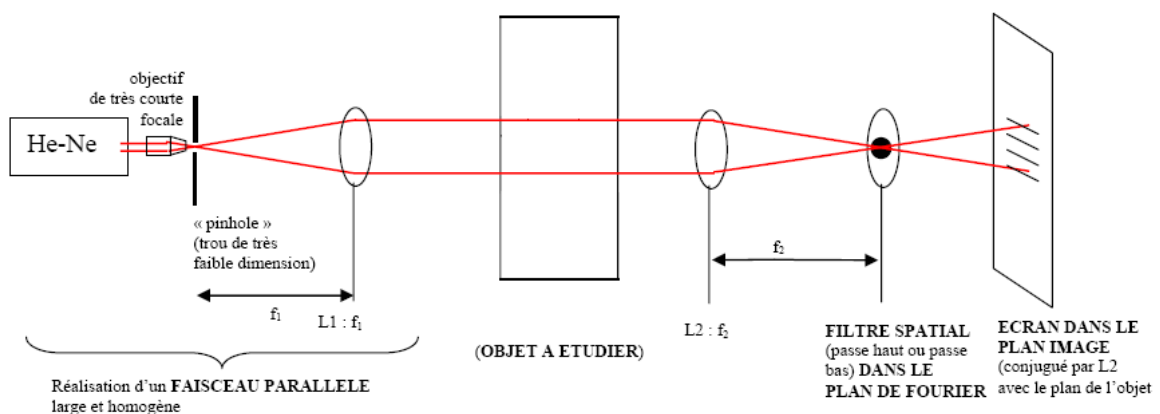


Figure C.10 : Montage optique de strioscopie, extrait de (Zanier 2002)

Le filtrage spatial permet d'éliminer le fond continu (basse fréquence) ; ainsi on observe que la lumière diffusée par les défauts recherchés. Ce filtrage est réalisé en plaçant un obturateur de dimension adaptée ((Gheen et al. 1997) pour une étude de l'impact de

la forme de ce masque) au foyer d'une lentille avant imagerie sur une caméra CCD. Le contrôle peut s'effectuer en transmission ou en réflexion. La mesure se faisant par l'intermédiaire d'un capteur CCD, sa résolution est limitée par celui-ci.

De plus, le déplacement de la caméra dans le plan longitudinal permet l'inspection de différents plans (Bertin 2005) (Figure C.11)

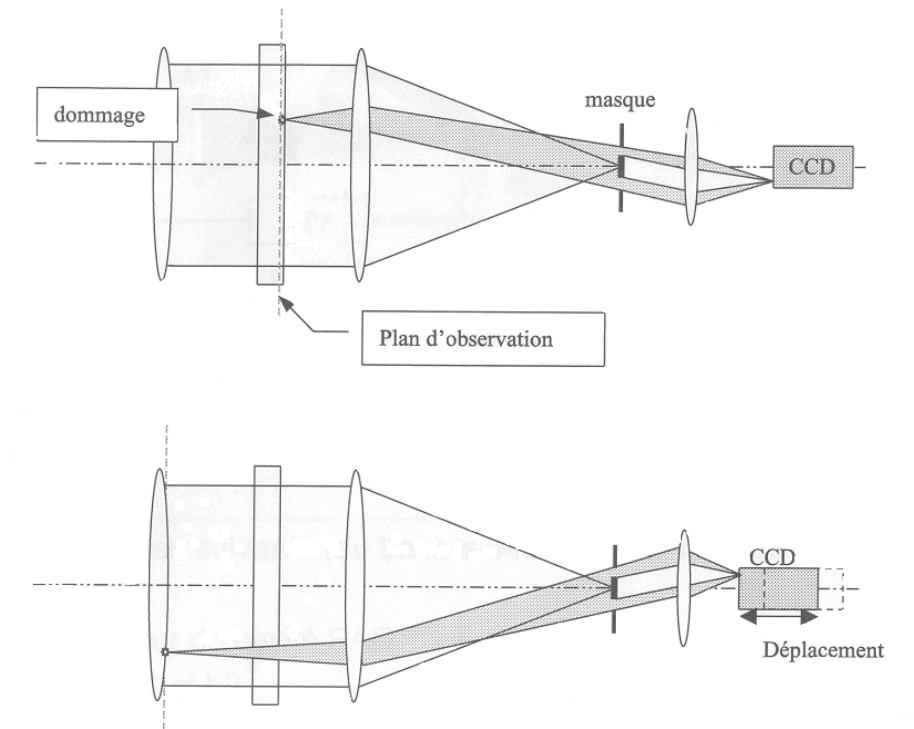


Figure C.11 : Observation de deux plans par déplacement du capteur CCD, extrait de (Bertin 2005)

Il est à noter qu'il existe une évolution de ce montage, où le cache est remplacé par une lame demi-onde (technique de contraste de phase); ce montage permet d'augmenter encore le contraste des défauts sur l'image (Philbert and Gryson 1962).

Spécifications techniques :

- Résolution en X et Y : dépend du mode d'acquisition de l'image et de la surface étudiée.
- Sensibilité : inconnue (à tester).
- Vitesse d'acquisition : de l'ordre de la seconde pour acquérir un champ correspondant

à une pièce entière.

Cette méthode est utilisée par l'appareil de détection MICO commercialisé par la société Bertin et fait également l'objet d'un brevet (Flauger 1996).

C.2.2.1.3 Avantages et inconvénients

Avantages :

- Coût relativement peu élevé
- Rapidité de la mesure
- Champ mesuré étendu
- Grande distance de travail
- Méthode utilisable sur tout type de pièces
- Méthode relativement insensible aux vibrations et aux gradients de température

Inconvénients :

- Sensibilité peut-être insuffisante et inadaptée
- Informations très limitées sur la morphologie des défauts détectés
- Nécessité d'un capteur CCD de très grande taille (donc d'un coût élevé) pour avoir un positionnement correct des défauts sur une grande pièce

C.2.2.2 La méthode d'illumination avec pose prolongée

C.2.2.2.1 Principe de la mesure

Le principe de cette méthode de détection, proposée dans (Rainer 1997), est très simple : on recueille grâce à un capteur CCD le flux diffusé par les défauts recherchés lorsque la pièce est éclairée par la tranche et en incidence rasante.

C.2.2.2.2 Réalisation technologique

Le montage original est donné par la Figure C.12. On utilise une source de lumière blanche pour éclairer la pièce par ses tranches à l'aide de 12 sources linéaires, et en incidence rasante, grâce à un réseau de fibres optiques. L'éclairage par les tranches permet de détecter les défauts de la matière (inclusions, bulles) et les défauts situés sur la surface

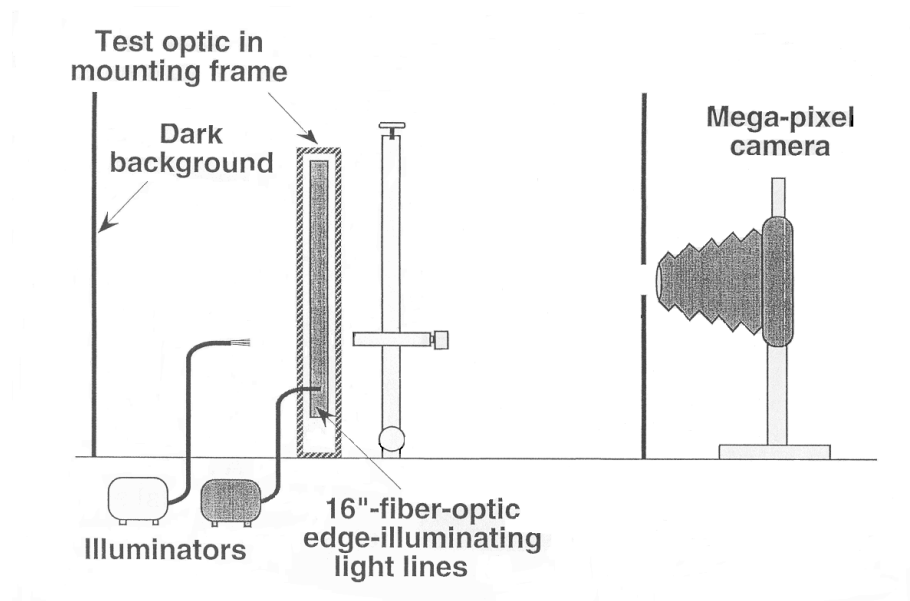


Figure C.12 : Système complet de contrôle de pièces optiques (un microscope monté sur une table XY pour caractériser les défauts détectés est également représenté sur le schéma), extrait de (Rainer 1997)

(piqûres, rayures...). L'éclairage en incidence rasante, permet quant à lui la détection de traces de nettoyage et de défauts de traitement. À noter que cette méthode entraîne également la détection d'artefacts dont il faudra déterminer la nature par un examen plus poussé.

Spécifications techniques :

- Résolution en X et Y : dépend du mode d'acquisition de l'image et de la surface étudiée.
- Sensibilité : inconnue (à tester).
- Vitesse d'acquisition : environ 1 minute pour acquérir un champ correspondant à une pièce entière

Contraintes de mesure :

- La matière dont la pièce est constituée ne doit pas être trop diffusante (sous peine de noyer le signal correspondant aux défauts dans le bruit)
- Le dispositif d'éclairage doit être adapté à la pièce contrôlée et par conséquent n'est pas universel (besoin d'un dispositif dédié pour chaque type de pièce)

C.2.2.2.3 Avantages et inconvénients

Avantages :

- Coût relativement peu élevé (hors CCD)
- Champ mesuré étendu
- Grande distance de travail
- Méthode insensible aux vibrations et aux gradients de température

Inconvénients :

- Sensibilité peut-être insuffisante et inadaptée
- Informations très limitées sur la morphologie des défauts détectés
- Méthode qui perd beaucoup de sa sensibilité sur des pièces traitées
- Nécessité d'un capteur CCD de très grande taille (donc d'un coût élevé) pour avoir un positionnement correct des défauts sur une grande pièce

Cette méthode est utilisée pour le contrôle des plaques amplis du projet américain équivalent au MegaJoule, le NIF et fait également l'objet d'un brevet (Henley and Bryan 1998).

C.2.2.3 Les techniques de diffusométrie

C.2.2.3.1 Principe de la mesure

Ces capteurs sont du type point par point : on teste séquentiellement des petites surfaces de la pièce contrôlée grâce à un flux lumineux intense. Les imperfections de surfaces (défauts de surfaces, poussières. . .) sont simplement détectées par l'action diffusante qu'elles ont sur ce flux.

C.2.2.3.2 Réalisation technologique

Il existe de très nombreux montages appliquant ce principe de base, décliné selon trois philosophies. La première consiste à mesurer directement le flux diffusé par un éventuel défaut ; on utilise alors une parabole intégrante (Chen et al. 2004), ou des montages de type dark-field (Hue et al. 2001). La seconde consiste à mesurer le flux spéculaire (Quan et al. 2000) ; une baisse significative de ce flux trahit en effet la présence d'un défaut.

La troisième consiste à utiliser ces deux approches dans le même montage (Kato et al. 1984),(Okada et al. 1999) et (Imaino and Latta 1999). Ces types de montage peuvent être raffinés par différentes techniques permettant d'améliorer le rapport signal sur bruit, comme par exemple l'utilisation de deux polarisations de l'onde incidente (Lalezari and Knollenberg 1989), d'une détection hétérodyne (Tsuji et al. 1999) ou homodyne (Klooster et al. 2004) du flux diffusé.

Le montage est complété par un mécanisme qui permet le scan du spot de mesure sur la surface à mesurer, utilisant généralement des modulateurs acousto-optiques ou des miroirs (pour un montage très performant, voir (Rangarajan et al. 2004)).

D'autre part, on peut obtenir des informations sur la taille (Alexander and Hallett 1999), (Hue et al. 2003), (Quan et al. 2000), (Sankaran and Stover 1999), ou sur la nature du défaut détecté (Sekine et al. 2000), (Welch 1996), (Okada et al. 1999), (Pladellorens et al. 2002) en étudiant la quantité ou la répartition spatiale de la lumière diffusée et réfléchi.

Spécifications techniques :

- Résolution en X et Y : dépend de la taille du spot de mesure (typiquement de l'ordre de la dizaine de μm).
- Sensibilité : dépend du montage. Cependant, pour des réalisations correctes, elle semble suffisante pour l'application visée (dans (Lorincik et al. 1996), on arrive à détecter des points de $1.5\mu\text{m}$ de diamètre et 40nm de profondeur).
- Vitesse d'acquisition : dépend du mécanisme de scan et de la taille du spot de mesure.

Contraintes de mesure :

- La matière dont la pièce est constituée ne doit pas être trop diffusante ou rugueuse (sous peine de noyer le signal correspondant aux défauts dans le bruit).
- Méthode sensible à la lumière ambiante.

C.2.2.3.3 Avantages et inconvénients

Avantages :

- Méthode relativement insensible aux vibrations et aux gradients de température
- Sensibilité importante

- Possibilité d'obtenir des informations utiles à la caractérisation des défauts
- Méthode utilisable sur tout type de pièces

Inconvénients :

- Nécessité d'un mécanisme de scan complexe et coûteux
- Petit champ de mesure
- Relative lenteur de l'acquisition (typiquement pour couvrir une étendue de 1m^2 pour un spot de $10\mu\text{m} \times 10\mu\text{m}$: temps de mesure de l'ordre de l'heure)
- Distance de travail courte

Cette méthode est utilisée par l'appareil SurfScan P1 commercialisé par la société KLA Tencor. On trouvera un exemple de son application au contrôle de semi-conducteurs dans (Bechtler and Velidandla 2003)

C.2.2.4 Conclusion

Le principal avantage de ces méthodes est le fait qu'elles fournissent en une mesure rapide, une information sur un champ de mesure très étendu. Cependant, la précision en X et Y est directement liée au mode d'acquisition et, au vue des larges surfaces à analyser, limitée. Ceci n'est en fait pas gênant si l'on ne cherche pas à imager le défaut mais juste à le détecter. Le point sensible de ces méthodes est l'amplification des défauts, qui doit être adaptée à ce qu'on cherche à voir et assurer une sensibilité correcte.

Ces capteurs sont donc potentiellement intéressants pour l'application que nous cherchons à développer. En effet, ils permettraient, sous réserve qu'ils aient une sensibilité suffisante, de détecter les défauts sur de grandes surfaces en un temps restreint. Il faut noter qu'ils peuvent également fournir une information, certes limitée, utile à la caractérisation du défaut.

C.3 Tableau récapitulatif des principales caractéristiques des méthodes

| Méthode | D | C | Spécificités | Précision en Z | Définition en X,Y | Coût | Durée d'une acquisition | Champ mesuré |
|---|---|---|--|----------------|----------------------------|------|-------------------------|-----------------|
| Focalisation dynamique conventionnelle | X | X | Sur pièce relativement propre | ≅ 1 nm | ≅ 1 µm | +++ | ≅ 10e-3 s | - |
| Focalisation dynamique par codage chromatique | X | X | Sur pièce relativement propre | ≅ 10 nm | ≅ 0.5 µm | +++ | ≅ 10e-6s | - |
| Microscopie Bright Field | X | X | | - | ≅ 0.5 µm | + | ≅ 10e-4 s | ≅ 500µm X 500µm |
| Microscopie Dark Field | X | X | Pour des matériaux peu diffusant | - | ≅ 0.5 µm | + | ≅ 1 s | ≅ 500µm X 500µm |
| Microscopie Normarski | X | X | Sensibilité en Z en dessous de 10 nm Possibilité d'avoir une mesure quantitative en Z | - | ≅ 0.5 µm | + | ≅ 10e-3 s | ≅ 500µm X 500µm |
| Photogrammétrie microscopique | X | X | Reconstruction lourde algorithmiquement | ≅ 100 nm | ≅ 0.5 µm | + | < 1s | ≅ 500µm X 500µm |
| Microscopie interférométrique | X | X | Sensibilité aux vibrations | ≅ 1 Å | ≅ 0.5 µm | ++ | ≅ 5 s | ≅ 500µm X 500µm |
| Interférométrie holographique | X | | la pièce étudiée doit être suffisamment diffusante | < 1 nm | dépend de la configuration | + | < 5 min | ≅ 1m² |
| Stioscopie | X | | sensibilité inconnue | - | dépend de la configuration | + | < 1 s | ≅ 1m² |
| Illumination avec pose prolongée | X | | sensibilité inconnue | - | dépend de la configuration | + | 1 min | ≅ 1m² |
| Diffusométrie | X | | Possibilité d'avoir des informations utiles à la caractérisation du défaut | - | ≅ 10 µm | ++ | ≅ 10e-6s | ≅ 10µm X 10µm |

Capacité pour la caractérisation = C

Capacité pour la détection = D

En ce qui concerne la microscopie optique, on donne les estimation pour un grossissement x200

C.4 Annexes : Techniques non retenues

C.4.1 La microscopie acoustique

C.4.1.1 Principe de la mesure

Le principe de base de la microscopie acoustique (Attal 1998) est le même que celui de la microscopie optique, l'onde utilisée n'étant plus optique mais acoustique (typiquement des ultrasons avec $\lambda \approx 1\mu\text{m}$). On dispose de deux modes d'études : un mode volumique basé sur la focalisation de l'onde, et un mode surfacique, basé sur l'étude des interférences entre les ondes réfléchies et les ondes de surface, qui fournit une signature acoustique, caractéristique de la structure du matériau employé.

C.4.1.2 Réalisation technologique

Le microscope acoustique comprend un générateur d'ultrasons, constitué par un transducteur piézoélectrique, qui convertit un signal électrique incident en signal acoustique (Figure C.13). Cette couche piézoélectrique est déposée sur l'une des faces planes et polies d'un petit barreau de saphir synthétique préalablement métallisé. Ce barreau de saphir comporte, sur la face plane opposée à celle du transducteur, un petit dioptre sphérique creusé et poli d'un diamètre de l'ordre d'une centaine de μm , qui focalise l'onde acoustique sur l'objet à étudier. Le deuxième milieu constituant le dioptre est un liquide ; le rapport entre les vitesses de l'onde acoustique dans le saphir et dans le liquide étant petit, l'aberration sphérique de focalisation est limitée par rapport à la microscopie optique. Lorsque l'on travaille par réflexion, le système émetteur joue aussi le rôle de récepteur en utilisant l'effet piézoélectrique inverse.

Spécifications techniques :

- Résolution en Z : en dessous du μm . On dispose aussi d'une information plus qualitative grâce au mode surfacique.
- Résolution en X et Y : dépend de la fréquence des ultrasons : environ 1 μm , mais on peut descendre jusqu'à 0.3 μm .
- Vitesse d'acquisition : dépend de la mécanique du scan (environ 10cm/s pour les meilleures réalisations).

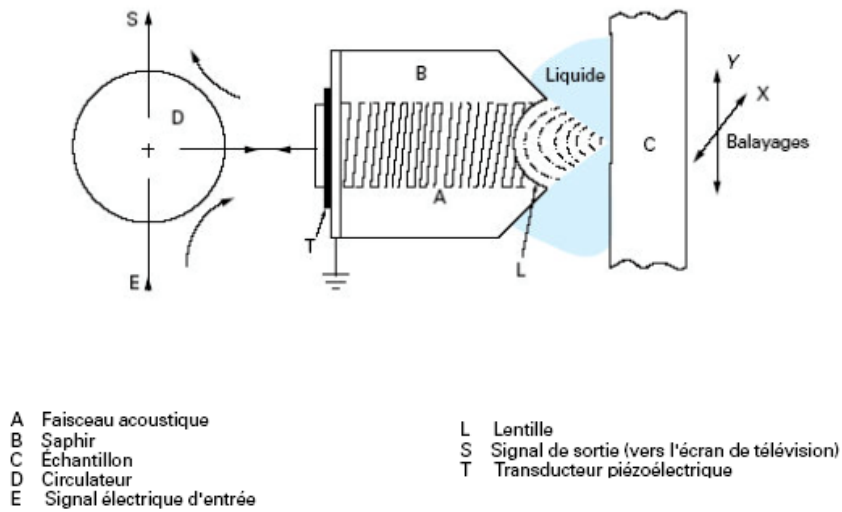


Figure C.13 : Schéma de la tête de mesure d'un microscope acoustique, extrait de (Attal 1998)

Les spécifications techniques fournies ici correspondent à l'appareil Evolution commercialisé par la société Sam Tec.

C.4.1.3 Avantages et inconvénients

Avantages :

- Fournit une information en 3D (plus riche en vue de l'interprétation)
- Bonne résolution en X, Y, Z
- La vision acoustique de la profondeur peut se faire sur des matériaux optiquement opaques
- Complémentarité avec la microscopie optique

Inconvénients :

- Coût élevé
- La pièce à étudier doit être plongée dans l'eau
- Lenteur de la mesure (typiquement pour couvrir une étendue de 1m^2 avec un maillage de $1\mu\text{m}$: temps de mesure de l'ordre de la semaine!)

C.4.1.4 Conclusion

Le principal avantage de cette méthode est de pouvoir faire des relevés 3D précis dans des matériaux optiquement opaques ; dans le cas qui nous intéresse, on ne peut tirer parti de cet avantage. D'autre part ses inconvénients sont très handicapants (vitesse et étude de la pièce dans l'eau), ce qui explique que nous ayons écarté cette solution.

C.4.2 La mesure par réflexion

C.4.2.1 Principe de la mesure

Une source lumineuse émet un cône de lumière, la lumière réfléchiée et diffusée par la pièce est mesurée par un détecteur (Charron 2004a). L'intensité lumineuse mesurée par le détecteur est fonction de la distance qui le sépare de la pièce et des caractéristiques optiques de la surface de celle-ci : lorsque la distance est très faible, le récepteur ne reçoit pas de lumière, lorsqu'elle augmente, l'intensité collectée passe par un maximum et décroît à grande distance. On peut donc calculer une altitude z correspondant à une petite surface éclairée.

C.4.2.2 Réalisation technologique

Pour compenser l'influence de la surface mesurée sur la réponse d'un capteur de mesure de distance par réflexion, on utilise deux capteurs en parallèle avec une disposition différente du deuxième récepteur (Figure C.14).

Le rapport des sorties des deux ensembles est calculé pour créer une sortie fonction de la distance mais relativement insensible aux variations de réflectance de la pièce mesurée (Figure C.15). Dans la plupart des réalisations industrielles, des fibres optiques sont utilisées pour « déporter » la source de lumière et les détecteurs au niveau du point de mesure : cela permet plus de souplesse et la miniaturisation de la zone mesurante.

On associe à ce capteur une table motorisée en X et Y qui donne un maillage selon ces axes. La reconstruction tridimensionnelle finale de la surface est effectuée par un ordinateur.

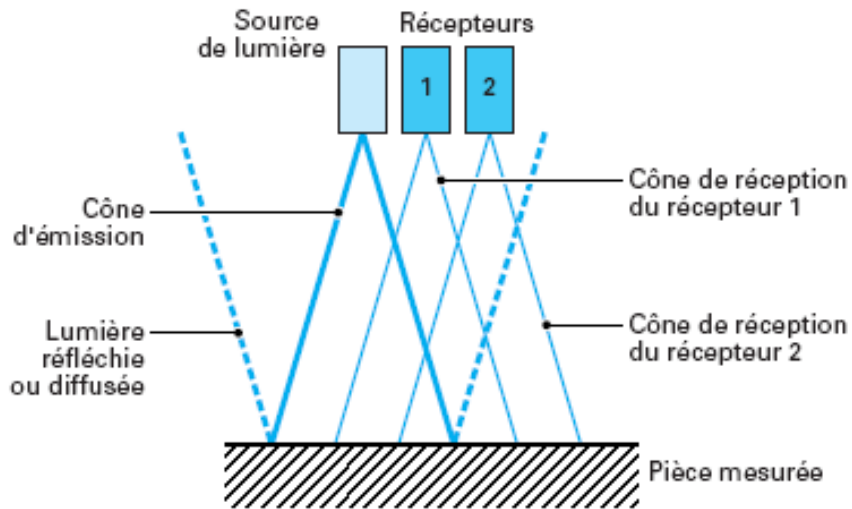


Figure C.14 : Schéma de principe de la mesure par réflexion, extrait de (Charron 2004a)

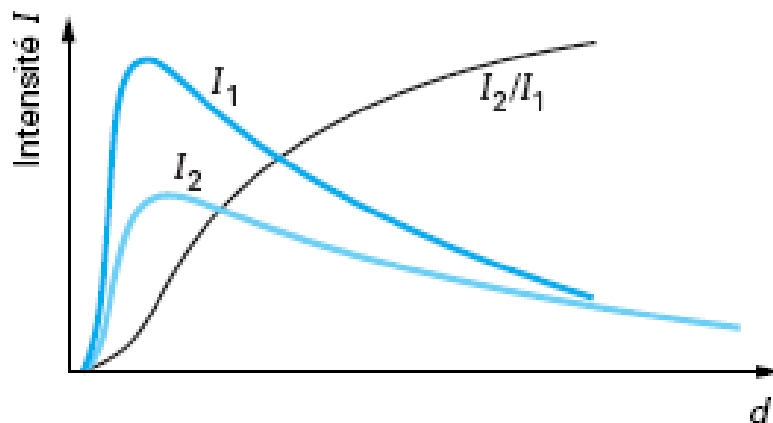


Figure C.15 : Courbes des intensités lumineuses recueillies par les capteurs n°1 et n°2 en fonction de la distance à la pièce mesurée, extrait de (Charron 2004a)

Spécifications techniques :

- Résolution en Z : de l'ordre de quelques nm
- Résolution en X et Y : dépend de la grandeur de la surface mesurante (on peut descendre en dessous de 50 μm)
- Vitesse d'acquisition : typiquement 20 kHz ; on peut monter à 1MHz au détriment de la précision de mesure

C.4.2.3 Avantages et inconvénientsAvantages :

- Excellente résolution en Z
- Mesure très rapide
- Coût relativement peut élevé

Inconvénients :

- Résolution en X et Y insuffisante
- Etalonnage sur pièce nécessaire
- La surface de la pièce doit avoir une réflectivité relativement uniforme

C.4.2.4 Conclusion

La faible résolution en X et Y nous interdit l'utilisation de ce genre de capteur pour l'application qui nous intéresse. Ce capteur reste cependant intéressant pour l'acquisition d'une distance ou d'un déplacement : il est simple, robuste et donne rapidement des mesures très précises en Z.

C.4.3 La technique IIM (Imaging Interferometric Microscopy)**C.4.3.1 Principe de la mesure**

Une partie de l'information sur un objet est contenue dans le signal diffracté dans des directions qui ne sont pas prises en compte par les moyens d'éclairage conventionnels. La méthode IIM a pour but de recouvrer ce signal grâce à un éclairage « extrêmement hors d'axe ». On a ainsi accès à un signal haute fréquence qui permet d'augmenter la résolution

du microscope. On jouit alors des avantages d'un objectif à grande ouverture numérique (bonne résolution), sans ses inconvénients (grande profondeur de champ, grand champ et grande distance de travail).

C.4.3.2 Réalisation technologique

On éclaire la pièce à étudier selon une direction extrêmement inclinée, de façon à recueillir les ordres de diffraction supérieurs à travers un dispositif de microscopie classique (Schwarz et al. 2003). L'ordre 0 est quant à lui reconstruit en le faisant interférer avec une onde de référence (Figure C.16). Le schéma du montage optique correspondant est donné dans la Figure C.17.

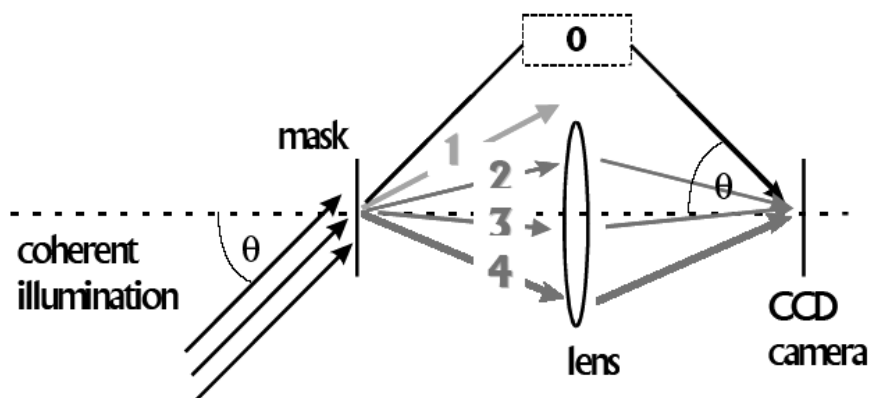


Figure C.16 : Principe de l'imagerie IIM, extrait de (Schwarz et al. 2003)

Spécifications techniques : Cette méthode étant présentée dans un article datant de 2003, seul un prototype visant à prouver sa faisabilité a été proposé ; les spécifications techniques de l'appareil ne sont donc pas connues. On sait cependant que l'on obtient, grâce à cette méthode, une résolution pour un objectif 20 X ($NA = 0.4$), comparable à celle d'un objectif « classique » 100x ($NA = 0.9$) ; les images obtenues sont également plus contrastées.

C.4.3.3 Avantages et inconvénients

Avantages :

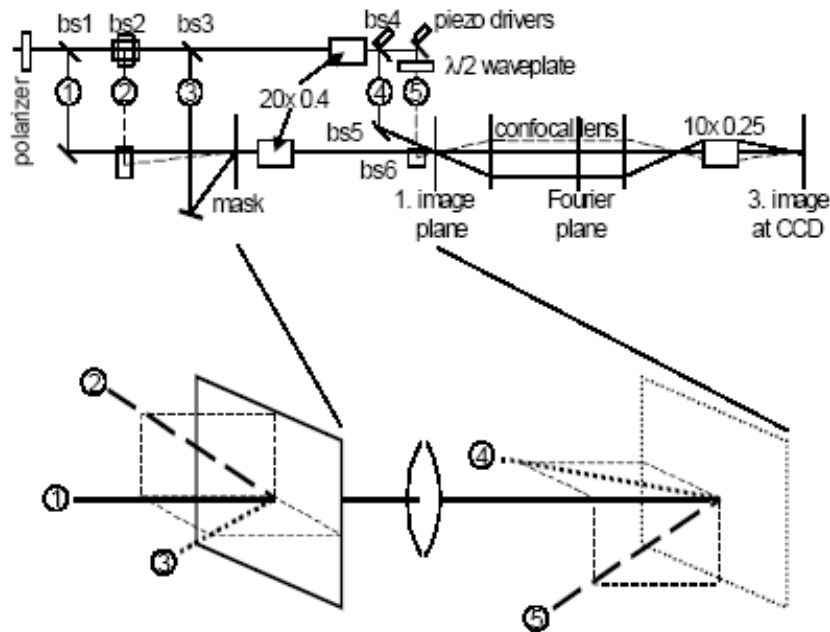


Figure C.17 : Montage optique associé à la technique d'imagerie IIM, extrait (Schwarz et al. 2003)

- Relative rapidité de la mesure par rapport à sa résolution (champs 25 fois plus grand pour une même résolution, par rapport à la microscopie classique)

Inconvénients :

- Méthode expérimentale donc coûteuse et pas forcément adaptée
- Recombinaison et traitement des signaux nécessitant de grosses ressources en calcul

C.4.3.4 Conclusion

Le principal avantage de cette technique est de régler certains problèmes inhérents à l'utilisation de techniques microscopiques (dus principalement au fort NA pour les grossissements classiquement utilisés). Malheureusement cette technique reste expérimental, met en jeu un montage assez lourd et complexe, et nécessite une grande puissance de calcul, ce qui explique que nous ayons écarté cette solution.

C.5 Bibliographie

Alexander, M. and Hallett, F. R. 1-6-1999. Small-angle light scattering : instrumental design and application to particle sizing. *Applied Optics*, 38 (19),p.4158-4163.

Arnison, M. R., Larkin, K. G., Sheppard, C. J. R., Smith, N. I., and Cogswell, C. J. 2004. Linear phase imaging using differential interference contrast microscopy. *Journal of Microscopy*, 214 (1),p.7-12.

Attal, J. 1998. Microscopie Acoustique. *Techniques de l'Ingénieur*, R1402.

Bakolias, C. and Forrest, A. K. 1997. Dark Field, Scheimpflug imaging for surface inspection. *Proceedings of SPIE*, 3029,p.57-68.

Bechtler, L. and Velidandla, V. 2003. Optical Surface Analysis : a new technique for the inspection and metrology of optoelectronic films and wafers. *Proceedings of SPIE*, 4944 (Integrated Optical Devices : Fabrication and Testing),p.109-116.

Benatmane, A. 2002. Développement de la Microscopie Interférométrique pour une Meilleure Analyse Morphologique des Couches Minces et Epaissees des Matériaux Semi-conducteurs et Optiques. PhD. Université Louis Pasteur - Strasbourg I - Discipline : Sciences pour l'Ingénieur (spécialité : Instrumentation et Microélectronique).

Bertin, T. 2005. MICO : Moyen d'Inspection des Composants Optiques. *Bertin Technologies*.

Bouchareine, P. 1999. Métrologie des Surfaces. *Techniques de l'Ingénieur*, R1390.

Cha, S., Lin, P. C., Zhu, L., Sun, P.-C., and Fainman, Y. 2000. Nontranslational Three-Dimensional Profilometry by Chromatic Confocal Microscopy with Dynamically Configurable Micromirror Scanning. *Applied Optics*, 39 (16).

Charron, J.-L. 2004a. Mesures sans contact - Méthodes Optiques (partie 1). *Techniques de l'Ingénieur*, R1332.

Charron, J.-L. 2004b. Mesures sans Contact - Méthodes Optiques (partie 2). *Techniques de l'Ingénieur*, R1333.

Charron, J.-L. 2004c. Mesures sans Contact - Comparatif. *Techniques de l'Ingénieur*, R1335.

Chatterjee, S. 2003. Design Considerations and Fabrication Techniques of Nomarski Reflection Microscope. *Optical Engineering*, 42 (8),p.2202-2212.

Chen, F., Brown, G. M., and Song, M. 2000. Overview of three-dimensional shape

measurement using optical methods. *Optical Engineering*, 39 (1),p.10-22.

Chen, W., Zeng, A., and Akbulut, M. 2004. System and methods for classifying anomalies of sample surfaces. Patent n ° US2004169852. 2-9-2004.

Flauger, T. 1996. Verfahren und Einrichtung zum Erkennen eines Defekts an einem Körper. Patent n ° DE19510535. 26-9-1996.

Flewitt, P. E. J. and Wild, R. K. 1994. Light Microscopy, in *Physical Methods for materials characterisation*, 1994.

Gheen, A., Wang, Y., and Wang, Z. 1997. New Optical Imaging Method for Lithography and High-Resolution Inspections. *Proceedings of SPIE*, 3051 (Optical Microlithography X),p.94-105.

Gomez, S., Hale, K., Burrows, J., and Griffiths, B. 1998. Measurements of surface defects on optical components. *Measurement Science and Technology*, 9,p.607-616.

Henley, F. J. and Bryan, M. A. 1998. Technique for determining defect positions in three dimensions in a transparent structure. Patent n ° US5790247. 4-8-1998.

Henzler, M., Kumpe, R., Frischat, H., and Kopp, F.-O. 2000. Device and method for examining the smoothness of a sample surface. Patent n ° US6128073. 3-10-2000.

Holmes, T. J. and Levy, W. J. 15-9-1987. Signal-processing Characteristics of Differential-Interference-Contrast Microscopy. *Applied Optics*, 26 (18),p.3929-3939.

Honlet, M. 1997. Nondestructive Testing of Complex Composite Materials and Structures Using Optical Techniques. *COFREND Congress on Nondestructive Testing* .

Hue, J., Muffato, V., Pellé, C., Quesnel, E., Garrec, P., and Baume, F. 2001. Automated Set-up for Extreme Ultraviolet Mask Lithography : The First Step to Count and Clean in One. *Proceedings of SPIE*, 4343 (Emerging Lithographic Technologies V),p.627-638.

Hue, J., Quesnel, E., Pellé, C., Muffato, V., Carini, G., Favier, S., and Besson, P. 2003. EUV Mask Blank : Defect Detection at 100 nm. *Proceedings of SPIE*, 5037 (Illumination optimization via multifunction approximation),p.860-871.

Imaino, W. A. and Latta, M. R. 1999. Surface inspection tool using reflected and scattered light. Patent n ° US5898492. 27-4-1999.

Jcong, S., Lai, C., Rekawa, S., Walton, C. C., and Bokor, J. 2000. Actinic Defect Counting Statistics over 1 cm² Area of EUVL Mask Blank. *Proceedings of SPIE*, 3997 (Direct creation of microstructures in fluorine-containing polymers),p.431-440.

Kato, K., Sakamoto, M., Yashikawa, S., Yamamiya, K., Kodama, H., and Ohsima, K. 1984. Apparatus for detecting optically defects. Patent n ° US4464050. 7-8-1984.

Kayser, D., Bothe, T., and Osten, W. 2004. Scaled topometry in a multisensor approach. *Optical Engineering*, 43 (10),p.2469-2477.

Klooster, A., Marks, J., Hanson, K., and Sawatari, T. 2004. Scanning Holographic Scatterometer for Wafer Surface Inspection. *Proceedings of SPIE*, 5375 (Metrology, Inspection, and Process Control for Microlithography XVIII),p.1314-1324.

Lalezari, R. and Knollenberg, R. G. 1989. Surface Analysis for the Characterization of Defects In Thin Film Processes. *Proceedings of SPIE*, 1164 (Surface Characterization and Testing II),p.222-235.

Lessor, D. L., Hartman, J. S., and Gordon, R. L. 1979. Quantitative surface topography determination by Nomarski reflection microscopy. I. Theory. *Journal of Optical Society of America*, 69 (2),p.357-366.

Li, Q., Gao, H., Xue, S., and Li, Y. 2001. Optical profilometer based on the principle of differential interference. *Optical Engineering*, 40 (5),p.833-836.

Lorincik, J., Marton, D., King, R. L., and Fine, J. 1996. Scanning Scattering Microscope for Surface Microtopography and Defect Imaging. *Journal of Vacuum Science & Technology B*, 14 (4),p.2417-2423.

Mäge, I., Pintér, B., Tuckermann, M., and Donzella, O. 2004. Defect Learning with 193-nm Resists. *Proceedings of SPIE*, 5375.

Montgomery, P. C. and Fillard, J. P. 1992. Peak Fringe Scanning Microscopy (PFSM) : Sub-Micron 3D Measurement of Semiconductor Components. *Proceedings of SPIE*, 1755 (Interferometry : Techniques and Analysis),p.12-23.

Okada, S., Sumimoto, T., Imade, M., and Miyaucha, H. 1999. Method and apparatus for inspecting surface conditions. Patent n ° US5125741. 19-5-1999.

Philbert, M. and Gryson, M. 1962. Réalisation et Contrôle par Strioscopie Interférentielle de Miroirs Plans Sphériques et Paraboliques. *Optical Instruments and Techniques*, Proceedings of the conference held in London, 1961,p.352.

Pladellorens, J. M., Caum, J., Tapias, M., Cadevall, C., Anto, J., and Fernandez, X. 2002. Analysis and characterization of surface defects in ophtalmic lenses. *Proceedings of SPIE*, 4780 (Surface Scattering and Diffraction for Advanced Metrology II),p.99-106.

Plaut-Aubry, R. 2000. Mesures de Contraintes Résiduelles par Interférométrie de Speckle. Mémoire présenté en vue d'obtenir le Diplôme d'Ingénieur C.N.A.M. en Instrumentation/Mesure - Option Contrôle industriel.

Puech, P. 2002. Cours d'Optique Géométrique et Ondulatoire. *Laboratoire de Physique des Solides de Toulouse*.

Quan, C., Wang, S. H., Tay, C. J., Shang, H. M., and Chan, K. C. 1-9-2000. Inspection of micro-cracks on solderball surface using a laser scattering method. *Optics Communication*, 183,p.19-27.

Rainer, F. 1997. Mapping and Inspection of Damage and Artifacts in Large-Scale Optics. *XXIX Annual Boulder Damage Symposium* .

Rangarajan, B., Singh, B., Subramaniam, R., and Templeton, M. A. 2004. Scatterometry based measurement of a moving substrate. Patent n ° US6704101. 9-3-2004.

Rioufreyt, P. 1990. Développement d'un Système de Métrologie Optique Tridimensionnelle sans Contact. Thèse de Doctorat en Physique ; Université Paris XI Orsay.

Roblin, G. 1999b. Microscopies Optiques à Balayage. *Techniques de l'Ingénieur*, R6714.

Roblin, G. 1999a. Microscopie Optique. *Techniques de l'Ingénieur*, R6712.

Sankaran, V. and Stover, J. C. 1999. Advanced Particle Sizing Technique for Development of High-Accuracy Scanner Calibration Standards. *SEMATECH*.

Schwarz, C. J., Kuznetsova, Y., and Brueck, S. R. J. 2003. Demonstration of Imaging Interferometric Microscopy (IIM). *Proceedings of SPIE*, 5038 (Metrology, Inspection, and Process Control for Microlithography XVII),p.1-11.

Sekine, A., Iwa, Y., Soma, H., Miki, N., Hisozaki, H., and Yoshino, H. 2000. Surface inspection apparatus. Patent n ° US6104481. 15-8-2000.

Serry, F. M., Stout, T. A., Zecchino, M. J., Ragan, C., and Browne, P. A. 2004. 3D MEMS Metrology with Optical Profilers. *Veeco Instruments, Inc.*

Smigielski, P. 2001a. Holographie Optique - Interférométrie Holographique. *Techniques de l'Ingénieur*, R6330.

Smigielski, P. 2001b. Interférométrie de Speckle. *Techniques de l'Ingénieur*, R6331.

Spoto, G., Branciforte, M., Doddo, F., and Occhipinti, L. 2002. System and process for analysing surface defects. Patent n ° US2002191831. 19-12-2002.

STIL 2005. Principe du codage chromatique. www.stilsa.com/FR/prin.htm

- Szeliski, R. and Zabih, R. 1999. An experimental comparison of stereo algorithms. *International Workshop on Vision Algorithms*, p.1-19.
- Taillade, F. 2003. La Shearographie, une Méthode Optique Plein Champ, pour la Visualisation des Ondes Mécaniques. *Action Spécifique "Contrôle Non Destructif - Intégration Multi-Capteur" ESPCI*.
- Tezuka, Y., Ito, M., Terasawa, T., and Tomie, T. 2003. Design and Development of a Novel Actinic Inspection Tool for EUV Multilayer-Coated Mask Blanks. *Proceedings of SPIE*, 5038 (Metrology, Inspection, and Process Control for Microlithography XVII), p.866-878.
- Tsuji, T., Miyazaki, K., Takeuchi, S., Yoshii, M., and Nose, N. 1999. Optical inspection method and apparatus including intensity modulation of a light beam and detection of light scattered at an inspection position. Patent n° US5861952. 19-1-1999.
- Weiss, A. and Saranli, A. 2004. Method and apparatus for flat and patterned media inspection. Patent n° US2004109598. 10-6-2004.
- Welch, K. 1996. Utilization of near specular energy in the detection of surface defects. *Proceedings of SPIE*, 2862, p.119-128.
- Whitehouse, D. J. 1985. Instrumentation for measuring finish, defects and gloss. *Proceedings of SPIE*, 525 (Measurement and Effects of Surface Defects and Quality of Polish), p.106-123.
- Whitehouse, D. J. 1997. Surface Metrology. *Measurement Science and Technology*, 8, p.955-972.
- Zanier, S. 2002. Optique de Fourier - Détramage et Strioscopie. *4ème année INSA Toulouse*.

Bibliographie

(Adam 2001) Adam, S. et al., "Utilisation de la transformée de Fourier-Mellin pour la reconnaissance de formes multi-orientées et multi-échelles : application à l'analyse automatique de documents techniques," *Traitement du Signal* 18 (1) : 17-33 (2001)

(Arnison 2000) Arnison, M. R. et al., "Using the Hilbert transform for 3D visualization of differential interference contrast microscope images," *Journal of Microscopy* 199 (1) : 79-84 (2000)

(Arnison 2004) Arnison, M. R. et al., "Linear phase imaging using differential interference contrast microscopy," *Journal of Microscopy* 214 (1) : 7-12 (2004)

(Avouris and Gasser 1992) Avouris, N and L. Gasser. *Distributed Artificial Intelligence : Theory and Practics*. Kluwer Academic Press, Dordrecht (1992)

(Baker 2004) Baker, L. R. 2004. *Metrics for High-Quality Specular Surfaces*. vol. TT65, ed. Lavoisier (2004)

(Barbarà and Chen 2007) Barbarà, D. and Chen, P. "Using the Fractal Dimension to Cluster Datasets", Proceedings of the sixth ACM SIGKDD International Conference on Knowledge discovery and data mining, p.260-264, August 20-23, 2007, Boston, Massachusetts, United States (2007)

(Bartkowiak and Gorski 2006) Bartkowiak, A. and Gorski, K. "Personal Authentication from Palm Scans : An Introduction," 13th International Multi-Conference on Advanced Computer Systems - ACS 2006. (2006)

(Belkin and Niyogi 2002) Belkin, M. and Niyogi, P. "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," Technical Report TR 2002-01, Univ.

Chicago, Dept. Comp. Sci. and Statistics, January 2002 (2002).

(Bouchareine 1999) Bouchareine, P., "Métrologie des Surfaces," *Techniques de l'Ingénieur* R1390 (1999)

(Bouveyron 2003) Bouveyron, C., "Dépliage du ruban cortical à partir d'images obtenues en IRMf", Mémoire de DEA de Mathématiques Appliquées, Université J. Fourier, Grenoble 1 (2003)

(Bouyoucef 2007) Bouyoucef, E. K., "Contribution à l'étude et la mise en oeuvre d'indicateurs quantitatifs et qualitatifs d'estimation de la complexité pour la régulation du processus d'auto-organisation d'une structure neuronale modulaire de traitement de l'information.", Thèse de Doctorat - Université Paris XII-Val-de-Marne (2007)

(Bruske and Sommer 1997) Bruske, J. and Sommer, G. "Intrinsic Dimensionality Estimation with Optimally Topology Preserving Maps," Technical Report TR 9703. 1997. Computer Science Institute, Christian-Albrechts University Kiel. (1997)

(Buchala et al. 2005) Buchala, S. et al., "Analysis of Linear and Nonlinear Dimensionality Reduction Methods for Gender Classification of Face Images," *International Journal of Systems Science* 14 (36) : 931-942 (2005)

(Camastra and Vinciarelli 2001) Camastra, F. and A. Vinciarelli, "Intrinsic Dimension Estimation of Data : An Approach Based on Grassberger-Procaccia's Algorithm," *Neural Processing Letters* 14 (1) : 27-34 (2001)

(Camastra and Vinciarelli 2002) Camastra, F. and A. Vinciarelli, "Estimating the Intrinsic Dimension with a Fractal-Based Method," *IEEE Trans.on Pattern Analysis and Machine Intelligence* (2002)

(Charron 2004) Charron, J.-L., "Mesures sans Contact - Comparatif," *Techniques de l'Ingénieur* R1335 (2004)

(Chatterjee 2003) Chatterjee, S., "Design Considerations and Fabrication Techniques of Nomarski Reflection Microscope," *Optical Engineering* 42 (8) : 2202-2212 (2003)

(Chen, Brown, and Song 2000) Chen, F., G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," *Optical Engineering* 39 (1) : 10-22 (2000)

(Chen, Zeng, and Akbulut 2004) Chen, W., Zeng, A., and Akbulut, M., "System and methods for classifying anomalies of sample surfaces," Patent [US2004169852]. 2-9-2004.(2004)

(Chettaoui et al. 2005) Chettaoui, C., Djemal, K., Djouak, A., and Maaref, H., "Etude de formes des globules drépanocytaires par traitement numérique des images, " 3rd International Conference : Sciences of Electronic, Technologies of Information and Telecommunications (SETIT 2005).(2005)

(Choksuriwong, Laurent, and Emile 2005) Choksuriwong, A., Laurent, H., and Emile, B., "Comparison of invariant descriptors for object recognition," IEEE International Conference on Image Processing (ICIP). 1, 377-380. (2005)

(Chong, Raveendran, and Mukundan 2003) Chong, C.-W., P. Raveendran, and R. Mukundan, "A Comparative Analysis of Algorithms for Fast Computation of Zernike Moments.," *Pattern Recognition* 36 : 731-742 (2003)

(Cohen, Raz, and Malah 1995a) Cohen, I., S. Raz, and D. Malah, "Orthogonal Shift-Invariant Wavelet Packet : Décomposition and Representation," *Signal Processing* 57 (3) : 251-270 (1995)

(Cohen, Raz, and Malah 1995b) Cohen, I., Raz, S., and Malah, D. "Shift-Invariant Adaptive Local Trigonometric Decomposition," Proc.of the 4th European Conference on Speech, Communication and Technology , 247-250. (1995)

(Comon, Voz, and Verleysen 1994) Comon, P., Voz, J. L., and Verleysen, M. "Estimation of performance bounds in supervised classification," *European Symposium on Artificial Neural Networks 1994 - ESANN'94*. 37-42. (1994)

(Costa, Girotra, and Hero 2005) Costa, J. A., Girotra, A., and Hero, A. O. "Estimating Local Intrinsic Dimension with k-Nearest Neighbor Graphs," 13th IEEE Workshop on Statistical Signal Processing (SSP). 417-422. (2005)

(Costa and Hero 2004) Costa, J. A. and A. O. Hero, "Geodesic Entropic Graphs for Dimension and Entropy Estimation in Manifold Learning," *IEEE Trans.on Signal Processing* 52 (8) : 2210-2221 (2004)

(Cselényi 2005) Cselényi, Z., "Mapping the dimensionality, density and topology of

data : The growing adaptive neural gas," *Computer Methods and Programs in Biomedicine* 78 : 141-146 (2005)

(de Backer 2002) de Backer, S., "Unsupervised Pattern Recognition - Dimensionality Reduction and Classification", PhD, University of Antwerp. (2002)

(Deguillemont, Lecoecuche, and Dubus 1998) Deguillemont, D., Lecoecuche, S. and J.-P. Dubus, "Pattern recognition of industrial defects by multiresolution analysis with wavelet decomposition," *Proceedings of SPIE* 3458 (Part of the SPIE Conference on Wavelet Applications in Signal and Image Processing VI) : 210-220 (1998)

(Demartines 1994) Demartines, P., "Analyse de Données par Réseaux de Neurones Auto-Organisés", Thèse de Doctorat, Institut National Polytechnique de Grenoble (1994)

(Demartines 1992) Demartines, P., "Organization Measures and Representations of Kohonen Maps," First IFIP Working Group Workshop. (1992)

(Demartines and Blayo 1992) Demartines, P. and F. Blayo, "Kohonen Self-Organizing Maps :Is the Normalization Necessary?," *Complex Systems* 6 (2) : 105-123 (1992)

(Demartines and Héroult 1993) Demartines, P. and J. Héroult, "Vector Quantization and Projection Neural Network," *Lecture Notes in Computer Science* 686 (International Workshop on Artificial Neural Networks) : 328-333 (1993)

(Demartines and Héroult 1995) Demartines, P. and Héroult, J. CCA : "Curvilinear Component Analysis". Proceedings of 15th workshop GRETSI . 15-9-1995. (1995)

(Dennis and Schnabel 1983) Dennis, J. E. and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ : Prentice Hall.(1983)

(Derrode 1999) Derrode, S., "Représentation de Formes Planes à Niveaux de Gris par Différentes Approximations de Fourier-Mellin Analytique en vue d'Indexation de Bases d'Images.", Thèse de Doctorat - Université de Rennes I (1999)

(Derrode and Ghorbel 2001) Derrode, S. and F. Ghorbel, "Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description," *Computer Vision and Image Understanding* 83 (1) : 57-78 (2001)

(Derrode, Mezhoud, and Ghorbel 2000) Derrode, S., R. Mezhoud, and F. Ghorbel, "Comparaison de deux familles complètes de descripteurs de formes pour l'indexation de bases d'objets 2D à niveaux de gris," *Annales de Télécommunications* 57 (5-6) : 77-86 (2000).

(Duchene 2003) Duchene, C., "Traitement de données multidimensionnelles par Analyse en Composantes Curvilignes", Mémoire de D.E.A. Traitement des Images et du Signal de Cergy Pontoise (2003)

(Ducottet 1994) Ducottet, C., "Etude de l'Application des Transformées en Ondelettes au Traitement d'Images d'Écoulements Obtenues par Tomographie ou Microholographie", Thèse de Doctorat, Université Jean Monnet, Saint-Etienne. (1994)

(Egmont-Petersen, de Ridder, and Handels 2002) Egmont-Petersen, M., D. de Ridder, and H. Handels, "Image Processing with Neural Networks - A Review," *Pattern Recognition* 35 : 2279-2301 (2002)

(Ferber 1998) Ferber, J. 1998. *Multi-Agent Systems ; Towards a Collective Intelligence*. Reading, MA : Addison-Wesley. (1998)

(Flewitt and Wild 1994) Flewitt, P. E. J. and R. K. Wild. Light Microscopy. In *Physical Methods for materials characterisation*. (1994)

(Foresee and Hagen 1997) Foresee, F. D. and Hagen, M. T. "Gauss-Newton approximation to Bayesian regularization," Proceedings of the 1997 International Joint Conference on Neural Networks. (1997)

(Fritzke 1995) Fritzke, B., "A Growing Neural Gas Network Learns Topologies," *Advances in Neural Information Processing Systems* 7 (1995)

(Fukunaga 1982) Fukunaga, K. *Intrinsic Dimensionality Extraction*. Vol. Classification, Pattern Recognition and Dimensionality Reduction. Edited by L. N. Kanal and P. R. Krishnaiah. Amsterdam. (1982)

(Fukunaga and Olsen 1971) Fukunaga, K. and D. R. Olsen, "An algorithm for finding intrinsic dimensionality of data," *IEEE Trans .on Computers* 20 (C) : 176-183 (1971)

(Gallant, Braun, and Essen 1993) Gallant, J.-L., J. Braun, and D. C. V. Essen, "Selectivity of Polar, Hyperbolic, and Cartesian Gratings in Macaque Visual Cortex," *Science* 259 : 100-103 (1993)

(Garris, Wilkinson and Wilson 1991) Garris, M. D., Wilkinson, R. A., and Wilson, C. L. "Analysis of a Biologically Motivated Neural Network for Character Recognition," *Proceedings : Analysis of Neural Network Applications*. George Mason University, ACM Press. (1991)

(Gebbinck et al. 1993) Gebbinck, M. S. K. et al., "Application of Neural Networks for the Classification of Diffuse Liver Disease by Quantitative Echography," *Ultrason Imaging* 15 (3) : 205-217 (1993)

(Ghorbel 1994) Ghorbel, F., "A Complete Invariant Description for Gray Level Images by the Harmonic Analysis Approach," *Pattern Recognition* 15 : 1043-1051 (1994)

(Ghorbel 1998) Ghorbel, F., "Toward an Unitary Formulation for Invariant Pattern Recognition : Application to Image Coding," *Annales de Télécommunications* 53 (5) : 242-260 (1998)

(Gleason et al 1998) Gleason, S. S. et al., "Rapid Yield Learning through Optical Defect and Electrical Test Analysis," *Proceedings of SPIE* 3332 : 232-242 (1998)

(Gomez 1998) Gomez, S. et al., "Measurements of surface defects on optical components," *Measurement Science and Technology* 9 : 607-616 (1998)

(Gonzalez and Woods 2002) Gonzalez, R. C. and R. E. Woods. *Digital Image Processing*. 2nd edition. Addison-Wesley (2002)

(Grace and Spann 1991) Grace, A. E. and M. Spann, "A Comparison between Fourier-Mellin Descriptors and Moment Based Features for Invariant Object Recognition Using Neural Networks," *Pattern Recognition Letters* 12 : 635-643 (1991)

(Grassberger and Procaccia 1983) Grassberger, P. and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D* 9 : 189-208 (1983)

(Grother 1992) Grother, P. J., "Karhunen Loève Feature Extraction for Neural Handwritten Character Recognition," *Proceedings of SPIE* 1709 (Applications of Artificial Neural Networks III) : 155-166 (1992)

(Harb 2003) Harb, H., "Classification du Signal sonore en Vue d'une Indexation par le Contenu des Documents Multimédias", Thèse de Doctorat, Ecole Centrale de Lyon. (2003)

(Haro, Randall, and Sapiro 2006) Haro, G., Randall, G., and Sapiro, G., "Stratification Learning : Detecting Mixed Density and Dimensionality in High Dimensional Point Clouds," *Advances in Neural Information Processing Systems - NIPS 2006. Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems* , 553-560. (2006)

(Hartman, Gordon, and Lessor 1980) Hartman, J. S., R. L. Gordon, and D. L. Lessor, "Quantitative surface topography determination by Nomarski reflection microscopy. II : Microscope modification, calibration, and planar sample experiments," *Applied Optics* 19 (17) : 2998-3009 (1980)

(Heikkonen and Lampinen 1999) Heikkonen, J. and Lampinen, J., "Building Industrial Applications with Neural Networks," *Proc. European Symposium on Intelligent Techniques, ESIT'99. 15-6-1999.* (1999)

(Henzler et al. 2000) Henzler, M., Kumpe, R., Frischat, H., and Kopp, F.-O. , "Device and method for examining the smoothness of a sample surface," Patent [US6128073]. 3-10-2000. (2000)

(Hing 2004) Hing, S. J., Lee, S. J., Park, D.-C., and May, G. S. , "Modular Neural Networks for Semiconductor Manufacturing Process Modeling and Characterization : Application to Plasma Processing," *Artificial Neural Networks in Engineering ANNIE 2004* . (2004)

(Holmes and Levy 1987) Holmes, T. J. and W. J. Levy, "Signal-processing Characteristics of Differential-Interference-Contrast Microscopy," *Applied Optics* 26 (18) : 3929-3939 (1987)

(Hong et al.2004) Hong, S. J et al., "Automated Fault Detection and Classification of Etch Systems Using Modular Neural Networks," *Proceedings of SPIE 5378 (Data Analysis and Modeling for Process Control)* : 134-141 (2004)

(Hu 1962) Hu, M. K., "Visual Pattern Recognition by Moment Invariant," *IRE Transactions on Information Theory* 1 : 179-187 (1962)

(Imaino et al. 1999) Imaino, W. A., Juliana, A., Latta, M. R., Lee, C. H., Leung, W. C., and Rosen, H. J., "Surface inspection tool," Patent [US5867261]. 2-2-1999. (1999)

(Imaino and Latta 1999) Imaino, W. A. and Latta, M. R., "Surface inspection tool using reflected and scattered light," Patent [US5898492]. 27-4-1999. (1999)

(Jabr 1985) Jabr, S. N., "Surface-roughness measurement by digital processing of Nomarski phase-contrast images," *Optics Letters* 10 (11) : 526-528 (1985)

(Kalman and Kwasny 1998) Kalman, B. L. and Kwasny, S. C., "Diagnostic Screening of Digital Mammograms : Using Wavelets and Neural Networks to Extract Structure," Technical Report - Washington University.(1998)

(Kayser, Bothe and Osten 2004) Kayser, D., T. Bothe, and W. Osten, "Scaled topometry in a multisensor approach," *Optical Engineering* 43 (10) : 2469-2477 (2004)

(Kégl 2002) Kégl, B. "Intrinsic Dimension Estimation Using Packing Numbers," In *Advances in Neural Information Processing Systems*, Volume 15, pages 681–688, Vancouver, Canada, Dec. 2002. The MIT Press.(2002)

(Kieckhäfer 1996) Kieckhäfer, J. , "Device for inspecting transparent objects and/or objects provided on one side with an opaque coating for material defects," Patent [WO9618882]. 20-6-1996.(1996)

(Kihira 1997) Kihira, H., "Method of and apparatus for detecting defect of transparent sheet as sheet glass," Patent [US5691811]. 25-11-1997.(1997)

(Kittler 1986) Kittler, J. Feature Selection and Extraction. In *Handbook of Pattern Recognition and Image Processing*.Academic Press. (1986)

(Knollenberg 1990) Knollenberg, R. G. , "Surface analysis system and method," Patent [US4893932]. 16-1-1990. (1990)

(Knollenberg, Hoxie, and Clinton 1996) Knollenberg, R. G., Hoxie, B. C., and Clinton, E. U. , "Surface defect inspection system and method," Patent [US5493123]. 20-2-1996. (1996)

(Kohonen 2001) Kohonen, T. *Self Organizing Maps*. Vol.30 of Springer Series in Information Sciences. 3rd edition. Berlin : Springer. (2001)

(Kohonen et al. 1996) Kohonen, T. et al., "Engineering Applications of the Self-Organizing Maps," *Proceedings of the IEEE* 84 (10) : 1358-1384 (1996)

-
- (Kutil and Uhl 2001) Kutil, R. and A. Uhl, "Parallel Adaptive Wavelet Analysis," *Future Generation Computer Systems* 18 (1) : 97-106 (2001)
- (Lampinen 1992) Lampinen, J., "On Clustering Properties of Hierarchical Self-Organizing Maps," *Journal of Mathematical Imaging and Vision* 2 : 261-272 (1992)
- (Lampinen and Oja 1995) Lampinen, J. and Oja, E., "Distortion Tolerant Pattern Recognition Based on Self-Organizing Feature Extraction," *IEEE Trans. on Neural Networks*. 6[3], 539-547. (1995)
- (Lampinen, Smolander, and Korhonen 1995) Lampinen, J., Smolander, S., and Korhonen, M. , "Wood Surface Inspection System Based on Generic Visual Features," Proceedings of the Industrial Conference "Technical Diagnosis & Nondestructive Testing" in the International Conference on Artificial Neural Networks, ICANN'95 . (1995)
- (Lee et al.2000) Lee, J. A., Lendasse, A., Donckers, N., and Verleysen, M., "A Robust Nonlinear Projection Method," European Symposium on Artificial Neural Networks - ESANN'2000. (2000)
- (Lennon et al 2001) Lennon, M. et al., "Curvilinear Component Analysis for Nonlinear Dimensionality Reduction of Hyperspectral Images," *Proceedings of SPIE* 4541 (Image and Signal Processing for Remote Sensing VII) : 157-168 (2001)
- (Lessor, Hartman, and Gordon 1979) Lessor, D. L., J. S. Hartman, and R. L. Gordon, "Quantitative surface topography determination by Nomarski reflection microscopy. I. Theory," *Journal of Optical Society of America* 69 (2) : 357-366 (1979)
- (Levina and Bickel 2004) Levina, E. and Bickel, P. J., "Maximum Likelihood Estimation of Intrinsic Dimension," *Advances in Neural Information Processing Systems - NIPS* 2004. (2004)
- (Li et al. 2001) Li, Q. et al., "Optical profilometer based on the principle of differential interference," *Optical Engineering* 40 (5) : 833-836 (2001)
- (MacKay 1992) MacKay, D. J. C., "Bayesian interpolation," *Neural Computation* 4 (3) : 415-447 (1992)
- (Madani, Chebira, and Rybnik 2003a) Madani, K., A. Chebira, and M. Rybnik, "Data Driven Multiple Neural Network Models Generator Based on a Tree-like Scheduler," *Lec-*

ture Notes in Computer Science (Computational Methods in Neural Modeling) : 382-389 (2003)

(Madani, Chebira, and Rybnik 2003b) Madani, K., Chebira, A., and Rybnik, M., "Neural Network Based Evolutionary Treelike Multi-Models Generator Reducing Complexity on Both Data and Processing Levels," International Multi-conference on "Computational Engineering in Systems Applications" - CESA 2003. (2003)

(Marioge, 2000) Marioge, J.-P, *Surfaces Optiques*. Ed. EDP Sciences (2000)

(Martinetz and Schulten 1991) Martinetz, K. and Schulten, K., " A Neural Gas Network Learns Topologies," IEEE International Conference on Artificial Neural Networks. 1, 397-407. (1991)

(Nath et al. 1994) Nath, S., Wincheski, B., Fulton, J., and Namkung, M., "Classification of Magneto-Optic Images using Neural Networks," {ASNT} Conference Proceedings , 125-127. 1-3-1994 (1994)

(Oukhellou 1997) Oukhellou, L., "Paramétrisation et Classification de Signaux en Contrôle Non Destructif. Application à la Reconnaissance des Défauts de Rails par Courants de Foucault.", Thèse de Doctorat, Université Paris XI Orsay. (1997)

(Pilato 1998) Pilato, V., "Application des Réseaux de Neurones aux Méthodes de Mesure Basées sur l'Interaction Rayonnement Matière" Thèse de Doctorat, Université Paris XI Orsay. (1998)

(Ping et al. 2000) Ping, Z. et al., "Fourier-Mellin descriptor and interpolated feature space trajectories for three-dimensional object recognition," *Optical Engineering* 39 (5) : 1260-1266 (2000)

(Raginski and Lazebnik 2005) Raginski, M. and Lazebnik, S., "Estimation of Intrinsic Dimensionality Using High-Rate Vector Quantization," Proceedings of the 2005 Conference on Neural Information Processing Systems (Vancouver, Canada, December 2005), pp. 1105-1112 (2005)

(Raich, Costa, and Hero 2006) Raich, R, Costa, J. A., and Hero, A. O., "On Dimensionality Reduction for Classification and its Application," IEEE International Conference on Acoustics, Speech, and Signal Processing - ICASSP 2006. (2006)

(Rajpal, Chaudhury, and Banerjee 1999) Rajpal, N., S Chaudhury, and S. Banerjee, "Recognition of Partially Occluded Objects Using Neural Network Based Indexing," *Pattern Recognition* 32 : 1737-1749 (1999)

(Ravichandran and Trivedi 1995) Ravichandran, G. and M. Trivedi, "Circular-Mellin features for texture segmentation," *IEEE Trans. Image Processing* 4 : 1629-1640 (1995)

(Recce et al. 1996) Recce, M., Taylor, J., Plebe, A., and Tropicano, G., "High Speed Vision-Based Quality Grading of Oranges," 1996 International Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing (NICROSP '96). 136-145. (1996)

(Reschke, Bäumer, and Kross 1993) Reschke, T. J., S. M. Bäumer, and J. Kross, "Detection and Characterization of Local Defects on Polished Glass Surfaces using Differential Interference Contrast Microscopy," *Proceedings of SPIE* (16th Congress of the International Commission for Optics : Optics as a Key to High Technology) (1993)

(Roblin 1999) Roblin, G., "Microscopie Optique," *Techniques de l'Ingénieur* R6712 (1999)

(Roweis and Saul 2000) Roweis, S. T. and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science* 290 (5500) : 2323-2326 (2000)

(Rybnik 2006) Rybnik, M., "Contribution to the modelling and the exploitation of hybrid Multiple Neural Networks systems : application to intelligent processing of information.", Thèse de Doctorat - Université Paris XII-Val-de-Marne.(2006)

(Sammon 1969) Sammon, J. W., "A Nonlinear Mapping for Data Analysis," *IEEE Trans.on Computers* C-18 (5) : 401-409 (1969)

(Scholkopf, Smola, and Müller 1998) Scholkopf, B., A. J. Smola, and K. R. Müller, "Nonlinear Components Analysis as a Kernel Eigenvalue Problem," *Neural Computation (NIPS)* 10 (1998)

(Sekine et al. 2000) Sekine, A., Iwa, Y., Soma, H., Miki, N., Hisozaki, H., and Yoshino, H., "Surface inspection apparatus," Patent [US6104481]. 15-8-2000. (2000)

(Sheng and Lejeune 1991) Sheng, Y. and C. Lejeune, "Invariant Pattern Recognition Using Fourier-Mellin Transforms and Neural Networks," *Journal of Optics* 22 (5) : 223-228 (1991)

(Sheng and Shen 1994) Sheng, Y. and L. Shen, "Orthogonal Fourier-Mellin Moments for Invariant Pattern Recognition," *Journal of Optical Society of America* 6 : 1748-1757 (1994)

(Siegel and Gunatilake 1998) Siegel, M. and Gunatilake, P., "Remote Enhanced Visual Inspection of Aircraft by a Mobile Robot," Proceedings of the 1998 IMTC Conference, IEEE , 49-58. (1998)

(Silvén, Niskanen, and Kauppinen 2000) Silvén, O., M. Niskanen, and H. Kauppinen, "Wood Inspection with Non-Supervised Clustering," *Machine Vision and Applications* 13 (5) : 275-285 (2000)

(Tanaka et al. 1995) Tanaka, K., Ishiide, H., Sugihara, T., Utsunomiya, A., and Maki-mae, T., "Inspection apparatus with improved detection of surface defects over large and curved surface," Patent [US5237404]. 17-8-1995 (1995)

(Tarel 1996) Tarel, J.-P., "Recalage géométrique avec plusieurs prototypes," Rapport de Recherche - Institut National de Recherche en Informatique et en Automatique. 2988. (1996)

(Tenenbaum, de Silva, and Langford 2000) Tenenbaum, J. C., V. de Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science* 290 (5500) : 2319-2323 (2000)

(Teoh Beng Jin, Ngo Chek Ling, and Thian Song 2004) Teoh Beng Jin, A., D. Ngo Chek Ling, and O. Thian Song, "An efficient fingerprint verification system using integrated wavelet and Fourier-Mellin invariant transform," *Image and Vision Computing* 22 : 503-513 (2004)

(Tian, Itoh, and Yatagai 1995) Tian, X., M. Itoh, and Y. Yatagai, "Quantitative surface evaluation using synchronous Nomarski microscope," *Proceedings of SPIE* 2576 : 137-143 (1995)

(Tomie 2003) Tomie, T., "Method and apparatus for inspecting multilayer masks for defects," Patent [US2003067598]. 10-4-2003. (2003)

(Verleysen 2001) Verleysen, M., "Learning high-dimensional data," LFTNC'2001 - NATO Advanced Research Workshop on Limitations and Future Trends in Neural Computing. 22-24. 2001. (2001)

(Veropoulos 2001) Veropoulos, K., "Machine Learning Approaches to Medical Decision Making", PhD, University of Bristol. (2001)

(Veropoulos, Campbell, and Learmonth 1998) Veropoulos, K., Campbell, C., and Learmonth, G., "Image Processing and Neural Computing Used in the Diagnosis of Tuberculosis. Colloquium on Intelligent Methods in Healthcare and Medical Applications. (1998)

(Veropoulos et al. 1998) Veropoulos, K. et al., "The Automated Identification of Tubercle Bacilli in Sputum : A Preliminary Investigation," *Analytical and Quantitative Cytology and Histology* 21 (4) : 277-281 (1998)

(Vigneron 1997) Vigneron, V., "Learning Methods for Ill-Posed Problems. Application to γ -Spectrometry.", Thèse de Doctorat - Université d'Evry - Val d'Essonne. (1997)

(Voiry et al. 2005) Voiry, M., Houbre, F., Amarger, V., and Madani, K., "Toward Surface Imperfections Diagnosis Using Optical Microscopy Imaging in Industrial Environment", Proceedings of the Workshop IAR & ACD 2005 , 139-144. 18-11-2005. (2005)

(Voiry et al. 2006a) Voiry, M., Amarger, V., Madani, K., and Houbre, F., "Combining Image Processing and Self Organizing Artificial Neural Network Based Approaches for Industrial Process Faults Clustering," Proceedings of 13th International Multi-Conference on Advanced Computer Systems - ACS 2006 2, 129-138. 18-10-2006. (2006)

(Voiry et al. 2006b) Voiry, M., Madani, K., Amarger, V., and Houbre, F., "Toward Automatic Defects Clustering in Industrial Production Process Combining Optical Detection and Unsupervised Artificial Neural Network Techniques," Proceedings of the 2nd International Workshop on Artificial Neural Networks and Intelligent Information Processing - ANNIIP 2006 , 25-34, 25-8-2006 (2006).

(Voiry et al. 2007) Voiry, M., Madani, K., Amarger, V., and Bernier, J., "Impact of Data Dimensionality Reduction on Neural Based Classification : Application to Industrial Defects," Proceedings of the 3rd International Workshop on Artificial Neural Networks and Intelligent Information Processing - ANNIIP 2007, 56-65. (2007)

(Volker 2001) Volker, T. Committee Machines. In *Handbook for Neural Network Processing*. CRC Press.(2001)

(Weiss and Saranli 2004) Weiss, A. and Saranli, A., "Method and apparatus for flat and patterned media inspection," Patent [US2004109598]. 10-6-2004. (2004)

(Whitehouse 1985) Whitehouse, D. J., "Instrumentation for measuring finish, defects and gloss," *Proceedings of SPIE 525 (Measurement and Effects of Surface Defects and Quality of Polish)* : 106-123 (1985)

(Whitehouse 1997) Whitehouse, D. J., "Surface Metrology," *Measurement Science and Technology* 8 : 955-972 (1997)

(Wyant 1985) Wyant, J. C., "Optical profilers for surface roughness," *Proceedings of SPIE 525 (Measurement and Effects of Surface Defects and Quality of Polish)* : 174-180 (1985)

(Yasuda, Bannai, and Onami 1999) Yasuda, T., Bannai, H., and Onami, S., "Towards Automatic Construction of Cell-Lineage of *C. elegans* from Nomarski DIC Microscope Images," *Genome Informatics* 10 :144-154 (1999)

(Yeh et al. 2005) Yeh, M.-C., Lee, I.-H., Wu, G., Wu, Y., and Chang, E. Y., "Manifold Learning, a Promised Land or Work in Progress?," *IEEE International Conference on Multimedia and Expo ICME 2005*. (2005)

(Young et al. 1998) Young, D. et al., "Towards automatic cell identification in DIC microscopy," *Journal of Microscopy* 192 (2) : 186-193 (1998)

(Young and Gray 1996) Young, D. and Gray, A. J., "Cell identification in Differential Interference Contrast microscope images using edge detection," *Proc.7th British Machine Vision Conference [BMVA Press]*, 133-142. Edinburgh.(1996)

(Zhang 2000) Zhang, G. P., "Neural Networks for Classification : A Survey," *IEEE Trans.on Systems, Man, and Cybernetics - Part C : Applications and Reviews* 30 (4) : 451-462 (2000)

Bibliographie de l'auteur

Chapitre de livre

Voiry, M., Amarger, V., Madani, K. et F. Houbre : Design and Prototyping of an Industrial Fault Clustering System Combining Image Processing and Artificial Neural Network Based Approaches. *In Advances in Information Processing and Protection*, Part III, pages 351-360. Jerzy Pejas and Khalid Saeed, Springer US, 2007. ISBN 978-0-387-73136-0.

Voiry, M., Madani, K., Amarger, V. et J. Bernier : Optical Devices Diagnosis by Neural Classifier Exploiting Invariant Data Representation and Dimensionality Reduction Ability. *In Lecture Notes in Computer Science (LNCS 4507) : Computational and Ambient Intelligence*, Proceedings of the 9th International Work-Conference on Artificial Neural Networks, IWANN 2007, San Sebastián, Spain, June 20-22, 2007, pages 1098-1105. Francisco Sandoval, Alberto Prieto, Joan Cabestany, Manuel Graña, Springer Berlin / Heidelberg, 2007. ISBN 978-3-540-73006-4.

Voiry, M., Amarger, V., Madani, K. et J. Bernier : ANN Based Defects' Diagnosis of Industrial Optical Devices. *In Encyclopedia of Artificial Intelligence*. Juan Ramón Rabuñal Dopico, Julian Dorado, Alejandro Pazos, Information Science Reference. ISBN : 978-1-59904-849-9.

Publication dans des revues internationales

Madani, K, Voiry, M., Amarger, V., Kanaoui, N., Chohra, A. et F. Houbre : Computer Aided Diagnosis Using Soft-Computing Techniques and Image's Issued Representation : Application to Biomedical and Industrial Fields, *International Journal of Computing* 5 (3) : 43-53 (2006).

Communication dans des congrès internationaux avec actes et comité de lecture

Voiry, M., Houbre, F., Amarger, V. et K. Madani : Toward Surface Imperfections Diagnosis Using Optical Microscopy Imaging in Industrial Environment. *In Workshop on Advanced Control and Diagnosis (IAR-ACD)*, Mulhouse, France, 17-18 Novembre 2005.

Voiry, M., Madani, K., Amarger, V., et F. Houbre : Toward Automatic Defects Clustering in Industrial Production Process Combining Optical Detection and Unsupervised Artificial Neural Network Techniques. *In 2nd International Workshop on Artificial Neural Networks and Intelligent Information Processing (ANNIIP)*, Setúbal, Portugal, 25-28 Août 2006.

Voiry, M., Amarger, V., Madani, K., et F. Houbre : Combining Image Processing and Self Organizing Artificial Neural Network Based Approaches for Industrial Process Faults Clustering. *In 13th International Multi-Conference on Advanced Computer Systems and Computer Information Systems for Industrial Management (ACS-CISIM)*, Międzyzdroje, Pologne, 18-20 Octobre 2006.

Voiry, M., Madani, K., Amarger, V. et J. Bernier : Impact of Data Dimensionality Reduction on Neural Based Classification : Application to Industrial Defects. *3rd International Workshop on Artificial Neural Networks and Intelligent Information Processing (ANNIIP)*, Angers, France, 10-12 Mai 2007

Résumé

Dans différents domaines industriels, la problématique du diagnostic prend une place importante. Ainsi, le contrôle d'aspect des composants optiques est une étape incontournable pour garantir leurs performances opérationnelles. La méthode conventionnelle de contrôle par un opérateur humain souffre de limitations importantes qui deviennent insurmontables pour certaines optiques hautes performances. Dans ce contexte, cette thèse traite de la conception d'un système automatique capable d'assurer le contrôle d'aspect. Premièrement, une étude des capteurs pouvant être mis en œuvre par ce système est menée. Afin de satisfaire à des contraintes de temps de contrôle, la solution proposée utilise deux capteurs travaillant à des échelles différentes. Un de ces capteurs est basé sur la microscopie Nomarski ; nous présentons ce capteur ainsi qu'un ensemble de méthodes de traitement de l'image qui permettent, à partir des données fournies par celui-ci, de détecter les défauts et de déterminer la rugosité, de manière robuste et répétable. L'élaboration d'un prototype opérationnel, capable de contrôler des pièces optiques de taille limitée valide ces différentes techniques. Par ailleurs, le diagnostic des composants optiques nécessite une phase de classification. En effet, si les défauts permanents sont détectés, il en est de même pour de nombreux « faux » défauts (poussières, traces de nettoyage. . .). Ce problème complexe est traité par un réseau de neurones artificiels de type MLP tirant partie d'une description invariante des défauts. Cette description, issue de la transformée de Fourier-Mellin est d'une dimension élevée qui peut poser des problèmes liés au « fléau de la dimension ». Afin de limiter ces effets néfastes, différentes techniques de réduction de dimension (Self Organizing Map, Curvilinear Component Analysis et Curvilinear Distance Analysis) sont étudiées. On montre d'une part que les techniques CCA et CDA sont plus performantes que SOM en termes de qualité de projection, et d'autre part qu'elles permettent d'utiliser des classifieurs de taille plus modeste, à performances égales. Enfin, un réseau de neurones modulaire utilisant des modèles locaux est proposé. Nous développons une nouvelle approche de décomposition des problèmes de classification, fondée sur le concept de dimension intrinsèque. Les groupes de données de dimensionnalité homogène obtenus ont un sens physique et permettent de réduire considérablement la phase d'apprentissage du classifieur tout en améliorant ses performances en généralisation.

Mots-clés : Composants optiques, Contrôle d'aspect, Capteurs sans contact, Microscopie Nomarski, Réseaux de neurones MLP, Réduction de dimension, Classification modulaire.

Study and design of an automatic system for optical devices diagnosis based on connexionist approaches

Abstract

In various industrial fields, the problem of diagnosis is of great interest. For example, the check of surface imperfections on an optical device is necessary to guarantee its operational performances. The conventional control method, based on human expert visual inspection, suffers from limitations, which become critical for some high-performances components. In this context, this thesis deals with the design of an automatic system, able to carry out the diagnosis of appearance flaws. To fulfil the time constraints, the suggested solution uses two sensors working on different scales. We present one of them based on Normarski microscopy, and the image processing methods which allow, starting from issued data, to detect the defects and to determine roughness in a reliable way. The development of an operational prototype, able to check small optical components, validates the proposed techniques. The final diagnosis also requires a classification phase. Indeed, if the permanent defects are detected, many "false" defects (dust, cleaning marks. . .) are emphasized as well. This complex problem is solved by a MLP Artificial Neural Network using an invariant description of the defects. This representation, resulting from the Fourier-Mellin transform, is a high dimensional vector, what implies some problems linked to the "curse of dimensionality". In order to limit these harmful effects, various dimensionality reduction techniques (Self Organizing Map, Curvilinear Component Analysis and Curvilinear Distance Analysis) are investigated. On one hand we show that CCA and CDA are more powerful than SOM in terms of projection quality. On the other hand, these methods allow using more simple classifiers with equal performances. Finally, a modular neural network, which exploits local models, is developed. We proposed a new classification problems decomposition scheme, based on the intrinsic dimension concept. The obtained data clusters of homogeneous dimensionality have a physical meaning and permit to reduce significantly the training phase of the classifier, while improving its generalization performances.

Key-words : Optical components, surface flaws diagnosis, Contactless sensors, Nomarski Microscopy, MLP Artificial Neural Network, Dimensionality reduction, Modular classification.

