UNIVERSITE DE CORSE-PASCAL PAOLI

ECOLE DOCTORALE ENVIRONNEMENT ET SOCIETE

UMR CNRS 6134 S.P.E.

THESE

présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ DE CORSE
en Informatique

Soutenue publiquement par

Jean-Sébastien GUALTIERI

le 19 décembre 2013

Application de réseaux de capteurs sans fils : apprentissage de chants polyphoniques corses

---

Directeurs :

M. Jean-François SANTUCCI, Professeur, Università di Corsica

M. Dominique FEDERICI, Professeur, Università di Corsica

Rapporteurs :

Mme Bégout Marie-Laure, Dr-HDR, Ifremer

M. Rousseaux Francis, Professeur, Université de Reims

Jury :

Mme Bégout Marie-Laure, Dr-HDR, Ifremer

M. Rousseaux Francis, Professeur, Université de Reims

M. Chabrier Jean-Jacques, Professeur, Université de Bourgogne

M. Aiello Antoine, Professeur, Université de Corse

M. Santucci Jean-François, Professeur, Université de Corse

M. Federici Dominique, Professeur, Université de Corse

Invité :

M. Jean-François Rossi, Directeur de la SITEC

UNIVERSITE DE CORSE-PASCAL PAOLI

ECOLE DOCTORALE ENVIRONNEMENT ET SOCIETE

UMR CNRS 6134 S.P.E.

THESE

présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ DE CORSE
en Informatique

Soutenue publiquement par

Jean-Sébastien GUALTIERI

le 19 décembre 2013

Application de réseaux de capteurs sans fils : apprentissage de chants polyphoniques corses

---

Directeurs :

M. Jean-François SANTUCCI, Professeur, Università di Corsica

M. Dominique FEDERICI, Professeur, Università di Corsica

Rapporteurs :

Mme Bégout Marie-Laure, Dr-HDR, Ifremer

M. Rousseaux Francis, Professeur, Université de Reims

Jury :

Mme Bégout Marie-Laure, Dr-HDR, Ifremer

M. Rousseaux Francis, Professeur, Université de Reims

M. Chabrier Jean-Jacques, Professeur, Université de Bourgogne

M. Aiello Antoine, Professeur, Université de Corse

M. Santucci Jean-François, Professeur, Université de Corse

M. Federici Dominique, Professeur, Université de Corse

Invité :

M. Jean-François Rossi, Directeur de la SITEC

pè i mò parenti

chì avete fattu ciò ch'eiu sò

Table des matières

REMERCIEMENTS **1**

RÉSUMÉ **3**

ABSTRACT **4**

RIASSUNTU **5**

CHAPITRE I : INTRODUCTION GÉNÉRALE **6**

1. CONTEXTE 7
2. PROBLÉMATIQUE ET APPROCHE DE RÉSOLUTION 9
3. ORGANISATION DU DOCUMENT 10

CHAPITRE II : GESTES ET CAPTATION **13**

1. ETAT DE L'ART 14
   1.1. Définition du geste *14*
   1.2. Contrôleur gestuel *15*
   1.3. Acquisition des données *16*
2. MIMOPHONIE 17
   2.1. La notation *19*
   2.2. La gestuelle *19*
3. CAPTATION DES GESTES 21
   3.1. Les interfaces gestuelles *21*
   3.2. Les systèmes de captation *22*
   3.3. Réalisation des gants *24*
      3.3.1. Contraintes et design électronique 24
      3.3.2. Support 25
      3.3.3. Flexion des doigts 25
      3.3.4. Orientation de la main 26
      3.3.5. Numérisation des capteurs 26
      3.3.6. Système de transmission sans fil 27
4. CONCLUSION 30

CHAPITRE III : RECONNAISSANCE DE GESTES **31**

1. ETAT DE L'ART 32
2. SYSTÈMES D'INFÉRENCE FLOUS 35
   2.1. Systèmes flous *36*
   2.2. Sous ensembles flous *37*

2.3. SIF : système d'inférence flou ... *38*
2.4. Fuzzification et défuzzification ... *39*
2.5. Description du système ... *39*
2.5.1. Entrées ... 40
2.5.2. Sorties ... 41
2.5.3. Règles ... 42
2.5.4. Résultats ... 42
3. FTM ... 44
4. DISCUSSION ET CONCLUSION ... 46
CHAPITRE IV : PRODUCTION SONORE ... 48
1. VOIX SYNTHÉTISÉES ... 49
1.1. Description de l'appareil vocal ... *49*
1.2. La synthèse de voix chantée avec le programme CHANT ... *50*
2. VOIX ENREGISTRÉES ... 54
3. CONCLUSION ... 55
CHAPITRE V : MISE EN ŒUVRE ET DÉVELOPPEMENT ... 56
1. DÉVELOPPEMENT DES MODULES ... 57
CHAPITRE VI : RÉSEAUX DE CAPTEURS SANS FIL : SUIVI D'ESPÈCES HALIEUTIQUES ... 63
1. TRAVAUX EN COURS ... 65
1.1. Suivi des araignées de mer ... *67*
1.1.1. Matériel utilisé ... 67
1.1.2. Informations sur l'espèce : Maja squinado ... 68
1.1.3. Suivi des individus ... 69
1.1.4. Conclusion ... 73
1.2. Suivi d'espèces temps réel ... *74*
2. CONCLUSION ET PERSPECTIVES ... 77
CHAPITRE VII : CONCLUSION GÉNÉRALE ... 79
1. BILAN DES TRAVAUX ... 80
2. PERSPECTIVES DE RECHERCHE ... 82
BIBLIOGRAPHIE ... 85
LISTE DES FIGURES ... 97

Remerciements

Ces travaux de thèse se sont principalement déroulés au sein de l'équipe Informatique de l'Università di Corsica. Cette équipe est intégrée au laboratoire Sciences Pour l'Environnement (SPE), UMR CNRS 6134 dirigée par Monsieur Paul-Antoine BISGAMBIGLIA, professeur à l'Università di Corsica. J'ai trouvé au sein de cette équipe une ambiance particulièrement favorable pour mener à bien ces travaux.

Je voudrais en premier lieu remercier M. Jean-François SANTUCCI, professeur à l'Università di Corsica, pour m'avoir guidé depuis ma première année à Université, pour m'avoir proposé ce sujet de thèse et pour avoir dirigé ces travaux. Ses compétences, ses conseils, son expérience et sa gentillesse m'ont été indispensables tout au long de ces années.

Je tiens également à remercier chaleureusement M. Dominique FEDERICI, professeur à l'Università di Corsica, pour avoir co-encadré ces travaux.

Ma reconnaissance va également à M. Antoine AIELLO, professeur à l'Università di Corsica et directeur de l'UMS CNRS 3514 STELLA MARE, pour ses conseils, pour l'intérêt qu'il porte à mes travaux en acceptant de participer à ce jury et pour m'avoir vivement encouragé à soutenir ma thèse avant la fin de l'année.

Je remercie M. Jean-François ROSSI, Directeur de la SITEC, pour m'avoir recruté et permis de mener mes travaux de thèse dans les meilleures conditions possibles.

Je souhaite exprimer toute ma gratitude à Mme Marie-Laure BEGOUT, Docteur-HDR au Laboratoire Ressources Halieutiques de La Rochelle et M. Francis ROUSSEAUX, Professeur de l'Université de Reims, pour l'intérêt qu'ils ont porté à ce travail en acceptant de juger ce mémoire avec autant d'attention.

M. Jean-Jacques CHABRIER, professeur émérite de l'Université de Bourgogne, a accepté de participer à ce jury et d'en être le président : qu'il trouve ici l'expression de ma profonde reconnaissance.

Un grand merci à toute l'équipe de l'association E Voce di U Cumune, et tout particulièrement à Nando AQUAVIVA, Claude BELLAGAMBA, Tony et Nicole CASALONGA pour leur grande gentillesse et pour leur disponibilité. Ils m'ont ouvert à la richesse de notre patrimoine musical et c'est grâce à eux si aujourd'hui je peux transmettre

cette passion à mes enfants. J'espère que notre collaboration et nos réunions à Pigna continueront encore longtemps.

Je tiens à remercier également M. Xavier RODET, professeur émérite à l'IRCAM, ainsi que les équipes analyse/synthèse et interactions musicales temps réel pour leur accueil à l'IRCAM et leur disponibilité.

J'exprime ici toute mon amitié à mes compagnons de thèse Stéphane GARREDU et Paul-Antoine BISGAMBIGLIA, qui ont contribué à rendre ces années toujours joviales, ainsi qu'à toute l'équipe de STELLA MARE pour leur solidarité et leur soutient inconditionnel.

Mes remerciements vont aussi à ma femme Hélène, ma sœur Anne-Laure, tout ma famille et mes amis qui, avec cette question récurrente, « quand est-ce que tu la soutiens cette thèse ? », bien qu'angoissante en période fréquente de doutes, m'ont permis de ne jamais dévier de mon objectif.

Ces remerciements ne peuvent s'achever, sans une pensée pour mes premiers fans (et correcteur des fautes d'orthographes de cette thèse) : mes parents. Leur présence et leurs encouragements sont pour moi les piliers fondateurs de ce que je suis et de ce que je fais. C'est à eux que je dédie mes travaux.

Résumé

Ce rapport présente les différentes étapes suivies afin de développer une méthode logicielle pour l'apprentissage de chants polyphoniques basée sur l'utilisation de capteurs sans fil. Cette étude réalisée au sein de l'UMR SPE de l'Université de Corse, en collaboration avec l'association E Voce di U Cumune et l'IRCAM, s'inscrit dans un projet universitaire de valorisation et de caractérisation du patrimoine culturel corse.

L'objectif de cette thèse est de proposer un outil d'aide à l'apprentissage de polyphonies. Pour cela nous avons utilisé un langage gestuel décrivant l'échelle musicale corse. Le but de nos travaux a été de capter ces gestes afin d'y associer, à chacun d'entre eux, une voix chantée. Ce processus qui va du langage gestuel au son fait appel à plusieurs notions qui sont :

- le contrôleur gestuel, qui regroupe un ensemble de capteurs au sein d'une même unité pour constituer le premier bloc fonctionnel du modèle d'instrument virtuel.

- la captation de gestes, qui permet l'acquisition, c'est-à-dire qui traduit en signaux électriques puis numériques, des gestes effectués.

- la reconnaissance, qui identifie les gestes captés et permet de différencier les postures de la main.

- la production sonore, la dernière étape du système, qui associe à chacun des gestes effectués une voix chantée à la hauteur de note souhaitée.

Cet enchainement nous a conduit à proposer un nouvel outil d'aide à l'apprentissage de chants polyphoniques qui offre la possibilité de manipuler les voix, le chant et la polyphonie sans n'avoir aucune connaissance musicale.

Abstract

This report presents the different steps to develop a software method for learning polyphonic songs based on the use of wireless sensors. This study within the UMR SPE University of Corsica, in collaboration with the association E Voce di U Cumune and the IRCAM, is part of a university project development and characterization of Corsican cultural heritage.

The objective of this thesis is to provide a tool for learning polyphony. To do this we used sign language describing the Corsican musical scale. The aim of our work was to capture these gestures in order to associate each of them with a singing voice. This process which established the link between gestural language and singing involves several concepts which are:

- Hand controller, which includes a set of sensors in a single unit to be the first building block of virtual instrument model.

- The gestural capturing which allows acquisition, i.e. which translates into electrical signals and digital, the gestures performed.

- Recognition that identifies the captured gestures and can differentiate hand posture.

- Sound production is the final stage of the system and associated with each of the gestures made a singing voice to match desired note.

This concatenation led us to propose a new tool for learning polyphonic songs which offers the possibility to manipulate voice, chant and polyphony without having no musical knowledge.

Riassuntu

Stu raportu prisenta e sfarente tappe seguitate da sviluppà un metudu "software" per l'amparera di canti pulifonichi fundatu nantu à l'adopru di cattadori senza filu.

Stu studiu realizatu à l'UMR SPE di l'Università di Corsica, in cullaburazione cun l'associu di E Voce di u Cumunu è l'IRCAM, si scrive in un prughjettu universitariu di valurisazione di u patrimoniu culturale corsu.

L'ughjettivu di sta tesa hè di prupone un arnese d'aiutu à l'amparera di pulifunie. Per quessa, avemu adupratu un linguaghju gestuale chì discrive una scala musicale corsa. U scopu di u nostru travagliu hè statu di chjappà issi gesti di manera à assucià, à ognunu, una voce cantata. Stu prucessu chì và da u linguaghju gestuale à u sonu chjama diverse nuzione chì sò :

-u cuntrollu gestuale, chì gruppa un inseme di cattadori in una stessa unità per cunstituisce u primu bloccu funziunale di mudellu di strumentu virtuale.

-a cattazione di i gesti, chì permette l'acquistu, vene à dì chì traduce in signali elettrichi eppo numerichi, di i gesti effettuati.

-a ricunniscenza, chì identificheghja i gesti cattati è permette di sfarenzià e pusture di a manu.

-a produzzione sonora hè l'ultima tappa di u sistema è assucieghja à ogni gestu effettuatu una voce cantata à l'altura di a nota vulsuta.

St´intrecciu ci hà cunduttu à prupone un nuvellu arnese d'aiutu à l'amparera di canti pulifonichi chì offre à pussibilità di manighjà e voci, u cantu è a pulifunia senza avè alcuna cunniscenza musicale

CHAPITRE I :

Introduction générale

1. Contexte

Aucune application ne permet l'apprentissage de chants polyphoniques corses, la transmission de ce savoir est encore véhiculée par l'oralité. C'est bien le signe de sa vitalité, de sa popularité, de son authenticité, mais également et paradoxalement de sa fragilité. Les évolutions de la société corse, la perméabilité aux influences extérieures transforment la culture musicale insulaire. La Corse n'échappe pas à la mondialisation, aux stratégies mercantiles, à la colonisation des consciences. Sous la pression de l'ingénierie du marketing la tradition du chant polyphonique corse peu à peu s'altère dans son esthétique et dans ses fonctions.

Depuis les années soixante-dix, la polyphonie s'est taillée une notoriété internationale. La présence des Nouvelles Polyphonies Corses lors de la cérémonie d'ouverture des Jeux Olympiques d'Albertville, en 1992, en illustre la consécration. Aujourd'hui, l'immersion quotidienne, subie en Corse comme ailleurs, dans cette expression musicale, crée une situation de « diglossie » où se confrontent deux langages. D'une part la modalité à laquelle appartient la musique traditionnelle et de l'autre la tonalité issue de la musique occidentale savante, influente par son tempérament, jusque dans les formes plus populaires des musiques de variété.

Le chercheur n'a pas à émettre de jugement de valeur sur les métamorphoses observables puisque la tradition n'est pas figée dans des règles immuables. Cependant, il ne peut ignorer l'héritage ethnographique des premières missions de collectage[1], recueilli au siècle dernier à une époque où l'enclavement d'une société rurale encore dynamique, ajouté à l'insularité rendaient plus difficile le brassage culturel. Cette documentation sonore déposée aujourd'hui au Musée de la Corse, étoffée par d'autres collectages plus récents, est accessible à l'écoute et révèle un patrimoine musical populaire surprenant par sa richesse et la diversité de ses sources.

La simplicité apparente du chant corse est trompeuse car outre les qualités vocales et une technique exigeante qu'il requiert, il emprunte une graduation que la musique tempérée ne traduit que partiellement.

[1] MNATP : Mission ethnographique Arrighi/Quilici, 1948 ; Collectage Félix Quilici pour l'ORTF, 1949

Plutôt que d'opposer les nouvelles technologies et des traditions ancestrales, nous allons, dans ce manuscrit, montrer comment nous avons développé une application logicielle permettant l'apprentissage de chants polyphoniques.

Les travaux que nous présentons font partie d'un projet universitaire de caractérisation et de valorisation du patrimoine culturel corse. Ces travaux sont menés par l'Université de Corse en collaboration avec l'Association « E Voce di U Cumune », qui depuis les années soixante-dix exerce une action de collectage et de valorisation du patrimoine musical corse et l'IRCAM (Institut de Recherche Acoustique/Musique)[2]. L'une des missions de l'IRCAM est d'accorder une place importante à la création et à la promotion de musiques nouvelles, aux frontières entre art et sciences. Cette ambition en a fait l'un des acteurs importants de l'histoire de la musique contemporaine, comme en attestent notamment ses contributions à la musique mixte. Le terme de musique mixte désignant toute œuvre de musique impliquant un interprète humain produisant un son acoustique (d'un instrument ou d'une voix) et des sons électroniques, échantillonnés, ou traités, "joués" par une machine.

Le but de nos travaux est la création d'une méthode logicielle permettant l'apprentissage de chants polyphoniques. Ce logiciel doit permettre à un chanteur unique de créer une polyphonie. Ceci est possible en remplaçant les autres chanteurs par des voix synthétisées [1]. D'où la problématique suivante : comment faire des polyphonies traditionnelles à partir de voix artificielles ?

Nous avons pour cela utilisé un langage gestuel : la mimophonie. Grâce à cet ensemble de gestes [2], le chanteur principal peut guider et jouer d'autres voix chantées. Nous avons donc entrepris la création de gants de données afin de capter ces gestes, afin de pouvoir y associer une voix.

Depuis 2005 une collaboration réunit l'Università di Corsica – Pasquale Paoli, l'IRCAM et l'association E Voce di U Cumune autour du projet de recherche : « Du geste au son, la polyphonie corse question de méthode. ». La recherche s'intègre au projet de création d'un « espace chant » au sein de la future médiathèque, prévue dans le programme « Inventaire et valorisation du patrimoine culturel » de l'université de Corse. Elle se donne pour objectif la création d'un outil pédagogique numérique qui permette l'analyse et l'apprentissage de la musique corse. Le projet s'articule autour de deux champs d'étude :

[2] http://www.ircam.fr - expert mondial en création d'outils musicaux

- De la mimophonie à la technologie, faisabilité technique, captation des gestes, synthèse de la voix.

- Recueil, analyse et caractérisation des chants corses.

Le premier est présenté dans ce manuscrit, le second est attribué à un autre doctorant : Claude Bellagamba.

2. Problématique et approche de résolution

Le but de nos travaux est la réalisation d'un outil permettant l'apprentissage de chants polyphoniques traditionnels. Pour cela, nous devons utiliser un modèle qui décrit les intervalles du chant traditionnel. Celui-ci doit être associé à un langage gestuel. En effet, l'enseignant et l'apprenant ont tous deux besoin à la fois de chanter et d'utiliser des gestes : le premier pour guider son élève, le second pour produire des sons. Il est donc nécessaire de créer un outil d'acquisition des données qui permettra de capter chacun des gestes. Une fois le geste capté, il faut y associer une voix chantées à la hauteur de note voulue. Nous avons dû pour cela utiliser des voix chantées de synthèse. Dans un premier temps, nous avons fait correspondre une voix à chacun des gestes par simple concordance. Un geste étant représenté par un intervalle de valeurs, il suffit d'y associer un son. Cependant, l'intégration d'un système de reconnaissance de gestes nous a vite semblé indispensable. En effet, de trop grandes variations dans les données d'acquisition des gestes, dues à la diversité des utilisateurs du système ou à l'imprécision des gestes, rendaient les correspondances aléatoires.

Dans le but de résoudre cette problématique, nous pouvons diviser notre projet en trois domaines liés entre eux : la captation des gestes et l'acquisition des données, la reconnaissance des gestes et la production sonore.

Comme précisé ci-dessus, nos travaux s'articuleront autour de trois domaines. La figure 1 reprend le raisonnement que nous avons suivi pour la réalisation de notre projet et qui va de l'étude des chants tradionels à la production de voix chantées.

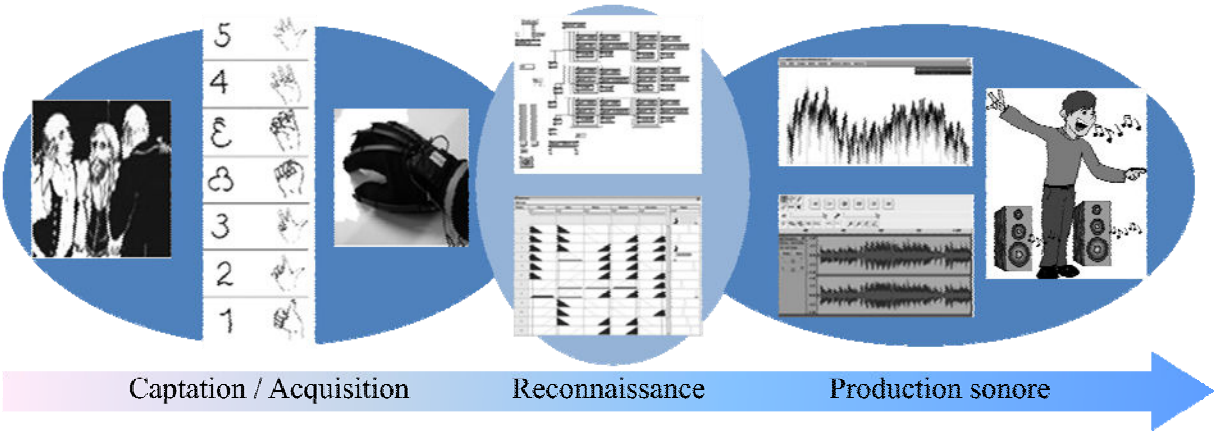


Figure 1 : processus "du geste au son"

La première étape de notre démarche a été d'utiliser un langage permettant de décrire l'échelle musicale traditionnelle corse. Dans une démarche pédagogique, ce langage a été couplé à un support gestuel. L'association de ce langage chiffré à un langage gestuel a été conçue par l'association E Voce di u Cumune et baptisée la mimophonie. Pour pouvoir utiliser ces gestes de manière numérique, il est nécessaire de pouvoir les capter et d'y attribuer une valeur en temps réel. La non-homogénéisation de chaque geste réalisé avec des gants de données nous a mené au développement d'une solution de reconnaissance des gestes. Puis il a fallu associer à chacun des gestes du langage un son musical.

3. Organisation du document

La réalisation d'un tel projet fait appel à plusieurs domaines différents. En effet, ce document aborde des notions telles que la musicologie, la captation et la reconnaissance de gestes et la synthèse de voix chantées. Ces différents thèmes, réunis ici dans un même processus (Cf. Figure 1), seront présentés indépendamment les uns des autres. Chacune des parties de ce rapport commencera donc par un état de l'art du domaine.

Le second chapitre décrit le processus de captation du geste qui va de la description du langage mimophonique à l'acquisition de données grâce à des gants. Nous commençons cette partie par la définition de notions – comme : le geste, le contrôleur gestuel et l'acquisition de données – utilisées dans le domaine musical.

Nous présentons ensuite le langage gestuel utilisé pour l'apprentissage de chants polyphoniques, la mimophonie. Après un bref historique sur le contexte dans lequel a été créé

ce langage, nous donnons une description de sa notion, utilisée pour la transcription de partitions ainsi que pour la création d'œuvres musicales, et de sa gestuelle qui permet l'apprentissage du chant.

Nous détaillons ensuite la captation du geste mimophonique. Cette captation est effectuée à l'aide de gants de données. Nous présentons la réalisation technique de ces gants.

Le troisième chapitre traite de la reconnaissance de gestes. Au vu du grand nombre de méthodes existantes, il ne nous semble ni judicieux ni pertinent de toutes les décrire. Nous faisons, dans ce chapitre, un état de l'art des systèmes les plus utilisés, notamment à l'IRCAM. Après avoir justifié de nos choix, nous décrivons les deux méthodes qui nous semblent les mieux adaptées pour la reconnaissance de gestes : les Systèmes d'Inférence Flous (SIF) et la bibliothèque FTM. Les SIF utilisent une représentation des connaissances sous forme de règles, une manière simple d'expliquer des procédés de décision. Ils utilisent pour cela des variables linguistiques qui sont proches de la représentation humaine. FTM est une bibliothèque partagée qui fournit dans MAX/MSP un ensemble d'objets utiles pour le traitement de la musique, le son et des données de capture de mouvement.

Nous terminons cette partie par une discussion sur ces deux méthodes de reconnaissance.

Dans le quatrième chapitre, nous décrivons comment, une fois le geste capté et identifié, nous l'associons à une voix chantée. Deux méthodes ont été utilisées, tout d'abord des voix enregistrées et couplées au système de reconnaissance par un *sampler*, ensuite des voix synthétisées, créées à partir de MAX/MSP. La première méthode, qui permet d'obtenir un rendu plus proche de la réalité, est certainement plus appropriée pour l'apprentissage et la représentation. La seconde, du fait que l'on peut en modifier tous les paramètres (hauteur, intensité, vibrato, etc.) favorise l'expérimentation.

Dans le cinquième chapitre, nous revenons sur le projet global de création d'un outil d'apprentissage du chant polyphonique en donnant une description du développement. A chaque étape du processus « geste-son » correspond un module développé avec MAX/MSP. Nous détaillons donc le module qui permet la captation de données émises par les gants, celui utilisé pour la reconnaissance du geste et enfin le module de production sonore.

Afin de valider notre application, nous l'avons proposé à un enseignant de chant polyphonique. Nous donnons ses impressions, les observations que nous avons faites lors de

la prise en main d'une personne non experte, ainsi que les améliorations que nous pouvons y apporter.

Ces travaux de doctorat m'ont permis d'acquérir des compétences aussi bien en acoustique, dans le domaine des capteurs sans fil et du développement d'application, mais également de méthodologie de travail et de gestion de projet.

Suite à mes travaux de recherche sur l'informatique musicale, j'ai pu apporter mes compétences dans un nouveau domaine scientifique et dans un nouveau projet de recherche portant sur la modélisation comportementale d'espèces halieutiques. Ces travaux portant, entre autre, sur l'utilisation de capteurs acoustiques, seront détaillés dans le chapitre six.

Enfin, en conclusion, nous revenons sur les différentes étapes de ces travaux et présentons plusieurs perspectives de recherche. En effet, outre le fait que l'on peut évidemment améliorer l'application précédemment présentée, la mise en œuvre de celle-ci nous a montré qu'il serait intéressant de développer d'autres modules complémentaires. Notamment, un éditeur de partitions utilisant les symboles de la mimophonie, ou un système qui permettrait d'ajuster en temps réel la hauteur des voix synthétisées par rapport à la tonalité du chanteur.

Ces travaux pourraient alors venir compléter l'espace musicologie de la médiathèque de l'Université de Corse.

CHAPITRE II :

Gestes et Captation

La relation entre le geste et le son a fait, depuis de nombreuses années, l'objet d'expériences et de travaux le plus souvent orientés à partir du geste pour générer une source sonore. Par exemple, le corps d'un danseur doté de capteurs permet une traduction musicale de l'expression kinesthésique. L'approche envisagée dans nos travaux concerne la production d'un son (une voix chantée) à partir d'un geste. Il faut tout d'abord trouver un langage gestuel et le rendre utilisable par notre outil.

Rappelons que le but de nos travaux est de développer un outil d'apprentissage des chants polyphoniques. Pour cela, nous utilisons un langage gestuel qui permet à une personne d'en guider deux autres. Ces travaux s'appuient sur un état de l'art que nous exposons ci-après. Nous présentons, dans la seconde partie, le langage mimophonique. Ce dernier étant défini par une représentation chiffrée et une représentation gestuelle. La troisième partie porte sur la captation de ces gestes. Puis nous terminons ce chapitre par une description des gants de données que nous avons réalisés.

1. Etat de l'art

1.1. Définition du geste

Plusieurs travaux ont proposé une classification des gestes en musique que le lecteur pourra retrouver dans les travaux [3], [4], [5] et[6].

Dans le but de considérer tout type de contrôleur nous avons cherché dans la littérature différentes classifications du geste, en particulier les gestes d'interaction dans un contexte musical. Dans [7] et [8], on trouvera une étude quasi-exhaustive des gestes musicaux. Nous avons retenu deux classifications significatives du geste :

- gestes primitifs : D'après Insook Choi [9][10], les gestes primitifs sont « les mouvements fondamentaux d'un sujet humain comme réponses dynamiques à un environnement». Chaque primitive peut incorporer d'autres mouvements, tant qu'elle caractérise l'intention première de l'exécutant. Il existe trois types de gestes primitifs :
    - basés sur des trajectoires (exemple : changements d'orientation),
    - basés sur la force (exemple : gradients de mouvements),
    - basés sur des motifs (exemple : mouvements quasi-périodiques),
- classification des gestes en trois classes proposée par François Delalande [11] :

• geste effecteur : regroupe l'ensemble des mouvements effectués pour produire mécaniquement un son,

• geste accompagnateur : rend compte des mouvements qui engagent le corps dans son entier : mimiques, gestes des épaules, ...,

• geste figuratif : expressions purement symboliques perçues par l'auditeur comme des articulations dans une mélodie : par exemple, un mouvement du corps vers le bas pour signifier la fin d'une phrase mélodique. Canazza et al. ont effectué une étude du geste figuratif sur l'expressivité du jeu de la clarinette [12] [13].

Nous nous intéressons particulièrement dans notre étude au geste effecteur, que Claude Cadoz définit comme *le geste instrumental*.

Comme leur nom l'indique, les *gestes instrumentaux* sont les mouvements de l'interprète directement reliés à la production du son. Ces gestes sont pour le musicien l'interface directe de contrôle du son de l'instrument. Dans le cas de notre application d'apprentissage des chants polyphoniques, les gestes de la main gauche guident le chant de la basse et ceux de la main droite le chant de la tierce.

Avant de continuer, il faut donc préciser ce que nous entendons par *geste*. Afin de se restreindre et pour éviter de discuter ici toutes les nuances possibles que peut revêtir ce mot, nous nous limiterons au seul domaine musical. Nous appelons ici *geste* les actions physiques effectuées par l'utilisateur afin de produire un son. Cette définition est en fait celle du *geste instrumental*. Cependant, le terme *instrumental* n'est pas approprié dans notre situation car le son retourné ne provient pas d'un instrument de musique mais d'une voix. Nous parlerons donc de *gestes de la mimophonie* ou de *gestes mimophoniques*.

1.2. Contrôleur gestuel

Le foisonnement des nouvelles interfaces de contrôle gestuel et la part grandissante attribuée au domaine du contrôle et de la performance lors des conférences en informatique et en acoustique musicales (ICMC, DAFx, NIME, ICA, ISMA, entre autres) montrent à la fois qu'il est de plus en plus aisé de réaliser de nouveaux contrôleurs originaux et qu'il y a une réelle motivation pour développer une réflexion approfondie sur les enjeux et conséquences de ces travaux en informatique musicale. Les problématiques abordées concernent notamment :

- la conception et la réalisation de nouvelles interfaces de contrôle gestuel [14][15][16] [17][18],
- la mise en correspondance entre données gestuelles et données sonores (mapping) et le contrôle de la synthèse sonore [19][20][21][22][23][24][25][26],
- l'étude de critères de comparaison entre différentes interfaces de contrôle [27][28],
- l'influence des gestes musicaux (notamment de production sonore) sur les caractéristiques timbrales du son [29][30][31][32][33],
- l'influence des gestes musicaux sur la perception visuelle et auditive [34][35][36],
- l'analyse et la synthèse de mouvements expressifs [37][38],
- l'étude des particularités motrices des gestes musicaux [39][40][41],
- l'étude du mouvements d'auditeurs écoutant de la musique [42].

Cette liste n'a pas vocation à être exhaustive mais cherche plutôt à montrer l'étendue du spectre d'investigations menées autour de l'étude du corps et du mouvement dans la performance musicale. Ainsi, les problématiques traitant du geste musical s'inscrivent dans différentes disciplines telles que l'acoustique, le traitement du signal et la biomécanique, domaines sur lesquels reposent les travaux présentés dans cette partie.

1.3. Acquisition des données

Sans rentrer dans le détail d'une énumération des systèmes de captation des gestes (on trouvera de telles études dans [8] et dans [43]), on distingue deux modes d'acquisition :

• Acquisition directe : utilisation de différents types de capteurs qui mesurent l'ensemble des mouvements impliqués dans un même geste. On mesure alors le plus souvent des grandeurs physiques comme des forces, des déplacements, des accélérations, etc. Le plus souvent, chaque type de grandeur nécessite l'utilisation d'un type de capteur particulier.

• Acquisition indirecte : le geste est déduit dans ce cas du son produit par l'instrument. Le capteur utilisé est alors un microphone (capteur de gradient de pression) placé devant l'instrument. Les paramètres gestuels sont extraits du son par des techniques d'analyse et de traitement de signal temps-réel, ce qui nécessite une grande puissance de calcul. Cette approche a été, par exemple, appliquée à la guitare [44].

Le système considéré dans notre étude fonctionne par l'acquisition directe des paramètres gestuels. Bien que nous souhaitions définir un modèle qui puisse s'adapter à tout type de contrôleur, ce système permet un déploiement plus simple et une familiarisation avec l'équipement plus rapide.

2. Mimophonie

Depuis une trentaine d'années les musiciens regroupés dans l'association E Voce di U Cumune s'attellent à une activité soutenue de recherches et d'études sur la musique corse. Fondée en 1978 Voce Cumune a son siège à Pigna et continue aujourd'hui encore d'œuvrer pour l'apprentissage, l'étude, la conservation du patrimoine musical corse. Si lors de sa création les activités avaient pour objectif la renaissance des instruments traditionnels (Cialamella, Pìfana, Pìrula et surtout la Cetera) elles s'orientèrent rapidement vers la tradition orale du chant et tout particulièrement le répertoire polyphonique. Donc, en même temps que la facture d'instruments, a débuté l'étude, la conservation voire la « restauration » du patrimoine vocal sacré et profane. Un profond travail d'écoute et d'analyse des fonds sonores, ainsi que des chantres détenteurs du patrimoine musical a permis de cerner la grammaire musicale essentiellement modale qui régit ces chants.

Dans *Les musiques de Corse,* l'auteur nous explique que l'architecture des intervalles est construite « en référence à la tonique conclusive » [45]. Les modes de *ré* et de *do* prépondérants sont souvent interprétés dans d'autres tonalités avec les altérations indispensables au respect du mode. Partant de ce constat, l'association se consacra à l'identification des intervalles de l'échelle musicale corse.

Nando Acquaviva, membre de l'association E Voce di U Cumune, découvrit rapidement la présence de micro-intervalles particuliers. À chaque degré identifié et classé à partir de la tonique fut attribué un chiffre arabe, les micro-intervalles étant symbolisés par un chiffre et son écriture spéculaire. Ces derniers furent l'objet d'une écoute particulière notamment pour celui du troisième degré, situé à une hauteur médiane entre majeur et mineur et logiquement baptisé « *terza mezana* ». Les autres micro-intervalles repérés, furent également désignés en langue corse, selon leur position sur l'échelle, enrichie ainsi de six hauteurs intermédiaires.

En conclusion, l'échelle musicale est donc constituée des douze demi-tons communs à la musique tempérée, auxquels s'ajoutent six micro-intervalles présents dans la musique corse. On pourra découvrir, dans la figure 2, l'étagement de ces dix-huit hauteurs.

	SIGNES	CORRESPONDANCES	NOMS	FRACTIONS	CENTS	NOTES
	8		OTTAVA SUPRANA	2/1	1200	DO
	٢		SENSIBULA	15/8	1088.27	SI
	٦	*Safiud-din Chromatique adouci Aristoxène*	SETTESIMA MEZANA	11/6	1049.36	LA ¼t
	7		SETTESIMA MINORE	16/9	996.09	SI♭
	6		SESTA MAIÒ	5/3	884.36	LA
	∞	*Safiud-din*	SESTA MEZANA	6561/4096	815.64	LA♭+
		Shruti Rohini		8/5	813.69	LA♭+
	∂	*Safiud-din*	SESTA MINORE	49/31	792.62	LA♭
	5		QUINTA	3/2	701.96	SOL
	4	*Safiud-din*	TRITONU	10/7	617.49	FA#
		Safiud-din Pythagoricien		729/512	611.73	FA#
		Tempéré			600	FA#
	⊦	*Safiud-din*	QUARTA CRESCIUTA	7/5	582.51	FA#--
	4		QUARTA	4/3	498.04	FA
	ε		TERZA MAIÒ	5/4	386.31	MI
	ε	*Safiud-din*	TERZA CUPERTA	256/209	351.17	RE ¼t
	∂		TERZA MEZANA	11/9	347.41	MI-
				50/41	343.56	MI-
				39/32	342.48	MI-
	3	*Chromatique amolli*	TERZA MINORE	6/5	315.64	MI♭+
I		*Pythagoricien Wasta du diatonique arabe*	DICESIMA DI U BASSU	32/27	294.14	MI♭
9	2	*Ton majeur pythagoricien*	SICONDA MAIÒ NUVESIMA DI U BASSU	9/8	203.91	RE
	Ω	*Ton mineur aristoxénien*	SICONDA CHJUCA	10/9	182.40	RE-
		Ton mineur vrai des arabes et des grecs (10 quart descendant)		65536/59049	180.45	RE-
ę	S	*Safiud-din ½ ton maj. aristoxénien*	SICONDA MINORE	16/15	111.73	RE♭
8	1		TONICA	1/1	0	DO

Figure 2 : Echelle musicale corse

Les recherches de l'association, loin d'être confinées à l'analyse musicale se prolongent par des actions de vulgarisation et de transmission des connaissances, au moyen de publications, de stages de formation, de créations musicales. Une démarche pédagogique et un outil didactique ont été conçus pour favoriser l'apprentissage de la polyphonie puis des actions menées en milieu scolaire.

Afin de transmettre la spécificité du chant traditionnel aux nouvelles générations, l'association à mis au point une méthode d'enseignement, avec un support mnémotechnique gestuel et chiffré. Cette méthode accessible à tous ne recourt pas à l'écriture musicale classique. Elle s'attache, avant tout, à structurer chez l'apprenant la conscience des particularités de la musique corse, pour préparer la maîtrise de l'échelle musicale traditionnelle [46] puis la pratique polyphonique.

Nous remarquons alors les deux aspects de la minophonie. D'une part la notation, qui permet l'écriture de partitions en utilisant les micro-intervalles. D'autre part, la gestuelle qui permet l'apprentissage des chants, où un chanteur peut en guider deux autres.

2.1. La notation

A l'observation, le chant traditionnel révèle une logique plus modale que tonale. La hauteur relative des sons y est déterminée en référence à la tonique. Les modes de **ré** (souvent articulé avec **do**) et de **sol** étant chantés dans d'autres tonalités avec les altérations nécessaires, la structure du mode n'est en rien modifiée, le chant est simplement interprété plus ou moins haut [47].

Les degrés de l'échelle musicale identifiés à partir de la tonique baptisée *Tònica*, chaque degré se voit ensuite attribuer un chiffre décrivant sa position :

1 pour la tonique : *tònica,* **2** pour le deuxième degré appelé : *seconda*, etc. Pour les altérations N. Acquaviva eut recours à une écriture en miroir en conservant le même chiffre (cf. Figure 2).

La tessiture de la basse parcourt une dixième, elle est symbolisée par un dix en chiffre romain (**X**), la suivante par un **9** et sa tonique supérieur par un **8**. L'octave basse est notée **I (**un en chiffre romain). La notation de la *terza mezana* réunit le chiffre **3** et son écriture spéculaire. Idem pour la sixte et la septième. Ces chiffres sont placés sur un support, qui symbolise une même octave, pour établir une portée qui permet une notation et une lecture synthétiques. Cette codification est utilisée à des fins de mémorisation, d'analyse et d'apprentissage.

2.2. La gestuelle

Nando Acquaviva codifia une gestuelle, pour guider à partir de la voix principale (*siconda*), les deux autres voix de la polyphonie corse, c'est-à-dire la basse : *bassu* et la tierce : *terza* plus aiguë que les autres. L'ambitus dans les polyphonies corses étant assez restreint il est possible d'indiquer tous les degrés avec les doigts de la main.

La gestuelle établie par Nando Acquaviva est fondée sur la représentation ordinaire des chiffres. Les signes dépassant le chiffre 5 et les altérations sont symbolisés par des positions plus particulières des doigts, voire de la main (cf. Figure 3).

Pour l'interprétation, quelques signes parfois empruntés à l'écriture conventionnelle viennent compléter le code mimophonique. Baptisés *i fiori*, ils indiquent l'ornementation (mélismes), la dynamique, le rythme, les pauses.

Cette méthode fut expérimentée lors de stages de formation organisés à Pigna, mais également auprès du public scolaire des écoles de la région de Balagne pour lesquelles Nicole Casalonga (membre de l'association E Voce di U Cumune) conçut et publia une progression pédagogique[3].

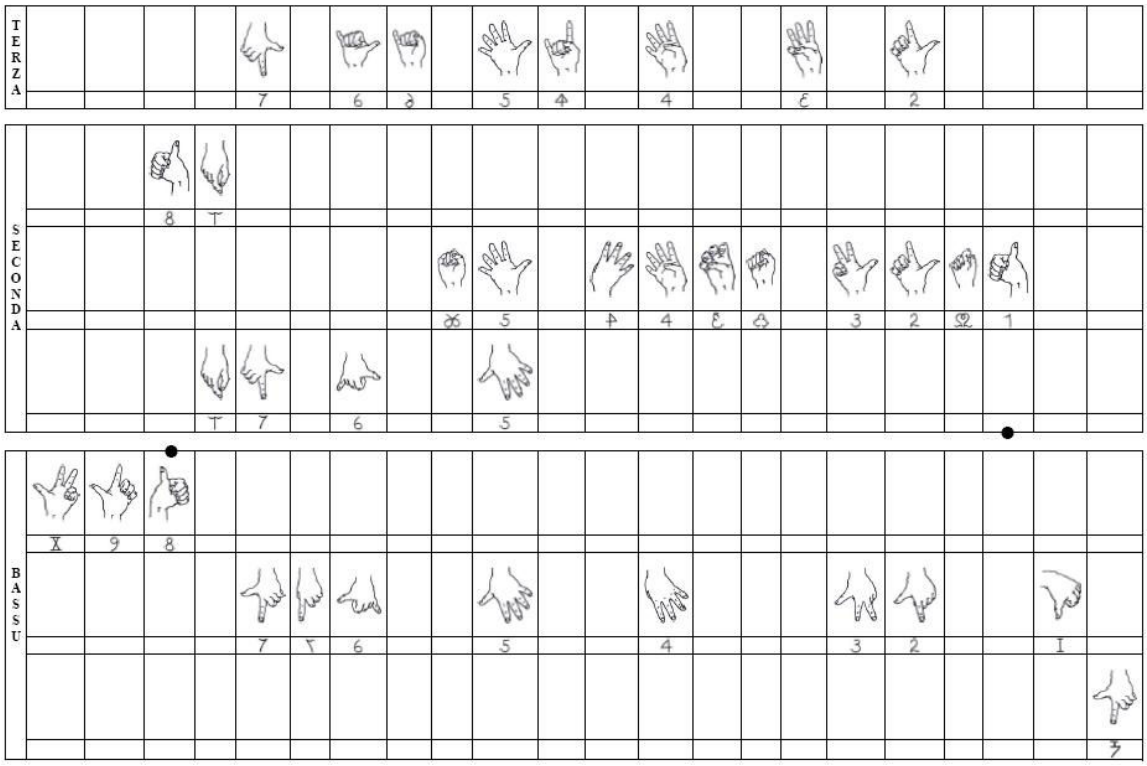


Figure 3 : correspondance geste-signe

Nous allons, pour résumer, prendre comme exemple l'intervalle de quarte. Celui-ci se situe entre la *terza maiò* et la *quarta cresciuta.* Sur la Figure 2, nous voyons que cette note se situe 498,04 cents plus haute que la tonique. Nous aurons donc un FA si la tonique est un DO, un DO si l'on est en SOL, etc. Dans l'échelle musicale telle que l'a représenté N. Acquaviva, cet intervalle est représenté, de manière écrite par un 4 et de manière gestuelle par la position de la main suivante : la main vers le haut, la paume en avant, le pouce plié et les autres doigts tendus.

[3] Nicole Casalonga, *Petit guide en 12 exercices pour découvrir le langage musical traditionnel, Quaderni Casa Musicale I,* Voce di u Cumune, Pigna.

Grâce à la mimophonie, chaque intervalle est caractérisé par un nom, une écriture chiffrée, une note et un geste. Il est nécessaire de pouvoir définir un dispositif en vue de reconnaître ces gestes.

3. Captation des gestes

Les chants polyphoniques corses sont traditionnellement constitués de trois voix : une tonique, une basse et une voix de tierce. Cette tradition du chant s'est transmise oralement à travers les générations et ne possède pas de trace écrite. C'est pourquoi il est utile de créer un outil d'apprentissage utilisant les micro-intervalles présentés précédemment.

Afin de pouvoir utiliser le langage mimophonique dans un logiciel d'apprentissage de chants polyphoniques, nous devons résoudre le problème de la captation des gestes. L'étape de captation permet de numériser les gestes de la mimophonie en utilisant une interface gestuelle.

3.1. Les interfaces gestuelles

Les interfaces gestuelles se proposent d'améliorer l'interaction homme-machine en utilisant la reconnaissance des gestes de l'utilisateur pour entrer des commandes. Selon Baudel. T. et Braffort, A. [48] on distingue deux types d'interfaces gestuelles :

- soit le système analyse la trace laissée par le geste (par exemple dans le cas de la reconnaissance de l'écriture ou de signes tracés au moyen d'une souris ou d'une tablette), auquel cas on parlera d'interfaces gestuelles *instrumentales*;
- soit il interprète le geste lui même, par exemple au moyen d'un gant numérique ou d'une caméra vidéo, auquel cas nous parlerons d'interfaces gestuelles *pures*. Avec une interface gestuelle pure, la machine n'analyse pas les modifications induites sur un périphérique, mais capte le mouvement même de la main : c'est celle-ci qui devient le périphérique.

On distingue trois types d'interfaces gestuelles pures:

• *Interfaces des réalités virtuelles*, basées sur la manipulation directe d'objets représentés en trois dimensions. L'utilisateur est plongé dans un monde virtuel, fonctionnant comme le

monde réel, mais dans lequel l'ordinateur interprète toutes ses actions comme des commandes. Krueger [49] décrit en détails ce type d'interfaces.

• *Interfaces Multi-Modales* : Etymologiquement, ces interfaces visent à interpréter en parallèle plusieurs des canaux de communication humains (voix, geste, direction du regard), pour améliorer l'interaction. La plupart des expériences menées dans ce domaine ont pour objectif de rapprocher l'utilisation d'une machine d'une communication naturelle entre humains [50] [51] [52].

• *Langages gestuels opératifs* : Ces interfaces utilisent la reconnaissance de langages gestuels spécialisés. K. Murakami et H. Taguchi [53] reconnaissent des mots de la langue des signes japonaise (langue utilisée par les sourds-muets). D. Sturman [54]reconnait un vocabulaire utilisé pour orienter les grues sur des chantiers. Morita et al. [55] utilisent un gant numérique et une baguette pour permettre à un chef d'orchestre de diriger un orchestre synthétique.

3.2. Les systèmes de captation

En général, un instrument virtuel est constitué d'un contrôleur de gestes qui convertit les mouvements de l'instrumentaliste en signal électrique et une interface qui génère du son en fonction du signal [56]. Les contrôleurs peuvent être conçus selon deux approches : l'une qui reproduit des interfaces traditionnelles ou qui s'en inspire, l'autre qui privilégie la création de nouvelles. Le principal intérêt de la première démarche est qu'elle conserve le bénéfice d'un geste *expert*, forgé par la culture et l'histoire. C'est le cas des contrôleurs Midi reproduisant les interfaces du saxophone, de la guitare, du violon ou du piano. Le même geste, appris et cultivé, peut ainsi contrôler des événements de natures très diverses. La seconde démarche ouvre au contraire l'accès à un univers gestuel inexploré, précisément en raison des contraintes mécaniques des instruments traditionnels. Afin de pouvoir appliquer ces nouveaux gestes aux instruments, il faut les capter [8].

C'est pourquoi de nombreux instruments virtuels comprennent des capteurs de pression, de position, de rotation et de translation, avec ou sans contact physique. Citons entre-autre le système Hands développé par l'institut Steim, les systèmes de capture vidéo, tels le BigEye, de Steim, ou l'Irisis, de Collectif et Cie [1] qui repèrent la position des objets au moyen d'une caméra et déclenchent des séquences qui ne sont pas nécessairement sonores (ils sont souvent utilisés avec des danseurs). Enfin, les gants de données, peuvent traquer les moindres mouvements de la main et des doigts pour les convertir en signaux numériques.

L'outil de captation des gestes que nous devons créer, sera destiné à un public non expert dans le domaine de l'informatique. Celui-ci doit donc être facile à mettre en place et à utiliser.

Il existe plusieurs systèmes d'acquisition de mouvements, notamment pour les mouvements de musiciens. L'exemple le plus représentatif de ce type de captation est le système VICON. Il repose sur le suivi (*tracking*) de marqueurs par un réseau de caméras rapides à haute résolution. Les marqueurs sont des réflecteurs passifs, non-filaires, et sont en conséquence peu invasifs. En plaçant des réflecteurs sur un sujet, dans le cas présent sur chaque doigt et la main, il est possible de mesurer des déplacements avec une précision spatiale en dessous de 1mm, à une fréquence pouvant atteindre 500Hz dans un volume de l'ordre de $1m^3$. Les marqueurs utilisés sont généralement sphériques ou demi-sphériques et leur taille est variable selon l'utilisation envisagée : de 5mm à plusieurs centimètres de diamètre. Toutefois, ce système présente quelques inconvénients. Malgré sa souplesse d'utilisation, le réseau de caméra est relativement encombrant : il est contraignant à déplacer et le système reste généralement attaché à une salle dédiée. De plus, de par la technologie employée, ce type de système de captation est relativement onéreux. Cependant, malgré ces inconvénients, ce système est l'outil idéal pour effectuer des mesures précises tout en minimisant la gêne pour l'utilisateur. Ce système a précédemment été employé sur des violoncellistes par [57].

Le système de capture vidéo ne paraît donc pas approprié.

Les gants de données sont une solution plus adaptée. Ils existent depuis de nombreuses années et ont été développés pour la réalité virtuelle. Ils permettent de repérer la position de la main dans l'espace ainsi que la flexion des doigts, ce qui permet en général de piloter un avatar d'une main virtuelle dans un programme infographique de réalité virtuelle.

Souvent relié par un fil à une interface, ces gants professionnels sortent cependant du budget à consacrer pour le développement de notre projet. De plus, l'interface logicielle ou le protocole sont rarement adaptés à un environnement comme MAX/MSP, prévu comme base de développement du moteur vocal des polyphonies virtuelles.

A l'opposé, nous avons écarté un contrôleur comme le gant P5. Extrêmement bon marché et facilement interfaçable avec MAX/MSP, ce contrôleur, prévu pour les jeux vidéos, manque de robustesse. De plus, il est impossible d'en utiliser deux simultanément à proximité l'un de l'autre.

Aussi, nous avons décidé, avec l'équipe temps réel de l'IRCAM, de réaliser notre propre système de gant de données avec une liaison sans fil assurant la transmission rapide à l'ordinateur tout en conservant une grande liberté de mouvement.

Nous avons dû créer deux gants (un pour chaque main) composés chacun de six capteurs : cinq capteurs de flexion, pour les doigts et un accéléromètre qui permet de relever les déplacements de la main (figure 4).

Nous utilisons pour la connexion sans fil, un dispositif d'émission-réception Bluetooth. Pour chacun de ces capteurs, nous sommes en mesure, grâce à une interface MAX/MSP, de générer une valeur. Nous pouvons donc attribuer à chaque geste du langage mimophonique six valeurs, une pour chaque capteur de flexion et une pour l'accéléromètre.

3.3. Réalisation des gants

Ce développement électronique a pour but de doter l'utilisateur d'un contrôleur bi-manuel permettant entre autre l'utilisation du langage mimophonique contrôlé par le geste et les postures de main.

Rappelons que la base des polyphonies corses requiert trois chanteurs. Dans l'apprentissage par la mimophonie, un chanteur dirige les deux autres. Basé sur un apprentissage de la tradition orale et de la pratique maître-élève, ce type de chants restait difficile sur le plan pédagogique, surtout qu'aucune notation/partition n'existait. Le chanteur dirigeant indique aux deux autres comment se placer par rapport à sa propre voix, en indiquant de la main à quel intervalle se placer.

Transposé dans le domaine numérique, il est apparu intéressant de doter un utilisateur d'un contrôleur lui permettant de piloter en temps réel un modèle audio-numérique de ces polyphonies. Notre contrôleur doit donc être capable de repérer la flexion de chaque doigt, ainsi que l'orientation de la main (vers le haut ou vers le bas).

3.3.1. Contraintes et design électronique

Les gants de données sont un support fiable et très utilisés dans ce domaine. Ils permettent de contrôler à la fois la position de la main dans l'espace et la flexion des doigts.

Nous devons donc utiliser un système de reconnaissance utilisant des gants de données. Ces derniers doivent être facilement interfaçables avec MAX/MSP, robustes et bon marché. De plus, il faut pouvoir utiliser les deux gants simultanément et proches l'un de l'autre.

Aussi, nous avons développé nos propres gants de données. Chacun est équipé de capteurs et d'un système de liaison sans fil qui permet d'effectuer aisément les gestes de la mimophonie et de capter ces derniers en temps réel.

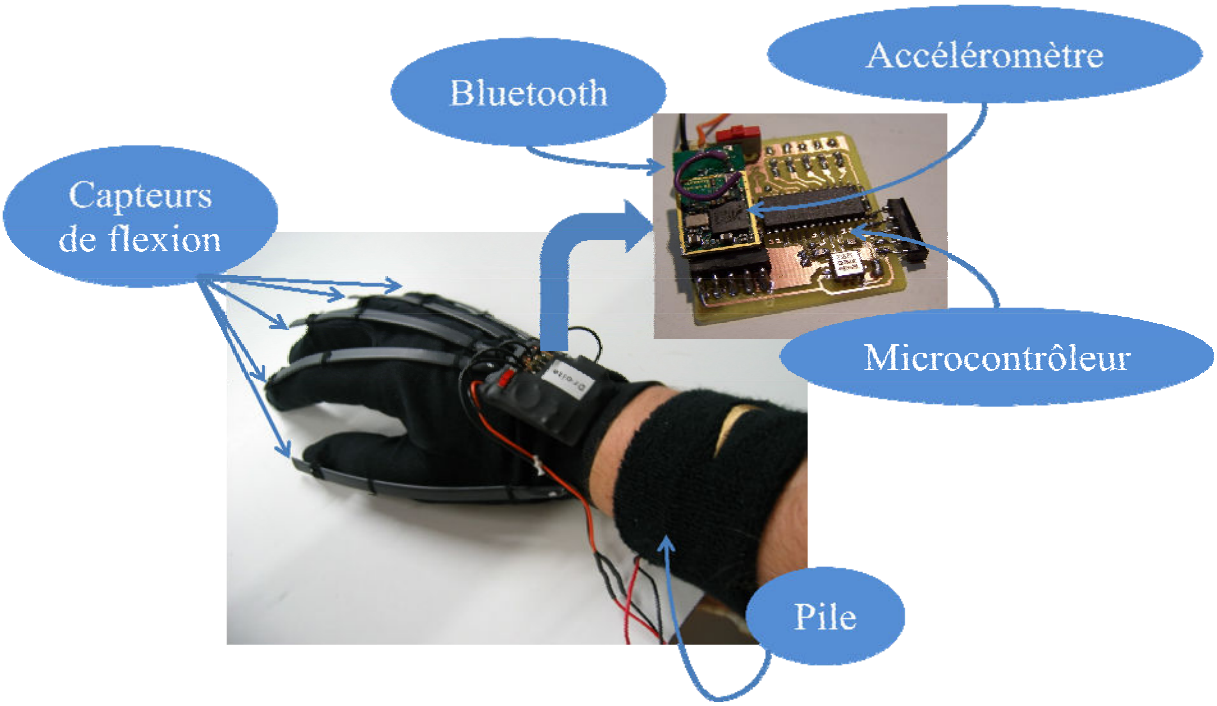


Figure 4 : Un des deux gants de données

3.3.2. Support

Pour fixer les capteurs, nous avons tout d'abord cherché des gants en lycra, souples et élastiques, comme ceux utilisés en réalité virtuelle. Impossible à trouver facilement, nous nous sommes rabattus sur des sous-gants en soie. En dépit des coutures (fines), le gant reste souple et agréable. La soie, plus résistante que le Lycra, a facilité la couture d'accrochage des capteurs.

3.3.3. Flexion des doigts

Nous utilisons des capteurs de flexion désormais classiques de chez JamecoElectronics[4]. Ces capteurs comportent une encre résistive entrant en contact avec un plastique métallisé. Lorsque le capteur est fléchi, la surface de contact entre l'encre et l'électrode augmente, faisant baisser la résistance électrique de l'ensemble. Ce capteur se comporte donc comme une résistance variable en fonction de la flexion.

Sur une échelle de 0 à 90° le capteur passe d'environ 1 MΩ à 10 kΩ.

Pour en extraire une tension électrique, il suffit de le monter en diviseur de tension avec une résistance talon « pull-down » d'environ 10 kΩ.

[4] http://www.jameco.com/

3.3.4. Orientation de la main

Nous aurions pu utiliser un capteur type « à goutte de mercure » pour déterminer si la main est orientée vers le haut vers le bas, cependant, ce capteur tout-ou-rien ne permet pas de détecter les transitions entre postures. Nous avons opté pour un capteur délivrant de l'information gestuelle en continu.

Le capteur est un accéléromètre ± 2g ADXL202 de chez AnalogDevices[5]. Ce capteur permet de mesurer l'accélération « statique » liée à la pesanteur ainsi que l'accélération dynamique d'un mouvement. Ainsi, on peut mesurer l'inclinaison du capteur ou l'accélération dans un plan parallèle au sol, mais on ne peut pas dissocier les deux informations lors d'un mouvement libre. Cependant dans notre cas, les mouvements sont très cadrés et quasi exclusivement dans le plan vertical du torse ce qui permet une lecture de l'orientation de la main dans ce plan, ainsi que l'information main en mouvement ou main statique.

De plus, l'accéléromètre comporte deux sorties, une pour chaque axe dans le plan de la détection. L'analyse combinée des deux axes permet de lever certaines ambiguïtés sur le geste.

L'accéléromètre fournit pour chaque axe une tension polarisée à la moitié de la tension d'alimentation lorsque le capteur est au repos et à 0° (à plat par rapport au sol). Une accélération positive provoque une élévation de tension, et réciproquement pour une accélération négative.

3.3.5. Numérisation des capteurs

Chaque gant est doté d'un petit circuit imprimé comportant un microcontrôleur de type PIC 18F2220 de chez Microchip[6]. Son horloge utilise un oscillateur interne ce qui permet de s'affranchir d'un encombrant quartz. Le microcontrôleur dispose d'entrées analogiques multiplexées pouvant être numérisées à l'aide du convertisseur analogique / numérique (CAN) 10 bits interne. Ainsi, à l'aide de ce seul composant, nous pouvons faire l'acquisition des 7 signaux des capteurs (5 capteurs de flexion et 2 axes sur l'accéléromètre) avec un très faible encombrement.

Les valeurs numériques image des capteurs sont ensuite envoyées sur le port série du microcontrôleur sous forme de trame pour être transmises jusqu'à l'ordinateur par une

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

[6] http://www.microchip.com/wwwproducts/Devices.aspx?dDocName=en010266

technologie sans fil. La valeur de chaque capteur est scindée en deux octets, l'ensemble des 10 bits de la valeur numérisée ne tenant pas sur un octet seul.

Afin de repérer sans erreur la valeur de chaque capteur, le paquet envoyé par le port série comporte un token de début trame valant 255 (0xFF). Suivent alors l'octet de poids faible et de poids fort (dans cet ordre) de chaque capteur. Du fait de l'utilisation du token de début de trame, le 8ème bit d'un octet ne peut être utilisé pour les données numérisées, sous peine de confondre par exemple le poids faible d'un capteur avec le token de début de trame. Aussi, le 8ème bit est réservé : il sera à 1 pour le token de début de trame, et à 0 pour les octets des valeurs numérisées (Figure 5).

C'est donc sans surprise que nous utilisons un format ressemblant à la norme MIDI pour les octets de nos capteurs. Les 10 bits sont scindés dans un format 7+3, c'est-à-dire 7 bits de poids faible (*Least Significant Bits* ou LSB) et 3 bits de poids fort (*Most Significant Bits* ou MSB). La valeur sera reconstituée lors de la réception sur l'ordinateur, dans Max/MSP.

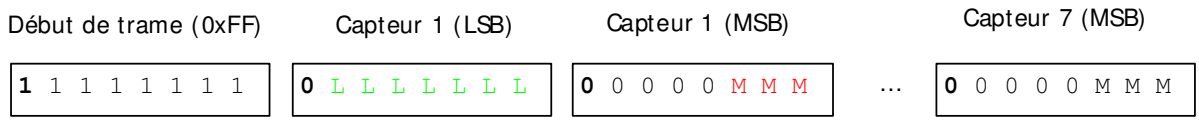


Figure 5 : Trame des capteurs numérisée et envoyée sur le port série

3.3.6. Système de transmission sans fil

Transmettre une liaison numérique sans fil n'est pas d'une grande difficulté technique. Cependant, en transmettre deux dans le même espace devient tout de suite plus compliqué. En effet, la législation sur les radiofréquences interdit les émissions libres. Des canaux spécifiques sont réservés à cette utilisation et c'est pourquoi on ne trouve dans le commerce qu'un nombre limité de transmetteur sans fil radio. Pour la plupart d'ailleurs, ils émettent sur un canal radio fixe, ce qui empêche d'en utiliser plusieurs simultanément.

Des variantes existent, on peut trouver par exemple un émetteur pour la gamme 433 MHz, et un autre pour la gamme 868 MHz, ce qui aurait permis de résoudre notre problème de gants de données sans fils, cependant, cette solution n'est pas pérenne et n'autoriserait pas de gérer deux personnes équipées du même système.

Nous avons donc recherché une solution plus générique. Nous avons opté pour des émetteurs/récepteurs Bluetooth classe 2 (20 m de portée). Ces émetteurs utilisent la bande ISM 2.4 GHz, partagée avec entre autre la technologie de réseau sans fil 802.11 (WIFI).

Comme la plupart des protocoles réseaux sans fil récents, Bluetooth permet un partage efficace des canaux par les techniques d'évitement de collision et d'écoute avant transmission (*ListenBefore Transmit*). Ceci permet à plusieurs périphériques de partager le même canal radio en vérifiant si celui-ci est libre avant de transmettre. L'adjonction d'un délai aléatoire assure que statistiquement tous les périphériques auront un accès au média pour transmettre.

Nos modules Bluetooth proviennent de la société suédoise « Free2Move[7] ». Le module émetteur mesure environ 20 x 10 mm ce qui le rend particulièrement bien adapté pour le gant. Chaque module est doté d'un émetteur et d'un récepteur ce qui permet d'expédier des données capteurs à l'ordinateur mais ce dernier pourrait également au besoin envoyer des ordres à nos gants de données.

Du fait que de nombreux ordinateurs sont équipés de port Bluetooth en standard (surtout les ordinateurs portables) ou qu'il est extrêmement bon marché de l'en équiper via un petit périphérique USB. Nous avions initialement envisagé de recevoir les données des gants directement sur l'ordinateur via le port Bluetooth interne. L'émetteur Free2Move se comporte en effet comme un port série virtuel Bluetooth, ce qui permet de l'interfacer rapidement avec un objet serial dans Max/MSP.

Cette solution, élégante et attrayante est malheureusement peu viable au final. En effet, tout port Bluetooth géré sur un ordinateur passe par la pile Bluetooth, un fragment logiciel gérant la découverte et l'appariement (*pairing*) avec les périphériques Bluetooth (port série, oreillette, téléphone portable, souris. Le processus de découverte et d'appariement est fastidieux et l'arrêt subi d'un périphérique apparié résulte d'un périphérique « fantôme ». Le périphérique est toujours vu par l'ordinateur, mais la remise en route ne permet pas de rétablir la communication. Ce n'est qu'au bout d'un certain temps (*time out*) que l'ordinateur « relâche » le périphérique et autorise à nouveau sa connexion (ce temps se mesure en général en minutes).

Après de nombreuses recherches infructueuses, il a fallu nous rendre à l'évidence : il est impossible d'agir sur le *time out* de déconnexion des périphériques.

Pour contourner ce problème, nous avons utilisé des modules Bluetooth autonomes en guise de récepteur. Ceux-ci sont totalement externes à l'ordinateur et donc exempts de pile Bluetooth. Grâce au logiciel fourni par Free2Move, il est possible de configurer des paires de périphériques à l'association quasi immédiate et sans délais (< 1 seconde). De la même

[7] http://www.free2move.se

manière, il est possible de régler le *time out* sur un temps très court, ce qui autorise la reconnexion des gants quasi immédiatement après une mise hors tension.



Les modules autonomes ne sont rien de plus que la version classe 1 (portée 100m et sensibilité accrue) des modules placés sur les gants. Ces derniers comportent une sortie série RS-232 classique, facile à interfacer à un port série USB. Cependant nous les avons modifiés afin de prélever les signaux reçus et à les faire d'abord passer par une carte à microcontrôleur chargée de former des paquets à timing constant. En effet, à travers Bluetooth, les données « série » circulent bien plus vite qu'une liaison série classique. Il convient alors de les réceptionner pour reformer des paquets uniformes (en détectant le token de début de trame), puis de les envoyer en un bloc, cette fois vers un port série USB.

L'ensemble de ces modules et cartes est logé dans un boîtier en plastique équipé d'un concentrateur USB (hub) qui fournit également l'alimentation 5V à l'ensemble du système (Figure 6).

Ainsi, un unique cordon USB rejoint l'ordinateur, sans besoin d'alimentation additionnelle.

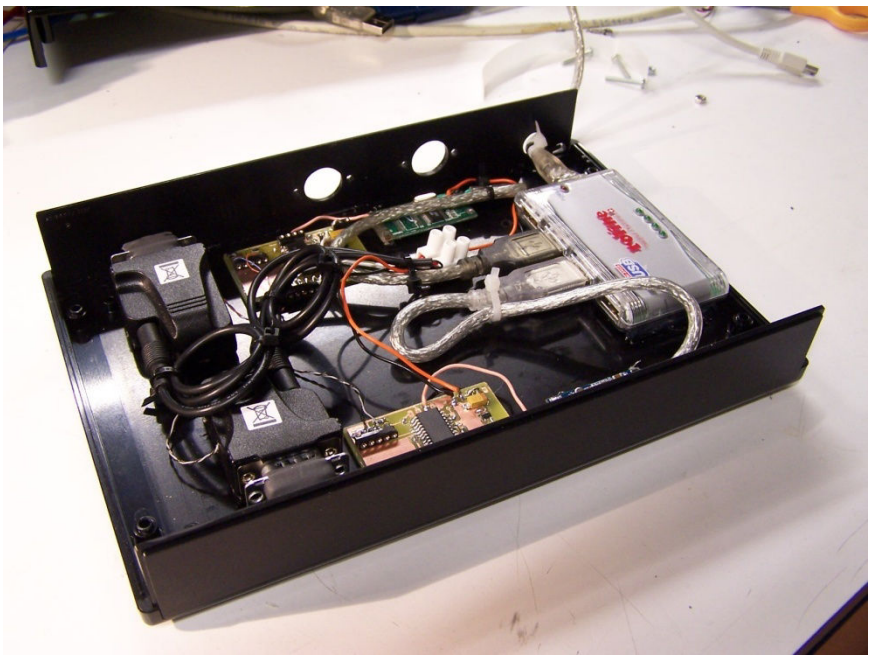


Figure 6 : Boitier de réception

4. Conclusion

Le chant polyphonique se fait à trois voix. L'objectif de nos travaux est de pouvoir soustraire deux chanteurs et de créer une polyphonie grâce à des gants de données. Chaque gant, équipé d'un ensemble de six émetteurs, permet de capter tous les gestes de la mimophonie. Ces gants sans fil vont être ensuite reliés à un ordinateur afin d'attribuer à chaque geste un ensemble de valeurs numériques en temps réel. Cette acquisition se fait à l'aide d'un module développé sous MAX/MSP et sera détaillé dans le chapitre 5.

A ce stade du projet, nous avons défini le langage gestuel qui convient à l'apprentissage de chants traditionnels corses, ainsi que le moyen de capter l'ensemble de ces gestes. Dans la section suivante nous allons formaliser la correspondance entre ces valeurs numériques et les gestes de la mimophonie.

CHAPITRE III :

Reconnaissance de gestes

Nous avons décrit précédemment (Cf. 2.2) le langage mimophonique. Nous avons vu qu'à chacun des intervalles de voix est associé un signe et à chaque signe un geste. Ces gestes permettent au chanteur principal (celui qui chante la tonique) de guider le chanteur de basse et celui de tierce. Dans le but d'élaborer une solution logicielle pour l'apprentissage et la valorisation des chants polyphoniques, nous avons créé deux gants de données (Cf. 2.4). Ceux-ci permettent de capter les gestes.

Il reste cependant à résoudre la question de la reconnaissance de ces gestes. En effet, l'interface MAX/MSP permet d'obtenir des valeurs pour chaque capteur. Celles-ci varient en temps réel en fonction de la flexion des doigts et des mouvements de la main. Cependant, pour un geste donné, les valeurs peuvent varier selon la taille de la main de la personne qui porte les gants, la flexion des doigts, la position des articulations du doigt par rapport au capteur, etc. Il nous fallait donc insérer à notre application un système permettant la reconnaissance de ces gestes.

Parmi plusieurs solutions possibles, deux nous ont semblé plus intéressantes : les systèmes d'inférence floue (SIF) et la conception d'un système de reconnaissance directement dans MAX/MSP, en utilisant les bibliothèques développées par l'équipe Interaction Musicale Temps Réel (IMTR) de l'IRCAM.

Nous présenterons tout d'abord dans ce chapitre un état de l'art sur la reconnaissance gestuelle dans le domaine musical. Nous décrirons ensuite les deux systèmes de reconnaissance cités précédemment. Puis nous terminerons par une discussion sur ces différentes techniques.

1. Etat de l'art

Avant de poursuivre ce chapitre sur la reconnaissance de gestes, il est nécessaire de donner la définition du *geste*. En effet, tout comme dans le domaine du contrôle gestuel (Cf. 2.1.a), le terme de « geste » peut avoir plusieurs définitions suivant le contexte. Dans le cas des gestes de la mimophonie, nous nous intéressons surtout à la posture du geste. C'est-à-dire à un geste figé dans le temps. Le geste et la posture ont respectivement une forme dynamique et statique (tout comme si l'on regarde un film image par image). Selon [58], "La coordination entre la posture et le mouvement conditionne l'efficacité du geste. En fait, les activités musculaires posturales, statiques ou dynamiques, permettent d'anticiper, accompagner et suivre l'exécution du mouvement d'une manière automatique" et plus tard

dans la même page: "[…] Ce qui est perçu comme une simple forme statique est le résultat d'un processus dynamique".

Notre objectif est d'aboutir à un système de reconnaissance capable de différencier les différents gestes de la mimophonie définis par un ensemble de signes de la main. Ce système doit permettre de reconnaître des postures au sein de mouvements constants. En effet, le chanteur qui utilise le langage mimophonique, pour guider les autres chanteurs, enchaine une série de gestes au rythme de son chant. Les transitions entre deux gestes ne devront pas être reconnues. Il s'agit donc d'identifier un évènement discret dans un ensemble de mouvements continus (ou de signaux spatio-temporels [59]).

L'analyse de mouvements (ou signaux spatio-temporels) passe par des techniques basées sur les principales approches suivantes : la mise en correspondance dynamique (DTW - Dynamic Time Warping) [60] [61], les approches neuronales [62] [63] et les Modèles de Markov Cachés (MMC - Hidden Markov Models, HMM) [64] [65] [66] [67] [68] [69] [70] [71].

Le DTW a été appliqué à la reconnaissance de gestes après avoir fait ses preuves en matière de reconnaissance vocale dynamique. C'est une technique de mise en correspondance dynamique basée sur des modèles connus. Le but est d'aligner le modèle et le signal à reconnaître par déformation non-linéaire de telle sorte que la différence entre les deux signaux devienne minimale [60] [61]. Cette méthode est efficace lorsque le nombre de phénomènes à reconnaître est faible (les variations temporelles en rythme et durée sont prises en compte). Cependant, elle nécessite un grand nombre de modèles lorsque l'on souhaite élargir le champ de reconnaissance. De plus, elle ne reconnaît pas les modèles non définis. Enfin, cette technique pêche par sa lourdeur, même si des améliorations ont été apportées pour accélérer les calculs [61].

Les approches neuronales sont particulièrement adaptées à la classification de modèles dans les ensembles de données [63] [62]. Le principal problème de ces techniques reste la prise en compte du temps. Dans une approche, le temps est vu comme un mécanisme externe. Dans ce cas, le TDNN (Time Delay Neural Network – réseaux de neurones à lignes de retard) est l'un des modèles les plus efficaces. Inspiré du réseau de neurones multi-couches, le TDNN transforme le problème temporel en un problème spatial. Il a montré son efficacité en reconnaissance de gestes en langage des signes [63], mais les systèmes basés sur ce principe ou leurs variantes restent d'une trop grande complexité et demandent une étape de pré-

traitement des données trop lourde. Dans une autre approche, le temps est intégré en tant que mécanisme interne au système [72]. Dans tous les cas, le principal inconvénient reste le coût de calcul prohibitif et la complexité de l'apprentissage ainsi que la difficile interprétation des résultats.

Les modèles de Markov cachés sont aussi très appliqués dans l'analyse de séries temporelles contenant des variations spatiales et temporelles [68]. Ils ont rencontré beaucoup de succès face aux problèmes de reconnaissance de l'écriture, de la voix et, plus récemment, des gestes [73] [64] [66] [65] [67] [71] en particulier les gestes de la langue des signes [69] [70]. Particulièrement adaptés aux signaux spatio-temporels [68], les modèles de Markov cachés semblent fournir le meilleur compromis en termes de flexibilité, de rapidité et de robustesse aussi bien vis à vis du bruit que des variations temporelles des mouvements. De plus, ils sont moins opaques et moins difficilement interprétables que les réseaux de neurones.

Un autre système de reconnaissance, basé sur la logique floue, nous a semblé intéressant : les Systèmes d'Inférence Floue (SIF) [74]. Les SIF reposent sur des règles de type « SI…ALORS… ». Ces règles sont formulées avec des termes linguistiques. Elles sont donc proches du langage naturel et d'une grande lisibilité.

Parmi les différentes solutions concernant les systèmes de reconnaissance, plusieurs nous ont semblé pertinentes : l'utilisation de FTM (qui permet l'utilisation d'outils directement dans MAX/MSP), les réseaux de neurones (RdN) et les systèmes d'inférence floue (SIF).

Même si l'utilisation de la bibliothèque FTM semble être la plus rapide à mettre en place, nous avons voulu travailler sur les deux dernières méthodes de reconnaissance. Notre choix c'est porté sur l'utilisation des SIF et cela pour plusieurs raisons.

Dans les RdN, la représentation des connaissances se fait de manière implicite (boîte noire) et compacte (propriété de parcimonie). Les poids du RdN partagent l'information acquise et n'ont pas de signification facilement interprétable. Ils peuvent par contre être utilisés pour traiter des problèmes de grande dimension, là où le SIF connait des problèmes d'explosion combinatoire. A l'opposé, dans les SIF, la représentation des connaissances est explicite (règles linguistiques) et extensive, chaque paramètre a une signification précise. Grâce à cette propriété la représentation des connaissances par SIF est moins compacte que par RdN. De plus, à la différence des RdN où la diversité est assez réduite, il existe plusieurs

possibilités pour le choix des fonctions d'appartenance, le choix des opérateurs et le choix des méthodes de défuzzifications [74].

Au niveau du déploiement, ces deux méthodes sont sensiblement identiques. Nous entendons par déploiement le temps et la complexité d'intégration de la méthode dans l'application, par exemple l'implémentation des algorithmes d'apprentissage (apprentissage par perceptron [75], méthode du gradient) ou des méthodes d'inférences (méthode Mamdani [76], méthode Takagi-Sugeno [77]).

Pour la phase d'apprentissage, basée sur des règles pour les SIF et par expérimentation pour les RdN, l'utilisation des SIF est plus adaptée à notre problématique. Nous connaissons en détail le système, ses entrées et ses sorties ; de ce fait l'imprécision induite via l'utilisation des gants est plus facile à contrôler et à appréhender grâce à l'utilisation d'ensembles flous. Ce premier avantage des SIF est directement lié à notre système et notre problématique, dans d'autres circonstances les réseaux de neurones auraient très bien pu être employés.

Pour conclure cette brève comparaison, les éléments qui nous ont le plus influencés sont les résultats obtenus. Après l'intégration des deux méthodes, la reconnaissance par SIF donne de meilleurs résultats. Elle est plus précise, meilleur pourcentage de reconnaissance, et surtout elle ne reconnaît pas de gestes non renseignés. Avec les réseaux de neurones, si l'on effectue un geste qui n'existe pas, le geste le plus proche est identifié et donc une note est jouée. Alors qu'avec les SIF, si aucun geste n'est reconnu, le système ne renvoie rien. Pour notre problématique, les systèmes d'inférence flous sont plus appropriés.

2. Systèmes d'inférence flous

La logique floue est une extension de la logique classique. Elle a été présentée par Zadeh [78] comme un cadre pour le raisonnement approximatif, une théorie mathématique ayant pour objet l'étude des systèmes flous.

Le raisonnement approximatif et la prise en compte de faits imprécis et incertains sont tout à fait naturels pour l'être humain. Pour raisonner sur de telles connaissances, la modélisation classique ne suffit pas ; en effet dans ce cas, les approximations sur les variables génèrent au final des erreurs relativement importantes. Selon Zadeh [79] la modélisation floue fournit des moyens approximatifs mais efficaces pour décrire le comportement des systèmes qui sont trop complexes ou trop mal définis pour admettre l'utilisation d'une analyse mathématique précise [80]. Le résultat de cette modélisation n'est plus une valeur unique

mais un intervalle de valeurs, il y a donc plus de chances qu'il soit cohérent avec les objectifs et contraintes de départ. Pour son interprétation, on peut utiliser des critères de décidabilité afin de proposer une solution en adéquation avec le domaine étudié.

L'étude de ces systèmes passe par la prise en compte d'imprécisions, d'incertitudes mais aussi par un raisonnement pertinent et performant sur le système dans sa globalité (variables d'entrées, de sorties, comportement). Les systèmes flous possèdent des variables d'entrée floues et fournissent des résultats sur des variables de sortie elles-mêmes floues. Ces notions de base sont présentées dans la première section. Le comportement d'un système flou est basé sur un ensemble de règles permettant de fournir des résultats à partir d'un comportement partiellement connu. La prise en compte du flou dans la démarche de modélisation est décrite dans les deux dernières sections.

2.1. Systèmes flous

Les systèmes flous utilisent une représentation des connaissances sous forme de règles floues, une manière normale d'expliquer des procédés de décision. Ils expriment des concepts en terme (variable) linguistique, près de la représentation humaine, par exemple, "vent rapide" au lieu de "vent supérieur à 40km/h". Il y a une correspondance entre les variables linguistiques et des valeurs numériques associées.

Une variable linguistique [78] est une variable prenant ses valeurs dans un ensemble de "mots" symboliques, définissant certaines catégories d'un ensemble de référence.

On appelle variable linguistique un triplet (V, X, T_v), tel que :

– V : une variable symbolique définie sur X ;

– X : un ensemble de référence ;

– T_v : $T_v = \{A_1, A_2, ...\}$ un ensemble fini ou dénombrable de sous ensembles flous normalisés, utilisés pour caractériser V.

Par exemple, une variable linguistique est définie par une variable symbolique V : "la position d'un doigt", par un ensemble de référence $X = R^+$, et un ensemble de sous ensembles flous (SEF) caractérisant les positions du doigt T_v = {droit, plié, fermé}.

Une proposition floue est définie à partir d'une variable linguistique (V, X, T_v) par la qualification "V est A", exemple : "le doigt est plié".

On appelle modificateur linguistique M un opérateur qui permet à partir de toute caractérisation floue A de V, de produire une nouvelle caractérisation où A appartient à $M(T_v)$, ce dernier étant l'ensemble des caractérisations obtenues en appliquant un

modificateur linguistique M. Le résultat de cette opération n'est pas forcément dans T_v. Par exemple un modificateur linguistique peut caractériser la position du doigt comme étant "peu pliée".

Une règle floue est une proposition floue de la forme "Si p alors q" utilisant une implication entre deux propositions floues quelconques p et q. Ces règles permettent de décrire le comportement d'un système flou à partir de la combinaison d'une ou plusieurs propositions floues. Par exemple, "si le pouce est plié et l'index est droit alors le geste est une seconde".

Les systèmes flous sont adaptés à la modélisation de systèmes à comportement complexe, ils peuvent rapprocher n'importe quel comportement numérique avec le niveau désiré de précision et un coût d'expression linguistique réduit [81].

2.2. Sous ensembles flous

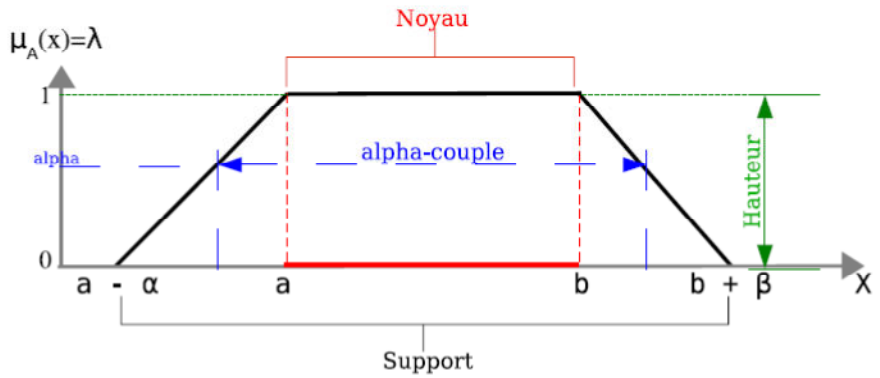


Figure 7 : Exemple de Sous Ensemble Flou

Les sous ensembles flous [82] ont été introduits afin de modéliser la représentation humaine des connaissances, et ainsi améliorer les performances des systèmes décisionnels utilisant la modélisation. Dans un ensemble de référence E, un sous ensemble flou A de ce référentiel est caractérisé par une fonction d'appartenance μ de A, qui associe à chaque élément x de E, le degré $\mu_A(x)$, compris entre 0 et 1, pour lequel x appartient à A. Cette fonction est l'extension de la fonction caractéristique d'un sous ensemble classique [83], [84]. Elle peut être représentée sous forme de fonction triangulaire, trapézoïdale ou parabolique (figure 7).

La notion de sous ensembles flous a pour but de permettre des gradations dans l'appartenance d'un élément x à une classe A, c'est-à-dire d'autoriser un élément à appartenir plus ou moins fortement à cette classe.

Pour les systèmes dont le comportement n'est pas précisément connu (incomplet) des outils comme le raisonnement approximatif et les systèmes d'inférence flous permettent à la fois de décrire son comportement mais aussi d'en exploiter les résultats.

2.3. SIF : système d'inférence flou

Les systèmes d'inférence flous (SIF) sont composés d'une collection de règles floues qui ont la forme générale : "Si p alors q".

La conception des SIF s'appuie en général sur des connaissances expertes pour la définition des termes linguistiques correspondant à chaque variable (ensemble de fonctions d'appartenance, SEF) d'une part, et sur des algorithmes d'apprentissage pour la génération des règles d'autre part. Par exemple, les termes linguistiques associés au langage gestuel (*seconda, terza minore, terza mezzana, quinta*, etc.) correspondent au langage mimophonique.

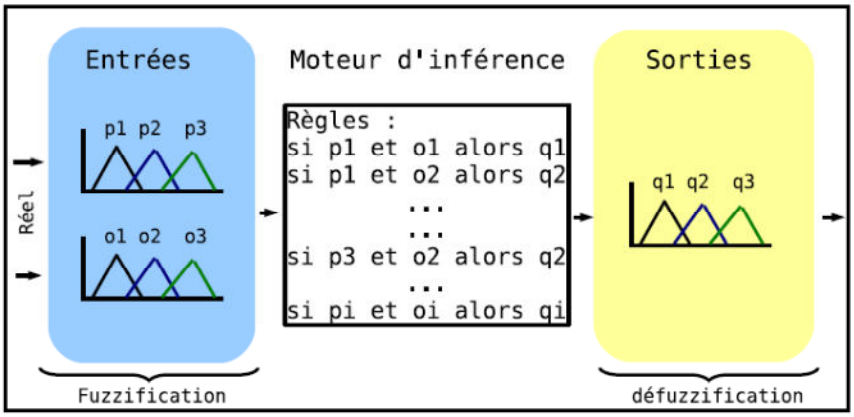


Figure 8 : Exemple de SIF

Les SIF sont à utiliser quand :

– il existe une expertise humaine que l'on veut exploiter et introduire dans un système automatique ;

– on veut extraire des connaissances à partir de données numériques, en les exprimant dans un langage proche du langage naturel ;

– on veut réaliser une interface homme-machine, donner des explications ou établir des diagnostics immédiatement interprétables.

La réalisation d'un système d'inférence passe par plusieurs étapes (figure 8) [74], [85], la fuzzification et la défuzzification des variables d'entrées et de sorties, et la réalisation d'un moteur d'inférence.

2.4. Fuzzification et défuzzification

L'opération qui consiste à transformer un ensemble réel en ensemble flou est appelée "fuzzification", et l'opération inverse est appelée "défuzzification". La fuzzification est l'étape qui consiste à quantifier à l'aide de termes linguistiques ou de valeurs floues les valeurs réelles d'une variable.

Fuzzifier des variables est une phase délicate du processus mis en œuvre par la logique floue, car il faut connaître toutes les variations possibles de la variable, sa ou ses fonctions d'appartenance.

La défuzzification est une phase décisionnelle, qui permet de transformer une valeur floue d'une variable en une valeur réelle. Il existe de nombreuses méthodes de défuzzification :

– méthode des hauteurs consiste à choisir comme valeur celle pour laquelle la hauteur vaut 1 (lambda = 1) ;

– méthode du centre de gravité (dite de Mamdani [76]) consiste à prendre pour valeur finale, en sortie, l'abscisse du centre de gravité de l'ensemble flou agrégeant les conclusions ;

– méthode des moyennes des maxima consiste à faire la moyenne des valeurs sommets des fonctions ;

– méthode des aires (dite de Sugeno [86]) consiste à prendre la médiane qui fait le partage de l'aire en deux.

Dans la partie suivante nous allons présenter notre méthode de reconnaissance de gestes basée sur les systèmes d'inférence floue.

2.5. Description du système

Le système de reconnaissance que nous allons décrire ici, est basé sur un langage gestuel. Chaque geste est capté à l'aide de gants de données et numérisé en temps réel grâce à une interface logiciel. Pour chaque entrée, nous avons donc six valeurs, correspondant aux six capteurs présents sur chaque gant. Le but du système de reconnaissance est de permettre d'associer à chaque ensemble de valeurs d'entrée une voix chantée.

Nous rappelons que chacun des trois chanteurs possède un ensemble propre de gestes. Il est donc nécessaire que ces trois ensembles de gestes soient traités indépendamment les uns des autres. Cependant, le système de reconnaissance des gestes reste le même pour les trois chanteurs. C'est pourquoi, nous ne présenterons dans cet article que le cas du chanteur de seconde.

2.5.1. Entrées

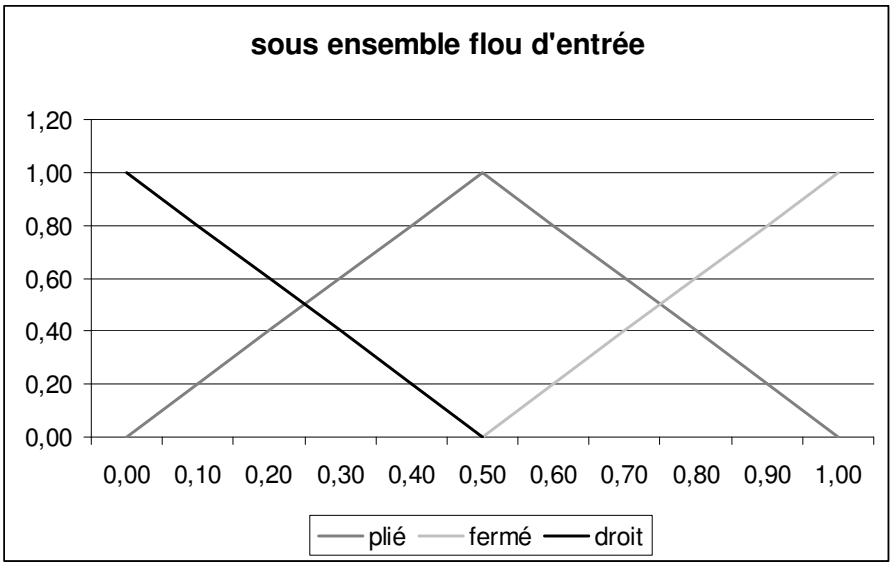


Figure 9 : d'un des Sous Ensembles Flous d'entrée

La première étape pour la réalisation et l'utilisation d'un SIF est l'identification des paramètres d'entrées.

Nous avons six entrées représentant les six capteurs. Il y en a un pour chaque doigt, plus un pour la main.

Les cinq entrées liées aux doigts sont modélisées à partir de trois SEFs caractérisant la position du doigt en fonction des coordonnées retournées par le capteur (Figure 9). Nous avons choisi d'utiliser des SEFs sous forme triangulaire car les gestes à modéliser peuvent être représentés à partir de trois états : Droit [0 ; 0.33], Plié [0.33 ; 0.66], Fermé [0.66 ; 1]. Chacun de ces états a une plage de valeurs comprise dans un intervalle. Nous constatons sur la figure 10 que la transition de coordonnées pour l'index entraîne un changement d'état et de geste. On dit qu'il y a transition entre deux gestes lorsque le système reconnaît un geste intermédiaire. Dans ce cas une transition correspond bien à la variation de la voix du chanteur, c'est un changement de note. Par exemple le passage progressif entre les gestes 1 et 2 est possible du fait de l'existence d'un état intermédiaire ; ce n'est pas le cas pour les notes non consécutives.

			
Pouce	[0; 0.33]	[0 ; 0.33]	[0; 0.33]
Index	[0.66; 1]	[0.33;0.66]	[0; 0.33]
Majeur	[0.66; 1]	[0.66 ; 1]	[0.66; 1]
Annuaire	[0.66; 1]	[0.66 ; 1]	[0.66; 1]
Auriculaire	[0.66; 1]	[0.66 ; 1]	[0.66; 1]
Position de la main	[0 ; 0.5]	[0 ; 0.5]	[0 ; 0.5]

Figure 10 : Sous Ensembles Flous pour l'index

La sixième entrée permet de représenter le capteur qui est sur la main, elle indique le sens du geste. Pour ce cas, seulement deux états sont nécessaires : Haut et Bas. Ceci nous permet de différencier par exemple les **1** et **I** (Figure 3).

2.5.2. Sorties

En sortie nous voulons obtenir le geste réalisé, parmi les quatorze gestes existants de la 'seconda'. Pour cela nous avons représenté la sortie à partir de quatorze SEFs.

La figure 11 présente nos SEFs de sortie. La réalisation des gestes n'étant pas continue, nous avons utilisé une échelle discrète pour les représenter. En effet ils sont pour la plupart indépendants. Les gestes identifiés comme très proches, avec transition, sont représentés par des SEFs imbriqués, exemple les trois premier gestes.

Les sorties sont identifiées à partir de l'échelle musicale (geste 1 => tonique du chant) définie par le langage mimophonique (partie 2.1)

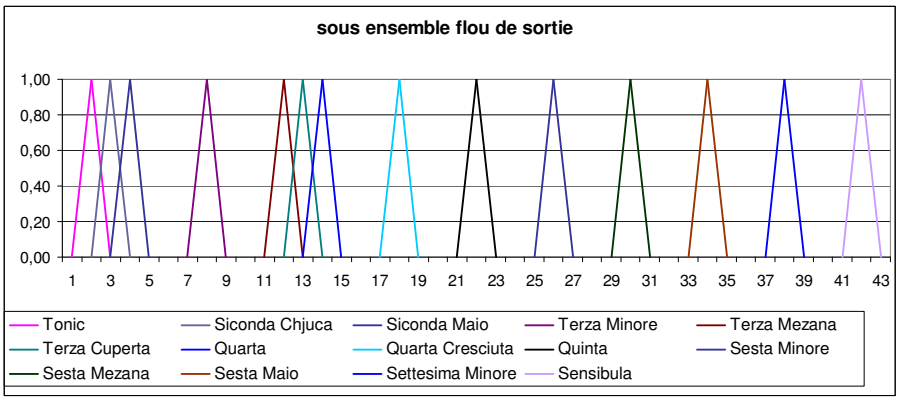


Figure 11 : SEFs représentant des sorties

Le lien entre les entrées et les sorties se fait à partir de règles. Chaque règle décrit la position de chaque doigt et de la main correspondant à un geste. Si l'ensemble des valeurs d'entrée vérifie une règle, il est alors associé à un geste.

2.5.3. Règles

L'étape suivante a été la réalisation des règles. Il y en a quatorze, une pour chaque geste du chanteur de seconde. Au vu du nombre restreint de règles, nous n'avons pas utilisé d'algorithme d'apprentissage, mais avons défini nos règles manuellement en associant l'état des entrées à un geste de la mimophonie (sortie). Avec cette méthode, les gestes non présents dans le langage ne sont pas reconnus. Par exemple, le geste « Pouce droit, Index fermé, Majeur droit, Annuaire fermé, Auriculaire droit » qui n'appartient pas au langage mimophonique ne sera ni reconnu par le système ni associé à une sortie (voix).

La figure 12 présente l'ensemble des règles pour un chanteur de seconde. Elles permettent, en fonction des coordonnées d'entrées, de définir le geste réalisé.

Leur principal intérêt est leur souplesse, car en représentant un sous ensemble de coordonnées par un sous ensemble flou ({Droit} – figure 10) nous ajoutons de la progressivité et de la flexibilité dans la prise en compte des coordonnées.

Les règles sont définies comme suit :

Règle	SI Pouce	ET Index	ET Majeur	ET Annuaire	ET Auriculaire	ET Position main	ALORS Signes
6	Droit	Fermé	Fermé	Fermé	Fermé	Haut	Tonica
4	Droit	Plié	Fermé	Fermé	Fermé	Haut	Siconda Chjuca
3	Droit	Droit	Fermé	Fermé	Fermé	Haut	Siconda Maio
2	Droit	Droit	Droit	Fermé	Fermé	Haut	Terza Minore
11	Fermé	Fermé	Fermé	Fermé	Fermé	Haut	Terza Mezana
7	Fermé	Plié	Plié	Plié	Plié	Haut	Terza Cuperta
9	Fermé	Droit	Droit	Droit	Droit	Haut	Quarta
12	Droit	Droit	Droit	Droit	Droit	Bas	Quarta Cresciuta
1	Droit	Droit	Droit	Droit	Droit	Haut	Quinta
10	Fermé	Fermé	Fermé	Fermé	Plié	Haut	Sesta Mezana
5	Droit	Fermé	Fermé	Fermé	Droit	Bas	Sesta Maio
13	Droit	Droit	Fermé	Fermé	Fermé	Bas	Settesima Minore
8	Plié	Plié	Fermé	Fermé	Fermé	Bas	Settesima Mezana
14	Droit	Fermé	Fermé	Fermé	Fermé	Bas	Ottava Suprana

Figure 12 : Ensemble des règles pour le chanteur de seconde

« SI Pouce est Droit ET Index est Fermé ET Majeur est Fermé ET Annuaire est Fermé ET Auriculaire est Fermé ET Position de la main est Haut ALORS Tonica »

2.5.4. Résultats

La dernière étape est la vérification des résultats. Nous pouvons fixer les coordonnées des capteurs et voir en fonction du geste réalisé lequel est réellement reconnu.

Nous avons testé les méthodes d'inférence les plus courantes (Moyenne des Max, Sugeno, Aires) afin de les comparer et de choisir celle qui nous paraissait la plus pertinente.

Le tableau figure 13, présente les résultats obtenus sur huit gestes représentatifs parmi les gestes de la mimophonie. Il met en évidence les valeurs obtenues pour quelques gestes avec trois méthodes de défuzzification. Nous constatons que pour des gestes sans transition comme le **3** et le **6**, les trois méthodes retournent la valeur attendue. Cependant, dans le cas des trois premiers gestes du tableau, seules les coordonnées de l'index changent. Nous sommes dans la même situation pour les gestes **1** et où seule la position du pouce est modifiée et entre les gestes et qui ne diffèrent que par la position de l'auriculaire. Dans ces trois cas, les méthodes de défuzzification de Sugeno et des Aires ne donnent pas les valeurs attendues. Par contre, nous pouvons remarquer qu'avec la méthode dite des moyennes des max le pourcentage de reconnaissance est de cent pour cent. Nous avons constaté que ces trois méthodes de défuzzification retournent de bons résultats pour des gestes sans transition. Cependant, seule la méthode dite des moyennes des max permet de reconnaître des gestes avec transition.

Signes / Entrées / Sorties	1		2	3			6	7
Coordonnées des entrées :								
- Pouce	**0,05**	0,05	0,05	0,05	**0,91**	0,91	0,05	0,1
- Index	**0,96**	**0,55**	**0,1**	0,1	0,96	0,96	0,9	0,07
- Majeur	0,9	0,9	0,9	0,04	0,9	0,9	0,89	0,89
- Annuaire	0,92	0,92	0,92	0,92	0,92	0,92	0,89	0,89
- Auriculaire	0,88	0,88	0,88	0,88	**0,88**	**0,48**	0,06	0,9
- Position de la main	0,06	0,06	0,06	0,06	0,06	0,06	0,94	0,87
Valeurs Attendues	1	2	3	6	9	20	23	26
Défuzzification par la méthode de Moyenne max	1	2	3	6	9	20	23	26
Défuzzification par la méthode Sugeno	0,14	1,767	2,767	6	11,519	17,959	23	26,477
Défuzzification par la méthode des Aires	0,095	1,518	2,724	6	12,161	17,255	23	26,655

Figure 13 : Exemple de résultats des tests de reconnaissance sur huit gestes

En s'appuyant sur ces résultats on peut considérer que dans notre cas, utiliser un système d'inférence pour répondre à notre problématique permet d'obtenir les aspects suivants : (i) la

reconnaissance de geste, est une solution très avantageuse, (ii) le système est facile à décrire, (iii) les règles sont peu nombreuses et (iv) les résultats sont conformes à nos attentes.

3. FTM

FTM est une bibliothèque partagée qui fournit un ensemble d'objets simples et temps réel et un ensemble de services optimisés pour être utilisés dans Max / MSP. L'idée de base de FTM est d'étendre les types de données échangés entre les objets dans un patch Max / MSP par des structures de données complexes telles que des séquences, des matrices, des dictionnaires, les tuples et tout ce qui pourrait sembler utile pour le traitement de la musique, le son et des données de capture de mouvement. FTM est la base de plusieurs ensembles de modules pour Max / MSP spécialisés dans le suivi de partitions, l'analyse/synthèse de son, la modélisation statistique et l'accès aux bases de données. Conçu pour des applications d'accompagnement automatique, de traitement du signal avancé et d'analyse gestuelle, les bibliothèques utilisent un ensemble commun de structures de données de base de FTM. Elles sont parfaitement interopérables et s'intègrent dans le paradigme de programmation modulaire de l'environnement Max/MSP [87].

La motivation première pour le développement de FTM a été la nécessité d'une représentation flexible de partitions liée au suivi de partition et une représentation efficace des matrices et des vecteurs permettant une implémentation de plusieurs algorithmes d'analyse/synthèse dans un cadre unifié.

Aujourd'hui, les modules externes pour les deux applications ont été développés dans des packages: *Suivi* et *Gabor*.

- Le paquet *Suivi* contient des modules effectuant le suivi de partition sur la base des Modèles de Markov Cachés (MMC) [88]. Ce paquet se compose principalement de deux objets pour le suivi de partition sur MIDI et entrée audio. Ils référencent les objets *track* (séquence d'éléments marqués dans le temps) qui contiennent les informations de la partition.
- Le paquet *Gabor* est une boîte à outils pour des applications d'analyse et de synthèse.

Un autre package, MnM («Music is Not Mapping") [89], offre un système unifié pour les différentes techniques de classification, de reconnaissance et de traitement des données de capture de mouvement, du son et de la musique.

Les fonctions actuellement mise en œuvre dans MnM comprennent notamment PCA (Principal Component Analysis), GMM (Gaussian Mixture Models), Matrix/Vector Statistics (min, max, mean, std, histogram, distance de mahalanobis) et HMM (Hidden Markov Models) qui est utilisé pour le suivi de geste [90].

Au vu de sa nature non parfaitement reproductible, la modélisation du geste par un processus stochastique s'est imposée dans ce domaine. Provenant initialement des travaux effectués sur la reconnaissance de la parole [68], les Hidden Markov Models (HMM), ou Modèles de Markov Cachés (MMC), sont utilisés dans la plupart des méthodes de reconnaissances de formes ou de processus temporels (langage, écriture manuscrite, . . .). Le modèle est assez souple pour pouvoir s'appliquer à de nombreux domaines mais nécessite d'être adapté selon le domaine considéré.

A partir d'une utilisation classique des MMC pour la reconnaissance du geste [91] [92], de nombreuses améliorations ont été proposées afin de prendre en compte la spécificité du phénomène observé. La plupart se concentrent sur une meilleure prise en compte des données d'apprentissage [93] mais supposent toujours qu'il y a plusieurs réalisations du même geste.

Une extension intéressante des MMC [94] [95] [96] a été proposée par A.D. Wilson afin de paramétrer le geste et donc de prendre en compte les déformations (linéaires ou non) qui peuvent apparaître dans l'exécution du geste lors de la reconnaissance et de pouvoir les mesurer. Mais la phase d'apprentissage reste fastidieuse : chaque geste servant à l'apprentissage doit être paramétré manuellement.

X. Ge propose également une alternative intéressante afin de trouver une forme donnée dans un flux d'information [97] [98], tout en prenant en compte de possibles déformations dans le signal observé. Ce dernier est modélisé par une semi-chaîne de Markov [99] ce qui permet de mieux prendre en compte les exécutions de geste à des vitesses différentes mais ne bénéficie plus de la simplicité d'implémentation d'une chaîne de Markov classique.

D'autres pistes sont néanmoins proposées. Certains partent de l'algorithme de Dynamic Time Warpping (DTW) pour obtenir une mesure de similarité entre deux séquences [100] [101]ou pour effectuer un alignement [102]. La plupart de ces méthodes ne sont pas orientées temps-réel car l'algorithme de DTW ne peut s'effectuer qu'une fois le signal entièrement disponible.

Des travaux ont été entrepris à l'IRCAM afin d'adapter cet algorithme à un alignement proche du temps réel (DTW à court-terme) [103].

Le module de reconnaissance de gestes que nous avons développé pour notre projet s'appuie sur la distance de Mahalanobis qui peut être comparée à une distance euclidienne tenant compte de la distribution en variance des données. L'idée générale est de calculer une distance qui est en fait l'écart de similitude entre deux séries de valeurs.

L'objet max qui permet de faire ça s'appelle `mnm.mahalanobis` et a été développé par l'équipe IMTR de l'IRCAM [104]. Cet objet est contenu dans les librairies FTM, MnM et gabor qui permettent de faire du traitement de matrices dans MAX/MSP [105] pour de la reconnaissance de gestes (MnM) et du son (gabor) [106].

En pratique, il faut apprendre chacun des gestes pour donner des valeurs de *mu* et *sigma* (deux paramètres de l'objet `mnm.mahalanobis` – Cf. 5.1) à chacun des gestes. Une fois les gestes chargés et liés au module, le système est prêt pour faire de la reconnaissance : chaque nouveau geste en entrée est comparé avec l'ensemble des gestes appris sur la base de sa distance de Mahalanobis aux gestes appris. La plus petite distance permet alors de dire quel geste a été reconnu (geste le plus proche, qui présente le plus de similitudes).

4. Discussion et conclusion

Nous avons présenté dans ce chapitre l'utilité pour notre projet de se doter d'un module de reconnaissance de gestes. Nous avons testé deux méthodes de reconnaissance littéralement différentes mais néanmoins intéressantes. La première méthode, basée sur les systèmes d'inférence flous, utilise pour guider son comportement un ensemble de règles. Nous avons en entrée de notre module de reconnaissance une série de sept valeurs correspondant aux cinq capteurs de flexion et aux axes x et y de l'accéléromètre. Dans le module de reconnaissance basé sur les SIFs, chaque valeur d'entrée correspond à un état du doigt qui peut être droit, plié ou fermé. L'ensemble des règles définies décrivent l'ensemble des gestes de la mimophonie. Chaque entrée est ensuite comparée à l'ensemble des règles. A chaque règle correspond une et une seule sortie qui est le nom du geste reconnu.

La seconde méthode que nous avons présentée se base sur la librairie FTM du logiciel MAX/MSP développé à l'IRCAM. La reconnaissance de gestes se fait par le calcul de la distance mahalanobis. Chaque liste de valeurs d'entrée est comparée avec l'ensemble des valeurs des gestes appris. Le geste qui a la distance la plus faible est celui qui ressemble le plus au geste d'entrée.

Le développement et l'implémentation de la méthode basée sur les SIF s'est avérée plus simple et plus rapide. En effet, le système de règles est assez explicite et compréhensible par une personne non experte. Cependant, dans un tel développement, les données d'entrées qui proviennent du logiciel MAX/MSP sont traitées dans un SIF et les sorties sont ensuite réintroduites dans MAX/MSP. La « qualité » du temps réel se trouve ici altérée. A contrario, le développement du module dans MAX/MSP nécessite une meilleure connaissance du domaine de la reconnaissance gestuelle. De plus, il est indispensable d'apprendre en amont l'ensemble des gestes du langage. Ce qui peut vite devenir long et fastidieux. Par contre l'intégration de ce module dans l'application, elle aussi développée avec MAX/MSP est plus fluide. C'est donc cette solution que nous avons privilégié dans notre développement.

Dans la création de notre outil d'aide à l'apprentissage de chants polyphoniques, nous pouvons à ce stade capter et reconnaitre l'ensemble des gestes du langage musical. Il faut maintenant associer à chacun des gestes et en temps-réel une voix chantée à la hauteur indiquée.

CHAPITRE IV :

Production sonore

Lors des chapitres précédents, nous avons tout d'abord défini le langage gestuel utilisé pour l'apprentissage de chants polyphoniques. Nous avons ensuite décrit le système de captation que nous avons choisi d'utiliser ainsi que les différentes techniques de reconnaissances de gestes que nous avons développées.

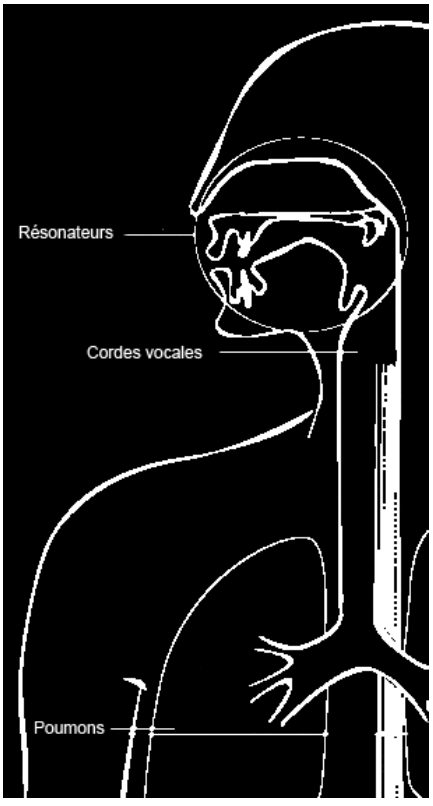
A ce stade du projet, nous sommes donc en mesure d'identifier chaque geste musical effectué. Pour que l'outil d'apprentissage soit complet, il faut pouvoir associer à chacun des gestes une voix chantée à la hauteur de note imposée par les gestes de l'apprenant.

Deux « voies » nous semblent intéressantes :

- La création de voix de synthèse, qui a l'avantage de permettre à l'utilisateur de manipuler tous les paramètres qui constituent la voix : le timbre, l'intensité, la texture, le changement de voyelles, le vibrato, etc.
- L'utilisation de voix enregistrées. Cette solution, simple à mettre en place, permet d'obtenir un rendu plus naturel. De ce fait, le sentiment d'être accompagné de chanteurs physiquement présents, augmente chez l'apprenant l'impression de réalisme.

1. Voix synthétisées

1.1. Description de l'appareil vocal



Afin de pouvoir synthétiser une voix humaine chantée avec un ordinateur, il est nécessaire de connaître son fonctionnement et les organes mis en jeu (poumons, cordes vocales, résonateurs).

Pour produire un son, il faut tout d'abord générer un souffle. Celui-ci va provoquer une pression sous les cordes vocales et les faire vibrer. Ces vibrations sont ensuite amplifiées et modulées par les résonateurs (pharynx, cavités nasales, sinus et bouche). L'appareil vocal peut être décrit en deux parties. La partie basse qui est la source et qui émet un son et donne la hauteur de la voyelle. La partie haute qui est le filtre et modifie la résonnance selon la voyelle.

La position des résonateurs chez l'homme crée dans le son généré des formants, c'est-à-dire des zones d'énergie renforcées dans un spectre par phénomène de résonance. Une voyelle est caractérisée par ses formants qui permettent de l'identifier.

La synthèse de voix chantées se base sur des modèles de formants émettant des fonctions d'ondes. La somme de plusieurs formants modélisés donne le signal souhaité.

La voix chantée, très proche de la voix parlée, se différentie par :

- sa prosodie, qui suit la phrase musicale.
- la tenue des voyelles. Dans la voix chantée, se sont les voyelles qui portent l'information de hauteur de la note. Les voyelles chantées sont souvent plus longues que dans la parole et contribuent davantage à qualifier le timbre de la voix.
- l'ambitus du chanteur. Dans le cas des polyphonies corses, nous parlerons de voix de basse, de seconde (ou tonique) et de tierce.
- le vibrato qui consiste à faire plus ou moins varier le son autour d'une hauteur « fixe ».

1.2. La synthèse de voix chantée avec le programme CHANT

Le programme CHANT a été conçu par l'équipe Analyse/Synthèse de l'IRCAM comme un outil attaché à l'analyse et à la synthèse de la voix chantée mais aussi à la production musicale et la composition [107]. Dans sa version de base, CHANT offre la possibilité d'influer sur une centaine de paramètres regroupés en rubriques : fréquence du fondamental, variations aléatoires du fondamental, vibrato, variations aléatoires du vibrato, spectre: formants et fondamental, pente du spectre, calcul automatique du spectre, intensité du son, enveloppe locale des formants et contrôle sur la synthèse. Ces paramètres peuvent être stockés dans un fichier avec leurs valeurs.

Le programme pour la synthèse de la voix reprend le modèle de production vocale de l'humain. Il s'articule sur les deux étages décrits dans le paragraphe précédent : la source qui émet un son pseudo périodique et le filtre qui donne une résonnance particulière selon la nature de la voyelle. Afin de simuler le fonctionnement des résonateurs de l'appareil vocal humain on peut utiliser soit des filtres agissant sur le spectre en fréquences, soit des FOF (Fonctions d'Onde Formantiques) opérant dans le domaine temporel. Dans le programme CHANT, cinq formants sont modélisés, en utilisant cinq filtres en parallèle. Chacun émet des fonctions d'ondes correspondant à un formant donné et le signal voulu est donné par la somme des cinq.

La figure 14 décrit ce principe. Sur l'entrée sont envoyées des impulsions régulières simulant les coups de glotte (K). Chaque fonction FOF_i fabrique des fonctions d'ondes correspondant à un formant donné $S(k)_i$. En sortie apparait la voix synthétisée (S).

Nous pouvons alors écrire que : $S = \sum_{i=1}^{5} S(k)i$

Chaque paramètre des FOF peut être ajusté pour modifier le son.

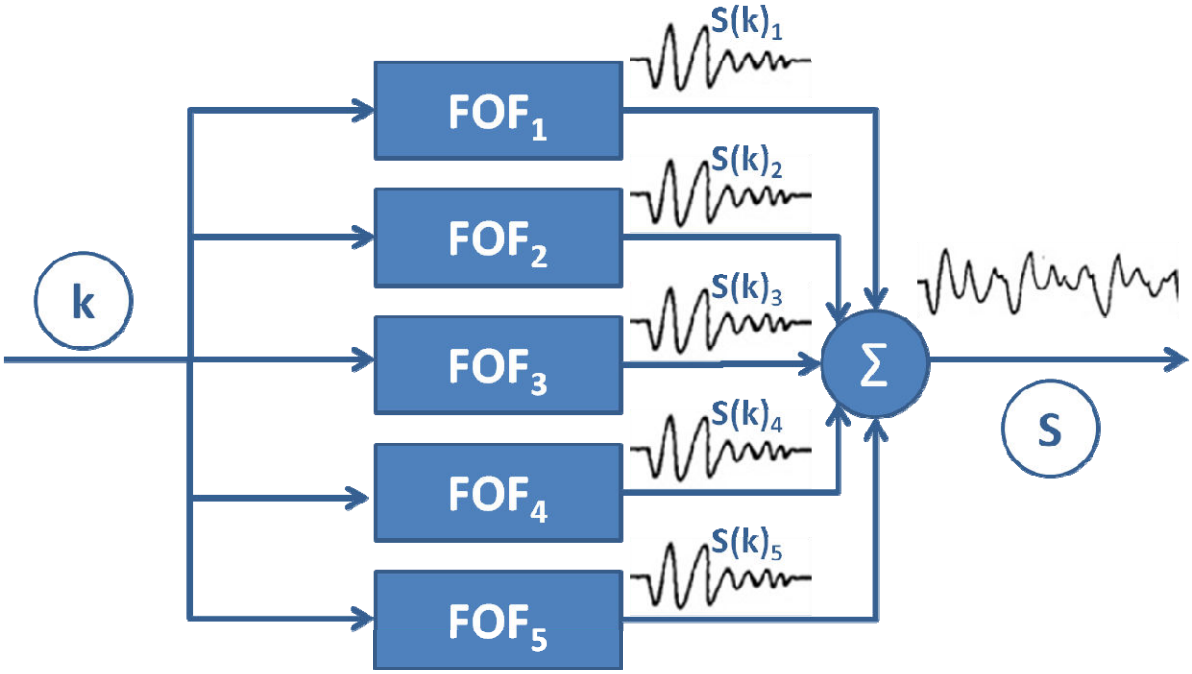


Figure 14 : Synthèse de voix basée sur FOF

Dans notre projet de synthétiser des voix pour le chant polyphonique corse, nous avons créé cinq voyelles chantées : lal, lél, lil, lol et loul.

Le module CHANT développé à l'IRCAM (Figure 15) permet de faire varier de nombreux paramètres comme par exemple la vitesse du vibrato, son amplitude, le timbre, etc. Nous voyons dans la figure ci-dessous qu'il existe un objet `p formdata` par voyelle à synthétiser.

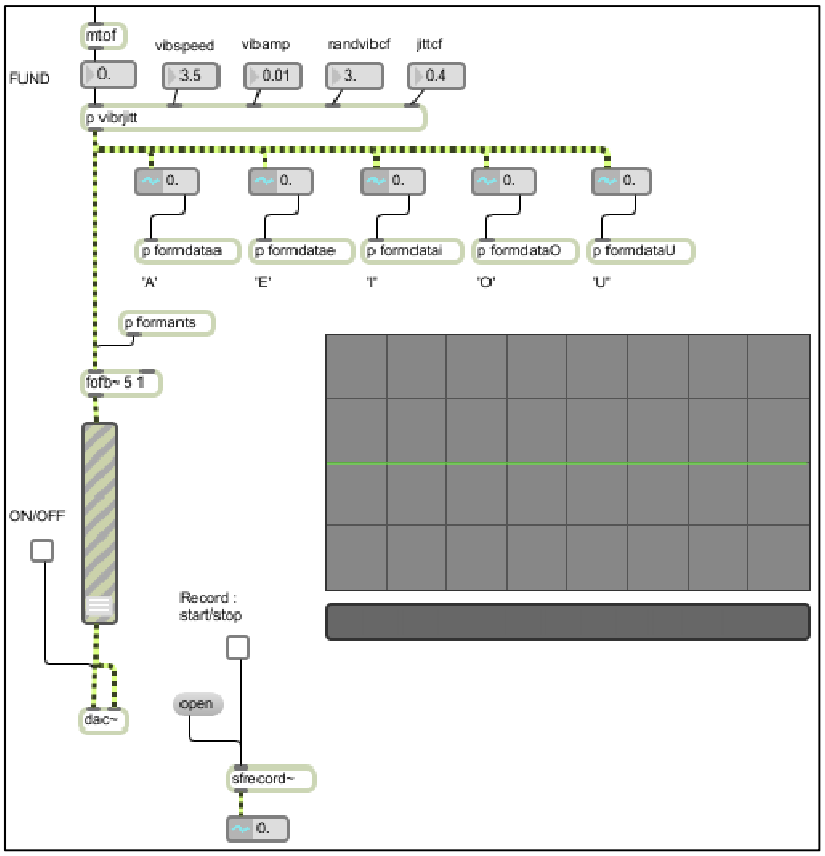


Figure 15 : patch principal du module CHANT

Chacun de ces objets fait appel à un autre module contenant la caractérisation formantique de la voyelle (Figure 16). Les paramètres de chaque formant peuvent être attribués soit en faisant varier chaque champ jusqu'à obtenir le son souhaité, soit en effectuant une analyse formantique d'une voyelle chantée dans un extrait sonore (avec un logiciel de traitement sonore tel qu'Audiosculpt par exemple). Il est alors possible d'extraire automatiquement les valeurs pour les cinq premiers formants d'une voyelle. La voix synthétisée se rapproche un peu plus du modèle.

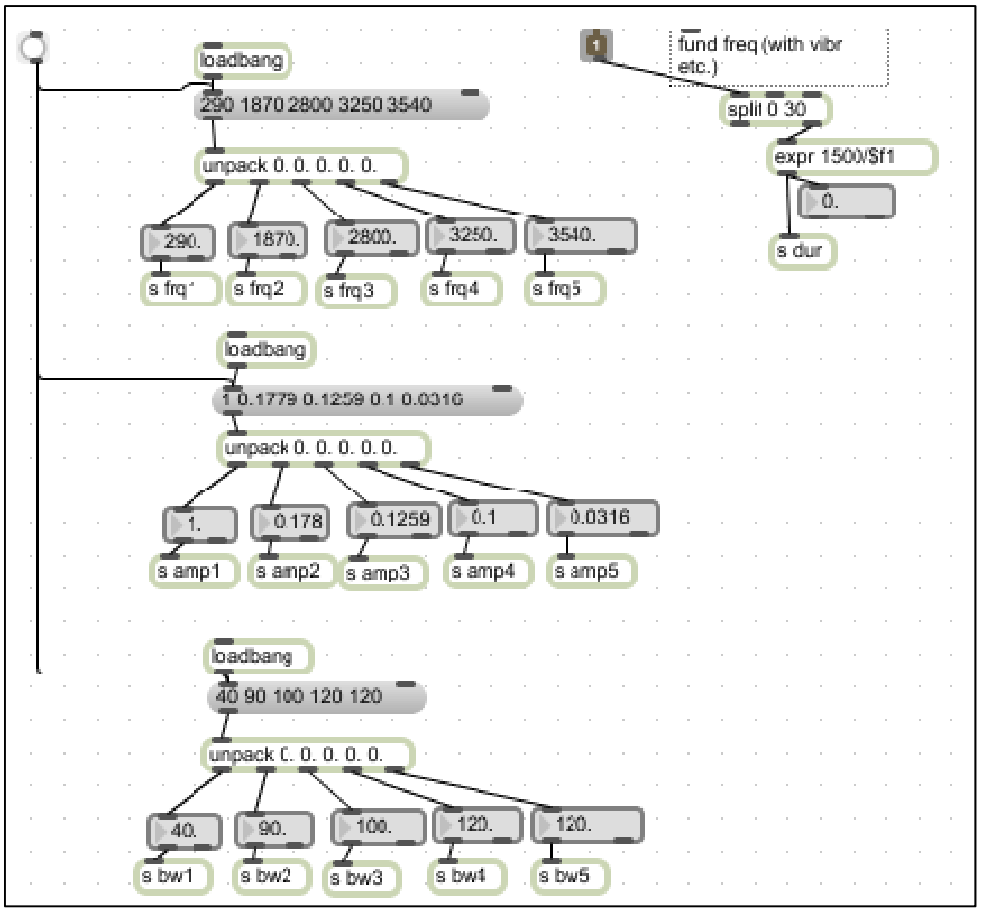


Figure 16 : description formantique pour la synthèse de la voyelle 'i'

L'ensemble des paramètres ajustés pour optimiser la qualité du son synthétisé agit comme une série de filtres sur la fonction FOF du module. On peut observer sur la figure 15 que nous avons ajouté au module CHANT une fonction d'enregistrement permettant de sauvegarder une piste audio de la synthèse retournée. Ceci à des fins de tests de comparaison avec une piste originale, lorsque les paramètres des formants de CHANT sont issus d'une analyse formantique.

Sur la figure 17, nous voyons les cinq formants issus du spectre d'une voyelle 'o' chantée dans Audiosculpt.

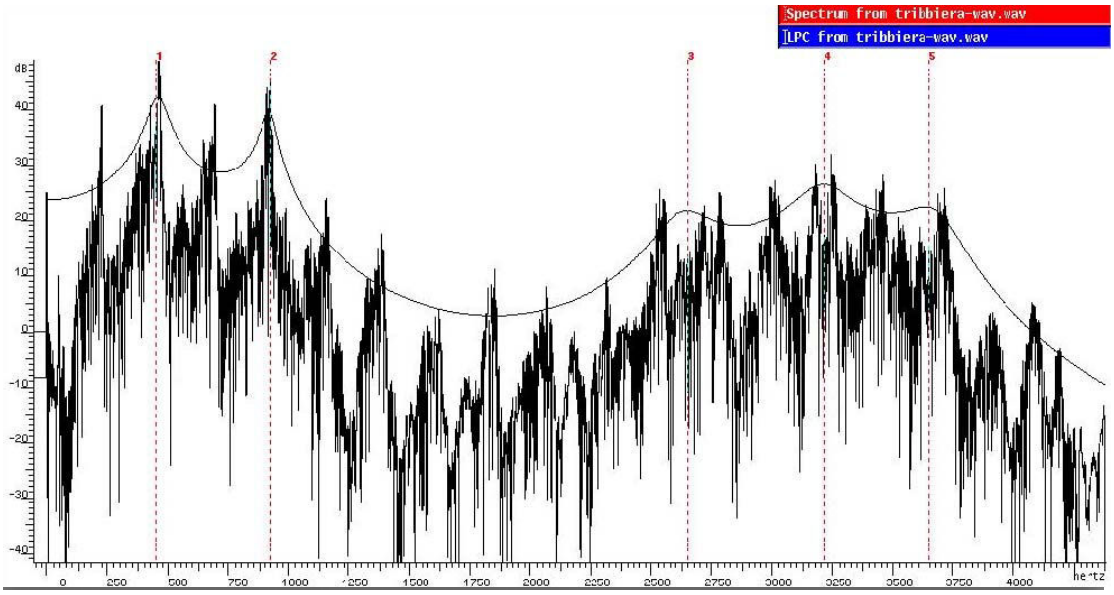


Figure 17 : les formants qui apparaissent sur le spectre d'une voyelle chantée

2. Voix enregistrées

Il semble pertinent de travailler avec des voix enregistrées pour plusieurs raisons. Tout d'abord, d'un point de vu pédagogique. Rappelons que l'outil développé doit permettre l'apprentissage du chant polyphonique. Bien que la qualité des voix synthétisées soit proche de la réalité, l'oreille humaine différencie encore les deux sons et ceci peut déranger un utilisateur peu habitué à ce genre d'exercice. Cette méthode de production sonore présente également l'avantage d'être simple à mettre en place et peu coûteuse en calcul.

Pour obtenir l'ensemble de notes correspondant aux intervalles de la mimophonie, nous avons enregistré un chanteur qui a parcouru l'intégralité de l'échelle musicale, pour une tonalité donnée. Chaque note est ensuite enregistrée dans un module et sera associée à un geste musical. Nous avons vu précédemment (Cf. 3.4) que chaque geste de la mimophonie était identifié par un module de reconnaissance. Ce dernier est alors connecté comme point d'entrée au module de production sonore. Les modules qui ont été développés dans ces travaux sont détaillés dans la partie 5.1.

3. Conclusion

Nous avons présenté dans cette partie la production de voix chantées qui boucle le processus qui va du geste au son. Nous avons vu comment, pour chaque geste de la mimophonie, nous avons attribué une note correspondant à la hauteur indiquée par les gants de données. Les deux solutions exposées présentent un intérêt certain dans notre application.

L'utilisation de voix synthétisée, créée à partir du module CHANT, a le grand avantage que l'on peut faire varier un grand nombre de variables assez rapidement. Et notamment la fréquence fondamentale qui est gérée ici comme une variable d'entrée à la fonction FOF. Il est donc possible de « régler » la tonalité d'un chant en fonction des capacités vocales de l'apprenant. Ce module de production sonore est parfaitement adapté à la création musicale.

La production de voix enregistrées et samplées offre un intérêt plus pédagogique. Ne nécessitant pas de connaissances particulières en matière de synthèse de voix, elle est plus simple à mettre en place. Il est par contre indispensable de posséder les enregistrements de l'ensemble des intervalles du langage. Les notes étant enregistrées, il n'est donc pas possible de faire varier la tonalité du chant, à moins d'avoir un enregistrement par tonalité. Cependant, le rendu des voix chantées est plus naturel et le timbre plus proche de la réalité. Ce qui offre pour l'apprenant un confort non négligeable.

CHAPITRE V :

Mise en œuvre et développement

MAX/MSP est un logiciel musical permettant de faire de la synthèse sonore, de l'analyse, de l'enregistrement, ainsi que du contrôle d'instrument MIDI. C'est un environnement visuel pour la programmation d'applications sonores interactives. Il a été développé par l'IRCAM dans le milieu des années 1980 et qui est maintenant distribué par Cycling74[8]. Il est l'un des logiciels musicaux parmi les plus utilisés par les musiciens professionnels et les amateurs.

1. Développement des modules

Nous avons vu dans les chapitres précédents comment capter les gestes de la mimophonie à l'aide des gants de données. Nous avons ensuite numérisé le signal des gants pour obtenir un ensemble de valeurs qui sont propres à chaque geste. Nous avons démontré la nécessité d'utiliser un module de reconnaissance des gestes. Et enfin comment un son est attribué à chaque geste. Ce même processus qui va du geste au son (Figure 18) nous a guidé pour le développement de l'application pour l'apprentissage de chants polyphoniques.

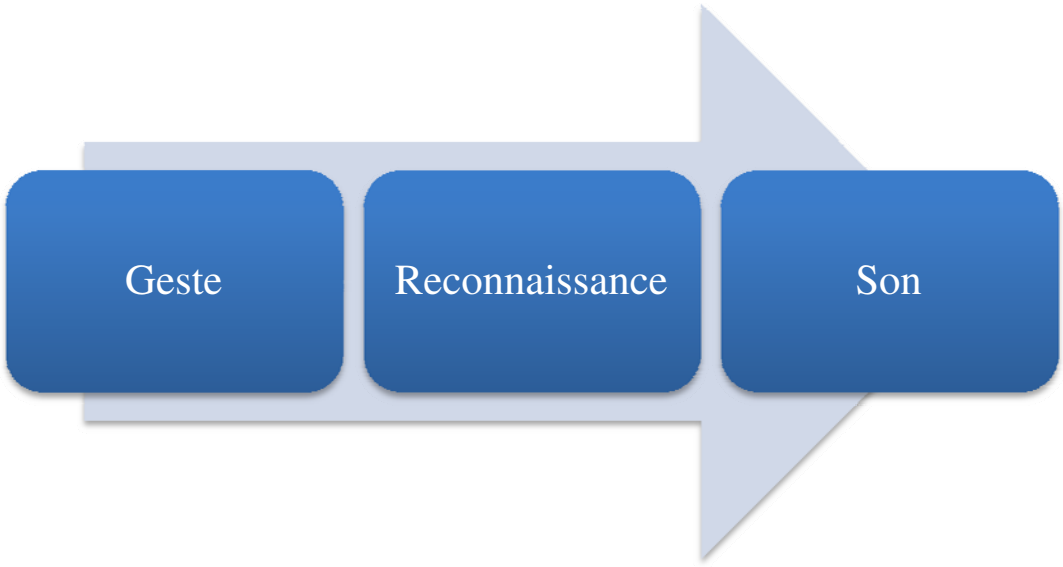


Figure 18 : processus du geste au son

Le premier module est donc celui qui va récupérer le signal des gants de données sans fil pour retourner une série de huit valeurs (Figure 19).

[8] www.cycling74.com

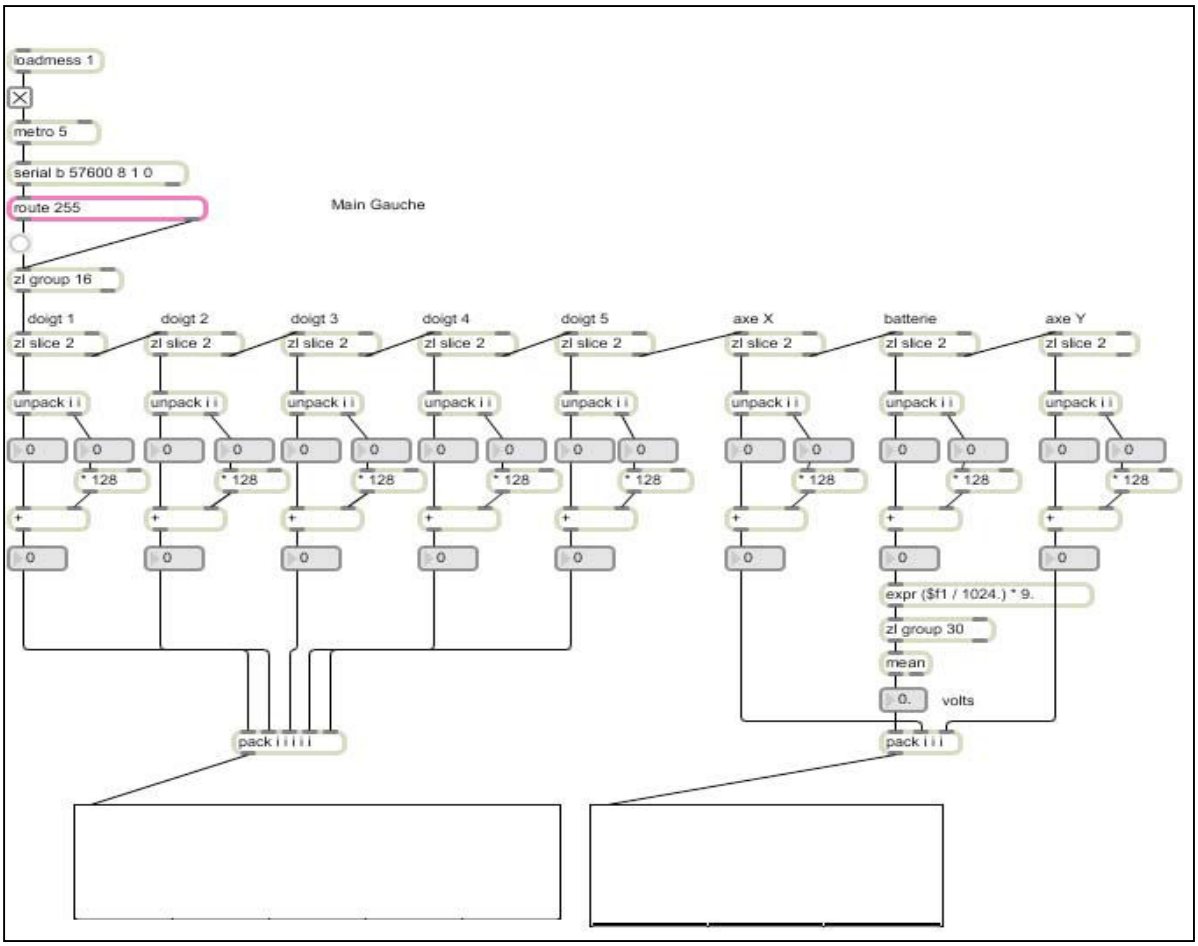


Figure 19 : module de réception des données issues des gants

Le premier objet `loadness` active automatiquement le module quand le fichier est ouvert. L'objet `metro` agit comme un métronome qui déclenche un « bang » à intervalle régulier (valeur passée en paramètre et en milliseconde). Le port série qui sera écouté par le module est désigné par l'objet `serial`. Cet objet ne fonctionne qu'avec des ports et des périphériques pris en charge par le pilote série standard. Il ne fonctionne pas avec un port USB, sauf si un adaptateur USB/série est utilisé. Le premier paramètre donne le port à écouter, ici la valeur `b` désigne le second port série. Le second paramètre représente le débit en baud du port série (ici 57600 bauds), le troisième paramètre spécifie le nombre de bits de données. Le dernier argument indique le nombre de bits d'arrêt du port série. L'objet `route` permet de reconstituer les trames de bits transmises des gants (Cf 2.4.5 Figure 5). L'objet `zl_group` regroupe les données reçues en entrée et les distribue en sortie sous forme de listes de longueur passée en paramètre. L'objet `zl_slice` propage la liste des données dans les différentes sous-listes, correspondant à chaque capteur. Ici les deux premiers éléments de la liste sont à associer au capteur du premier doigt et le reste de la liste est transmise au

`zl_slice` suivant et ainsi de suite. Les cinq premiers `zl_slice` correspondent aux cinq capteurs de flexion et aux cinq doigts de la main. Les trois autres sont les axes x et y de l'accéléromètre et le niveau de charge de la pile. L'objet `unpack` sépare la liste en éléments. Chacun d'eux est envoyé sur une sortie différente. Les deux valeurs sont ensuite additionnées dans l'équation suivante : $V = v1 + v2 \times 128$. Chaque valeur V de chaque capteur est alors regroupée dans une nouvelle liste par l'objet `pack`. Cette dernière liste contient pour chaque geste les valeurs de chaque capteur et pourra être traitée par le reste des modules de MAX/MSP. Le niveau de batterie est calculé en faisant la moyenne des trente dernières valeurs reçues par l'objet `mean`. Le résultat est exprimé en volt.

Les valeurs de sortie du module précédent sont transmises en entrée du module de reconnaissance du geste (Figure 20).

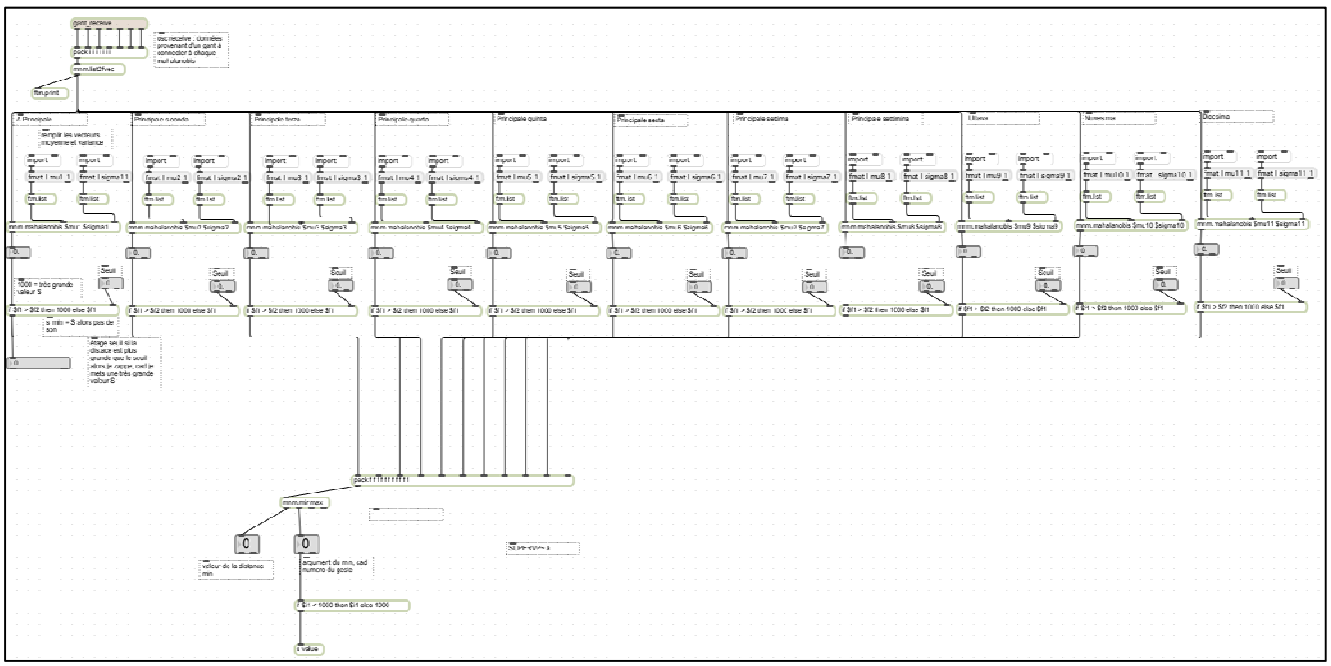


Figure 20 : module de reconnaissance des gestes provenant des gants de données

Le module de reconnaissance de gestes reçoit les valeurs issues du module précédent par l'objet `gant_receive` (qui est le nom du module). Cet objet qui reçoit deux listes les concatène en une seule grâce à `pack i i i i i i i`. Nous avons alors une liste de sept entiers (la valeur de puissance de la pile n'étant pas utile ici) qui décrit un des gestes. Cette liste est ensuite convertie en vecteur par l'objet `mnm.list2vec` afin que celle-ci soit utilisable par des objets FTM et MnM. Ce vecteur est ensuite transmis à l'ensemble des objets `mnm.mahalanobis` du module. Il y a autant de cet objet qu'il y a de gestes dans le langage.

Par exemple, si nous traitons de la main gauche, il s'agit donc du langage de la voix de basse, il y a donc onze gestes à tester.

L'objet `mnm.mahalanobis` (cf. 3.3) utilise la distance de mahalanobis qui est basée sur la corrélation entre des variables par lesquelles différents modèles peuvent être identifiés et analysés. C'est une manière utile de déterminer la similarité entre une série de données connues et d'autres inconnues. Il est donc nécessaire, pour chacun des gestes, d'avoir un fichier texte contenant les valeurs « références » du geste. Le meilleur moyen d'effectuer un bon apprentissage d'un geste est de le faire plusieurs fois avec le gant (Figure 21), d'enregistrer les valeurs et d'en faire une moyenne.

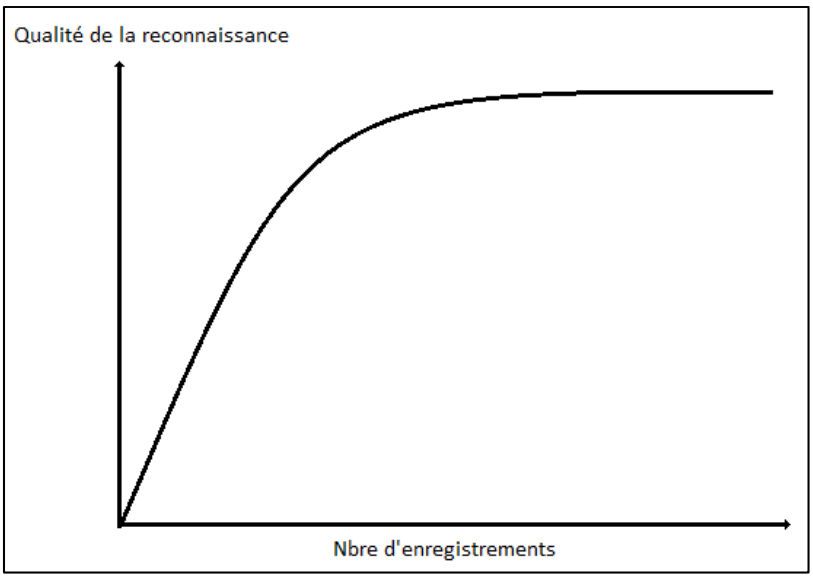


Figure 21 : la qualité de la reconnaissance d'un geste est liée au nombre d'enregistrements réalisés

Les valeurs références du geste contenues dans le fichier texte sont importées via l'objet `import` puis converties en une matrice FTM et stockées dans une variable par `fmat | mu1`. Nous avons ajouté une valeur seuil qui permet de tester la distance de mahalanobis retournée pour chaque geste. Ceci permet de ne pas traiter des valeurs excessivement élevées. La distance de mahalonobis retournée pour chaque geste est donc testée à la valeur du seuil fixée par l'instruction `if $f1 > $f2 then 1000 else $f1`. L'ensemble des distances sont concaténées dans une liste puis on y recherche (`mnm.minmax`) la valeur minimum. Le geste ayant la distance de mahalanobis la plus courte est donc le geste reconnu et sera transmis en sortie du module.

Au geste capté et reconnu, il faut maintenant lui associer un son. Si on utilise le module CHANT pour des voix de synthèse (Cf. 4.1.2), chaque geste correspond à un intervalle avec la

tonalité du chant. Si par exemple la tonalité est DO, si le geste reconnu est celui de la quinte, il faudra augmenter la note de DO de 701,96 cents dans le module CHANT. La hauteur de la note synthétisée sera un SOL.

Si on utilise des voix enregistrées (Cf. 4.1.1), il faut utiliser la valeur du geste reconnu dans un sampler (Figure 22) afin de trouver la note correspondante.

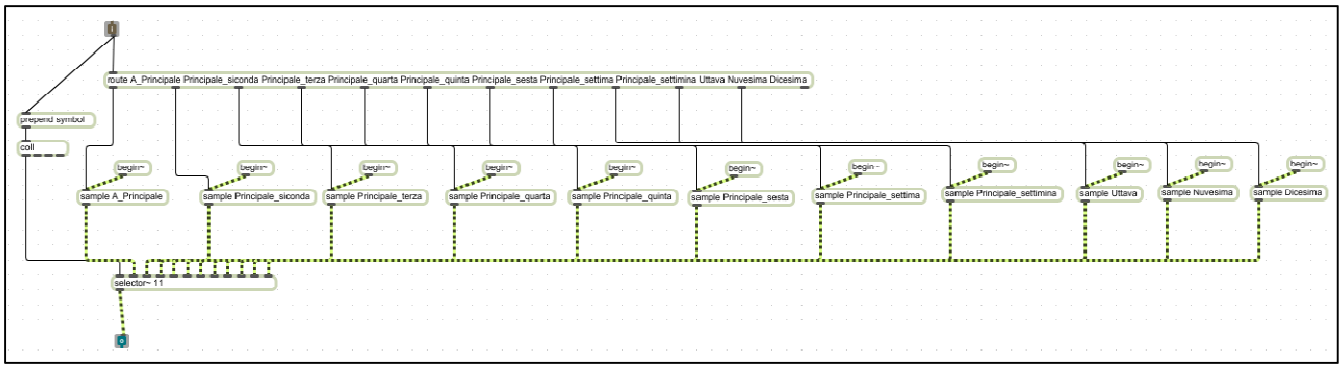


Figure 22 : module de sampler

En fonction du geste passé en entrée du module, l'objet `route` (qui contient la liste des sons enregistrés en paramètre) va sélectionner l'extrait sonore associé. Il va ensuite exécuter un dernier module (Figure 23) qui va jouer la piste. L'objet `buffer~` fonctionne comme un buffer de mémoire dans laquelle les échantillons sont conservés en attendant d'être sauvés, joués, modifiés, etc. L'objet `sig~` convertit des entiers en signal et l'objet `groove~` joue en boucle les échantillons audio stockés dans le buffer.

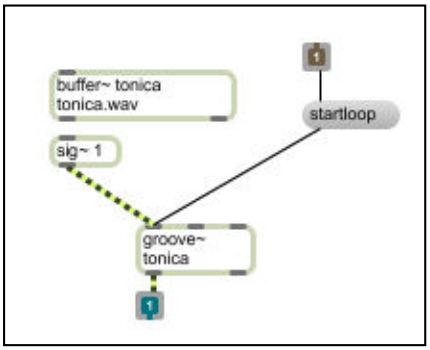


Figure 23 : module de lecture d'un son

Nous avons donc développé une série de modules avec le logiciel MAX/MSP et un ensemble de bibliothèques dédiées au geste et au son comme FTM et MnM. Ces modules permettent de capter les signaux émis par les gants de données sans fil. De traiter ces données, de les formater en liste d'entiers et de les transmettre à un module de reconnaissance des gestes basé sur le calcul de distance de mahalanobis. La liste de valeurs références ayant la distance de mahalanobis la moins élevée est considérée comme le geste reconnu. Ce geste est

ensuite transmis au module de production sonore, soit CHANT, soit un module de sampler qui jouera une piste sonore préalablement enregistrée.

CHAPITRE VI :

Réseaux de capteurs sans fil : suivi d'espèces halieutiques

Durant mes années de doctorat, j'ai acquis des compétences en informatique, en déploiement de capteurs et en acoustique que j'ai pu mettre à profit dans un nouveau projet de recherche scientifique.

En effet, j'ai intégré la plateforme de recherche de l'Université de Corse STELLA MARE (Sustainable TEchnologies for LittoraL Aquaculture and MArine REsearch) et participé au lacement d'un projet scientifique de modélisation comportementale d'espèces marines.

L'UMS-CNRS 3514 STELLA MARE, rattachée à l'INstitut Ecologie et Environnement (l'INEE), a pour objectif la maîtrise et la gestion intégrée des ressources halieutiques et littorales de Corse. La plateforme de recherche STELLA MARE, clairement axée sur l'ingénierie marine, est un environnement propice à de nouveaux projets scientifiques visant à améliorer la protection et/ou l'exploitation d'espèces halieutiques du littoral Corse.

Nos travaux de recherche s'appuient sur un constat qui a été fait par les professionnels de la pêche et qui nous a conduit à lister des espèces présentant un grand intérêt économique et qui sont le plus menacées. Pour le moment, nous ne savons pas quels facteurs (et dans quelle proportion) ont engendré cette mutation de la biodiversité marine : pêche abusive, changement climatique, intrusion de nouvelles espèces, etc.

Partant de là, nous avons proposé de développer, en partenariat étroit avec la Société d'Informatique et de TElématique de la Corse (SITEC), un "Système d'Information dédié au suivi d'espèces halieutiques du littoral corse". La modélisation comportementale d'espèces marines passe tout d'abord par un suivi in situ. L'objectif est de bâtir un système de collecte d'informations basé sur les réseaux, intégrant in fine le temps réel, permettant de fournir les données nécessaires à l'élaboration et la validation d'un modèle de simulation comportementale. Il est envisagé dans cette partie la mise en œuvre de réseaux de capteurs permettant dans un premier temps le tracking par géo-localisation, puis la prise en compte d'autres données propres à l'espèce (mode de chasse, proies prélevées -quantité et espèces-, zones d'activité en fonction de son âge, de la période annuelle, etc.…). Plus précisément, nous serons amenés à prélever, doter d'un équipement et relâcher des individus de grande taille des espèces sélectionnées afin de les suivre dans des profondeurs allant de 0 à 100 mètres. Cet objectif nécessitera la définition et éventuellement le prototypage de passerelles matérielles et logicielles en les différents milieux (marin, aérien), afin de permettre d'intégrer le suivi en temps réel.

Dans un premier temps, nos travaux de suivi s'organiseront autour de quatre espèces : l'oursin (*Paracentrotus lividus*), le denté commun (*Dentex dentex*), le homard (*Homarus gammarus*) et l'araignée de mer (*Maja squinado*). Le choix de ces espèces, respectivement de

la famille des échinodermes, des sparidés et des crustacés, s'est fait en étroite collaboration avec les professionnels (pêcheurs, aquaculteurs), sur des espèces présentes sur le littoral corse et qui ont un intérêt majeur au niveau de l'exploitation mais également de la sauvegarde. Le but de cette analyse comportementale étant de comprendre et de déterminer les facteurs impactant le milieu marin.

Les différentes données recueillies au sein de ce projet – au près des pêcheurs et issues du suivi des espèces – seront couplées aux données des biologistes et permettront de développer différents modèles comportementaux. A l'Université de Corse, nous utilisons déjà ce type d'outils dans différents projets comme la résolution d'équations différentielles à coefficients imprécis, la détection de fautes dans des circuits électroniques, la prise en compte des retards dans les circuits électroniques, la propagation de feux de forêt à partir d'une représentation cellulaire ou la réalisation d'un contrôleur pour la régulation de la température d'une serre. Nous nous basons sur DEVS (abréviation de Discret EVent system Specification) qui est un formalisme modulaire et hiérarchique destiné à la modélisation et à la simulation de systèmes à événements discrets. Ce formalisme a été proposé publiquement par Bernard P. Zeigler en 1976 dans son ouvrage : Theory of Modeling and Simulation. Chaque modèle comportemental créé pourra interagir avec d'autres et aussi donner une description précise du milieu étudié. Par exemple, une fois que nous aurons développé le modèle comportemental de l'oursin, nous pourrons, en faisant varier différents paramètres ayant un impact sur l'espèce (température de l'eau, quantités de phytoplancton, le nombre de prédateurs, etc.), observer par simulation son développement. Le but étant de trouver le niveau optimal de chaque paramètre afin que toutes les espèces se développent au mieux.

L'objectif de notre projet de modélisation comportementale est double :

- Premièrement, de collecter un maximum d'informations sur des espèces et leur milieu afin de pouvoir mettre en place un outil d'aide à la décision pour les professionnels de la mer.
- Deuxièmement, grâce aux compétences dans le domaine des réseaux de capteurs de l'Université de Corse et de la SITEC, de développer de nouveaux outils de suivi d'espèces halieutiques. L'objectif étant d'arriver rapidement à un suivi d'individus en temps réel.

1. Travaux en cours

Les nombreuses études comportementales d'espèces marines menées ces dernières années ont conduit à un développement rapide de nouvelles applications pour la télémétrie acoustique

[108] [109] [110]. Afin de mieux connaitre le cycle de vie des espèces aquatiques, l'utilisation des ressources et son rôle écologique dans un écosystème, il est essentiel de suivre avec précision leurs mouvements à plus ou moins long terme [110] [111].

Les capteurs acoustiques sont utilisés pour le suivi de poissons depuis de nombreuses années. Une des premières études de suivi réalisée par télémétrie acoustique a été faite en 1956 par la National Marine Fisheries Service sur des saumons chinook adultes (Oncorhynchus tshawytscha) et des saumons coho (Oncorhynchus kisutch) [112].

Depuis, les études visant à décrire les mouvements des poissons dans leur milieu naturel se sont considérablement développées, grâce à l'essor de la télémesure, c'est à dire la transmission à distance d'un signal porteur d'un résultat de mesure, qui s'est appuyée sur certaines avancées technologiques de l'électronique, en particulier la miniaturisation et les améliorations de la puissance d'émission et de l'autonomie des émetteurs.

Au cours des trente dernières années, les fournisseurs de télémétrie acoustique ont développé différentes approches en matière de systèmes de positionnement haute résolution sur le medium acoustique. Le suivi actif et la surveillance acoustique passive sont devenus de puissants outils pour quantifier les mouvements complexes à différentes échelles spatiales et temporelles (par exemple [113] [114] [115] et [110]). Bien que la surveillance active puisse fournir des informations de mouvement précis sur quelques individus dans une courte période, la surveillance passive est utilisée pour quantifier l'information dans le plus long terme (données de présence-absence) de plusieurs individus. Bien moins cher en temps et de ressources, la surveillance passive ne permet pas l'acquisition de données et le suivi de position aussi précis que le suivi actif [115] [116].

La technique du tracking actif consiste à suivre en direct une espèce équipée d'un émetteur à l'aide d'un hydrophone embarqué sur un bateau. Les variations de puissance du signal reçu (en dB) nous donnent des indications quant à la position de l'espèce (plus le signal est fort et plus le sujet est proche). Le suivi passif nécessite le déploiement d'un ensemble d'hydrophones statiques et immergés. Chaque hydrophone peut rester plusieurs mois en mer et enregistre tous les signaux émis à proximité. Il faut enduite télécharger les données sur une machine pour pouvoir les traiter. Cette méthode est utilisée pour des études de migration, la détection présence/absence d'individus et le déplacement en 3D [117].

Parmi les principaux fournisseurs de matériels de suivi, nous pouvons citer entre autre VEMCO[9], Lotek[10], HTI[11] et Sonotronics[12]. Chacun de ces fournisseurs fournit un ensemble

[9] vemco.com

d'émetteurs acoustiques, de taille et de puissance d'émission différentes, une solution de tracking actif et un dispositif de monitoring passif. Parmi les nombreux suivis d'espèces marines réalisés en utilisant ce type de matériel, nous pouvons citer ces quelques travaux qui concernent des espèces identiques ou très proches de celles que nous comptons étudier [118] [119] [120] [121] [122] [123] [124].

1.1. Suivi des araignées de mer

1.1.1. Matériel utilisé

Le suivi a été réalisé à partir d'un bateau équipé d'un hydrophone omnidirectionnel VH165 (VEMCO Ltd, Novia Scotia) et un récepteur VR100 (VEMCO Ltd, Novia Scotia) pour localiser et suivre les émetteurs acoustiques.

En mer, la précision de la méthode manuelle de suivi dépend de plusieurs facteurs. Tout d'abord, la profondeur de l'émetteur et la présence/absence d'une thermocline qui agit comme une barrière acoustique. Deuxièmement, la végétation, les rochers, les vagues (dans des conditions météorologiques sévères) et d'autres obstacles et interférences qui peuvent réduire la puissance du signal. Sur le même site d'étude, il a été estimé que la précision varie de 10 m à 50 m.

Chaque position GPS, prise manuellement, a été enregistrée avec la date, l'heure, l'ID de l'émetteur (identifiant unique dans le monde, attribué par le constructeur) et la puissance du signal. Les coordonnées ont été transformées en coordonnées géodésiques (WGS84) et interprétées comme un thème dans l'environnement QGIS[13].

Les émetteurs utilisés dans cette étude sont des émetteurs codés V9 et V13 de chez VEMCO. Ces émetteurs envoient une série de pings acoustiques à 69 kHz à une cadence aléatoire dans un délai moyen de 60 secondes. Cette salve de pings comprend un numéro d'identification qui permet l'identification de l'étiquette spécifique avec les données de télémesure et de détection. Le V13 est équipé de capteurs de température et de profondeur. Les caractéristiques physiques de ces émetteurs sont :

[10] www.lotek.com

[11] www.htisonar.com

[12] www.sonotronics.com

[13] www.qgis.org

	V13TP-1L	V9-1L
Diamètre (mm) :	13	9
Longueur (mm) :	45	24
Poids dans l'air (g) :	12	3,6
Poids dans l'eau (g) :	6	2,2
Puissance de sortie (dB re 1µPa @ 1m) :	150	146

Figure 24 : caractéristiques d'émetteurs VEMCO utilisés pour l'étude

Les émetteurs de positionnement acoustique ne dépassent pas 1,1% du poids du corps du poisson (tableau 1).

1.1.2. Informations sur l'espèce : Maja squinado

L'araignée de mer *Maja squinado* est avec le *Cancer pagurus* l'un des plus grands crabes des côtes européennes. Ce qui explique qu'il a été cité très tôt dans la littérature et par de nombreux auteurs s'intéressant à la vie marine. Les premiers auteurs qui ont traité de l'araignée d'une manière précise sont Rondelet en 1554 et Aldrove en 1606. Herbst a été l'un des premiers à la décrire en détail en 1788.

L'idée générale qui se dégage de la littérature est celle d'un régime alimentaire omnivore. Ces animaux s'adaptent à l'environnement dans lequel ils se trouvent ([125] et [126]). Ce type de comportement semble approprié pour les animaux adultes qui ont à migrer sur de longues distances à travers différents biotopes. Cependant, il est possible qu'au stade de juvénile, l'environnement soit plus stable et le régime alimentaire soit plus restreint.

La littérature montre que l'on peut trouver le *Maja squinado* sur tous types de fonds. Plusieurs auteurs ont noté que, en raison de la migration, le substrat est différent selon le moment de l'année. Toutefois, les informations sur le lien entre le substrat et la saison ne sont pas concordantes entre les auteurs [127] :

- coralligène en hiver, vaseux en été [128] [129]
- rocheux en hiver, herbiers et fonds sableux au printemps [130]
- sableux ou vaseux tout au long de l'année, sauf en fin de printemps rocheux [131] [132]
- rocheux ou gravillonneux en hiver, herbier et sableux ou rocheux au printemps [133].

En ce qui concerne la profondeur de l'habitat de la *Maja squinado*, la plupart des auteurs donnent de 0 à 10 mètres. Les pêches les plus profondes citées dans la littérature sont 120m

en Bretagne Sud [134], 145m en Israël [135], 170m dans la mer Adriatique [136]et 600m en mer Ligure [137].

De nombreux auteurs ont noté, en Méditerranée et dans l'Atlantique, un profond changement selon la période de l'année en raison de la migration : les zones profondes en hiver, les zones côtières au printemps et en été.

1.1.3. Suivi des individus

Cette étude a eu lieu en mer Méditerranée, sur le littoral corse. Le suivi a été réalisé dans le cantonnement de pêche de San Fiurenzu (sur la côte ouest du Cap Corse: 42°45'N - 9°18'E) entre le 11 décembre 2012 et le 3 mai 2013.

Le but de ces travaux est de suivre l'araignée de mer dans son environnement et de collecter des informations sur son comportement. Les mouvements peuvent nous apporter différentes informations sur l'espèce comme sa vitesse de déplacement, son activité nycthémérale, ses zones de reproduction, les périodes de migration, etc.

Le tableau ci-dessous décrit les caractéristiques des individus étudiés.

	Sexe	**Taille (mm)**	**Poids (g)**	**REPEPADA (%)**	**Type d'émetteur**	**Temps du suivi (jours)**	**Date de largage**
Maja 1	femelle	168	1230	0,29	V9	1	11/12/12
Maja 2	mâle	154	1063	0,33	V9	2	11/12/12
Maja 3	femelle	178	1474	0,81	V13TP	6	11/03/13
Maja 4	femelle	156	1102	1,08	V13TP	5	12/03/13

REPEPADA : Rapport Entre le Poids de l'Emetteur et Poids de l'Araignée Dans l'Air

Figure 25 : caractéristiques des araignées capturées pour le suivi

La valeur de mesure utilisée dans cette étude est la longueur du céphalothorax (carapace), soit de l'entaille entre les deux pics ou des "cornes" du rostre au point le plus postérieur de la coquille, conformément à la règlementation CEE n°3094/86 du Conseil des Communautés Européennes du 7 octobre 1986 (Journal officiel des Communautés européennes n°L288 du 11/10/86).

Sur la figure suivante, on peut voir une araignée de mer (Maja 3 dans le tableau précédent) équipée d'un émetteur VEMCO V13TP.



Figure 26 : Maja 3 équipée d'un émetteur acoustique avant largage

Un premier suivi a débuté au mois de décembre. Nous avons relâché deux individus équipé d'émetteur V9. Le suivi n'aura duré qu'une journée car nous avons perdu leur signal le lendemain. Plusieurs raisons peuvent expliquer cette situation :

- les araignées se sont abritées sous une roche (qui fait obstacle au signal acoustique),
- elles se sont déplacées plus vite qu'on ne le pensait et ont quitté la zone de recherche,
- l'émetteur s'est arrêté d'émettre (disfonctionnement),
- elles ont été capturées ou mangées.

Ces conclusions nous ont amené à nous doter d'émetteurs plus puissants. C'est pourquoi nous avons opté pour des V13TP, qui en plus d'avoir une émission de signal plus importante, donne la profondeur à laquelle se trouve l'araignée, paramètre qui s'est avéré précieux lors de la recherche d'un individu en mer.

Nous pouvons noter que l'une des deux araignées (Maja 2) est revenue près de la côte et a été localisée le 03 mai 2013 (soit cinq mois après son largage).

Sur la Figure 27, nous pouvons observer l'ensemble des points enregistrés lors des suivis en mer. A chaque araignée est attribuée une couleur et le dégradé donne une indication sur la puissance du signal enregistré, plus la couleur est foncée et plus le signal est fort.

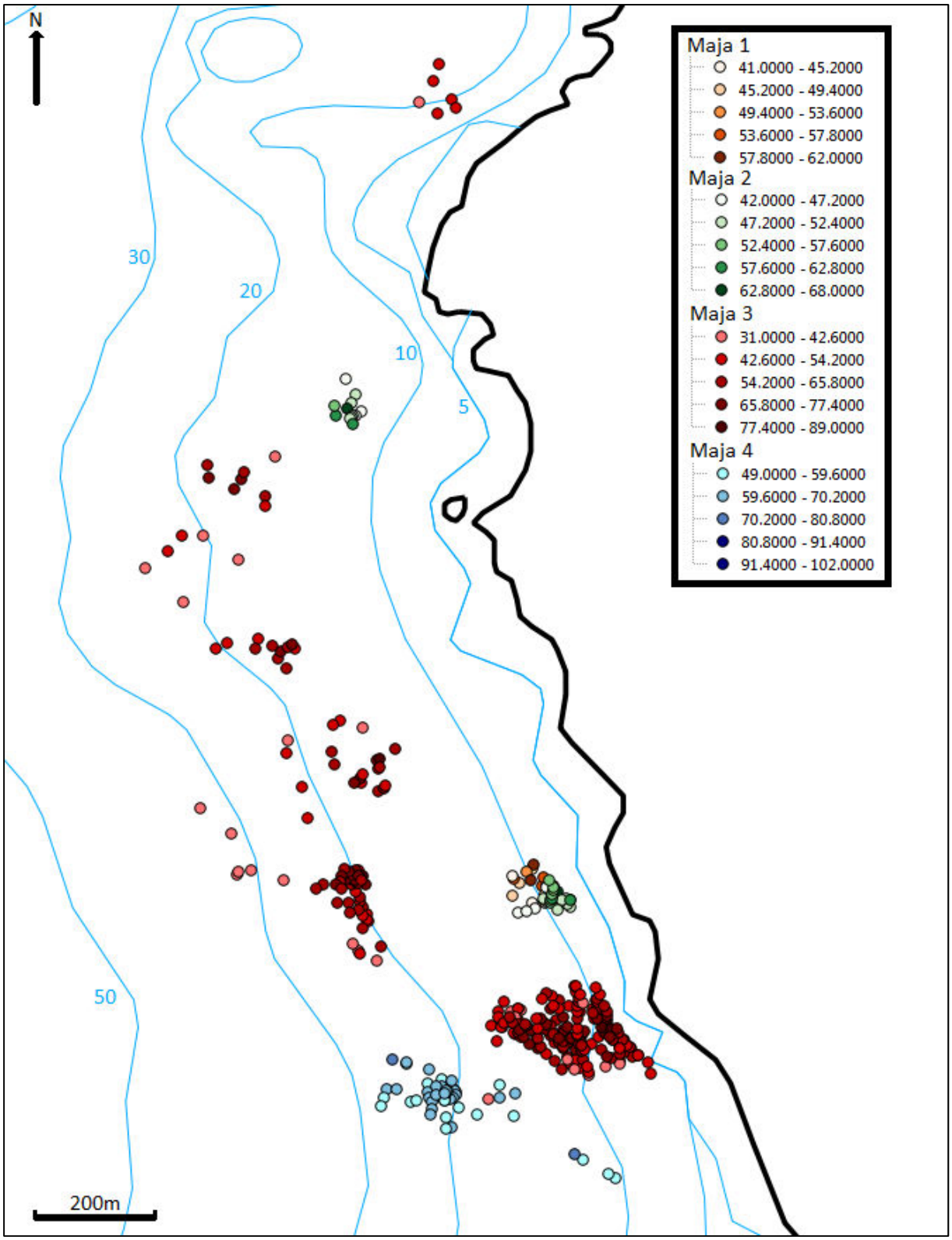


Figure 27 : ensemble des points de suivi des araignées de mer

Afin de visualiser le déplacement des *Maja squinado*, il faut relier les points dont la valeur du signal enregistrée est la plus forte. Par exemple, pour la Maja 3, on peut observer son

parcours en reliant entre eux tous les points ayant un signal capté supérieur à 70dB. Le résultat peut être observé sur la figure suivante.

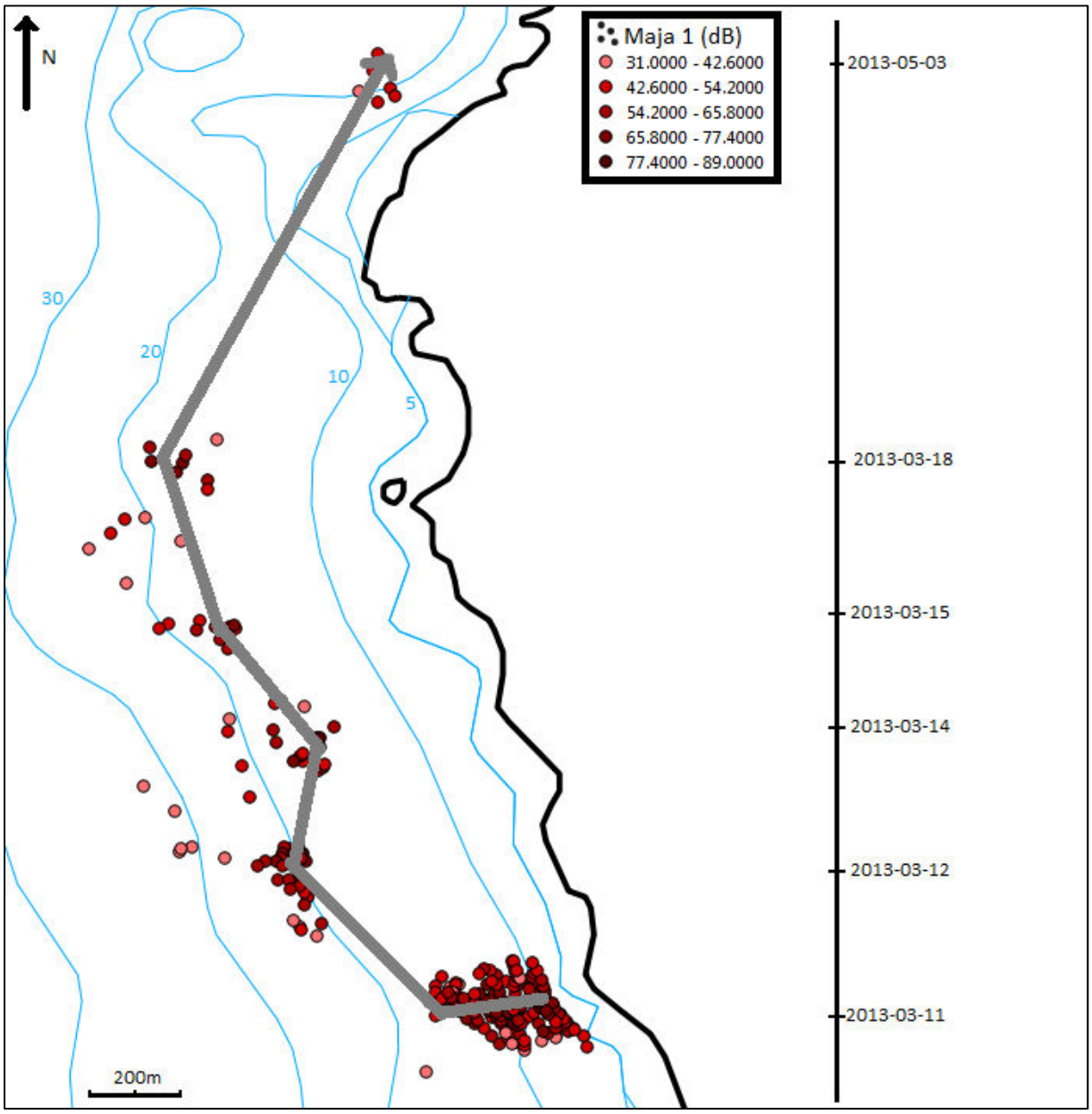


Figure 28 : trajectoire de Maja 3 pendant la campagne de suivi

Suite à ce suivi, nous pouvons observer que cette araignée de mer a suivi la ligne de bathymétrie des 20 mètres en direction du nord pour ensuite ce rapprocher de la côte (dans 10 mètres de fond) au mois de mai. Ce phénomène a également été observé pour la Maja 2 (Figure 27) qui était sortie de la zone de recherche pendant l'hiver (certainement pour aller vers des eaux plus profondes) et est revenue proche du littoral à la même date. Ce qui confirme les affirmations des professionnels de la mer et les quelques travaux effectués sur cette espèce [138]et [127]. Cette migration correspond avec la période de reproduction.

Le suivi de nuit montre également que la *Maja squinado* resterait plutôt sous roche la nuit. En effet, lors de tracking de nuit, nous avons pu observer que les signaux captés étaient bien plus faibles. Le graphique suivant fait apparaitre la moyenne des valeurs acoustiques enregistrées le jour et la nuit, ainsi que les valeurs maximum et minimum.

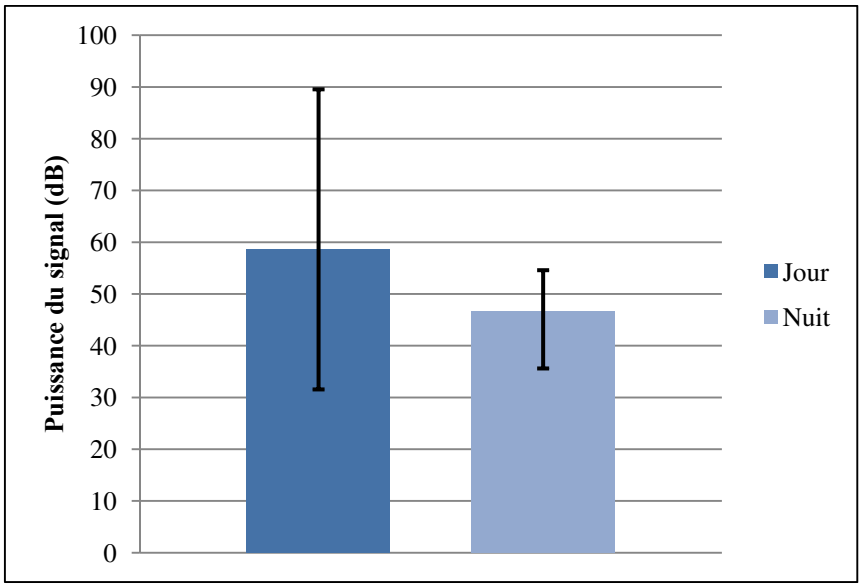


Figure 29 : comparaison entre les signaux enregistrés le jour et la nuit

1.1.4. Conclusion

L'étude de la *Maja squinado* apparaît comme un aspect stratégique en Méditerranée pour le développement de l'activité de pêche professionnelle et la relation avec son environnement. La plateforme STELLA MARE a pour objectif de faire sauter certains verrous scientifiques et permettre un développement harmonieux de la pêche. Ce suivi a porté sur l'étude de la *Maja squinado*, une araignée de mer, symbole de la déréglementation de l'activité de pêche. Les données sur cette espèce ne sont pas suffisantes et nous avons décidé de définir une base de données pour la construction d'un modèle de comportement. Suite à un suivi actif sur des individus équipés d'émetteurs VEMCO, nous avons recueilli des données sur ce comportement. Nous avons effectué onze jours de suivi sur une période de cinq mois et avons enregistré 390 positions. Nous avons observé un déplacement le long d'une ligne de profondeur de 20m pendant un mois et une remontée à 10m au mois de mai (qui correspond à la période de reproduction). Nous avons également observé que la *Maja squinado* reste dans son habitat pendant la nuit et s'est déplacée durant la journée. Maja 1 a parcouru environ 2,8km au cours de la période d'étude. Ces premiers résultats sont intéressants, mais ils ne sont pas suffisants et nous avons décidé de développer une méthode complémentaire.

1.2. Suivi d'espèces temps réel

D'après Albert C. Jensen [139] les premiers marquages et suivis de poissons remonteraient aux années 1860 avec le marquage de saumons et de truites en Ecosse et au Etats-Unis, dans le Maine. En 1895, un certain Hollister McGuire aurait marqué 5000 saumons de l'écloserie Clackamas et en aurait recapturé 32. Depuis les progrès de la science ont permis d'améliorer les suivis pour les rendre plus efficaces.

Le suivi d'espèces marines peut être réalisé de différentes façons. Les techniques les plus utilisées sont :

- Le marquage conventionnel (ou marquage visuel) qui consiste à fixer sur un individu une marque contenant un identifiant et un contact (numéro de téléphone ou adresse e-mail par exemple). Cette méthode de suivi « à moindre coût » permet de connaitre les points de largage et repêche d'un grand nombre d'individus. Cependant, elle ne permet pas l'acquisition de données précises sur le comportement d'une espèce, ni son parcours entre ses deux prises.
- Le marquage acoustique est certainement le plus intéressant. En effet, depuis la fin des années 60, les marques électroniques qui transmettent des signaux acoustiques pour suivre les déplacements des poissons individuels sont de plus en plus utilisées. Ces travaux ont permis de faire des progrès considérables dans la connaissance du comportement des espèces marines [140] [141] [142] [143] [144]. Cependant, cette technique est limitée dans la mesure où, dans la plupart des cas, on ne peut suivre qu'un seul poisson à la fois, chaque poisson ne peut être suivi que pendant une courte période (souvent quelques jours seulement) et le travail en mer réalisé à bord d'un bateau est coûteux.
- La marque-archive qui enregistre et stocke des données environnementales et comportementales. Pour ce suivi, aucune présence humaine n'est nécessaire, il est donc possible de surveiller le comportement et les déplacements d'un grand nombre de poissons en même temps pendant de longues périodes pouvant inclure une migration complète. La plupart des marques de stockage de données permettent de mesurer la pression (profondeur), la température (interne et externe), la lumière ambiante du jour, le cap au compas, la vitesse de nage, l'oxygène dissous ou l'activité trophique. Ces données environnementales peuvent être utilisées pour obtenir des informations sur le lieu et les déplacements des poissons [145] [146] [147].

La principale limitation de ces marques réside dans la nécessité de recapturer l'animal pour accéder aux données. Ceci implique le déploiement d'un nombre élevé de marques sur des

espèces ayant des taux d'exploitation élevés. De plus, il faut souvent attendre quelques années avant de récupérer un nombre significatif de données.

- La marque pop-up par satellite : afin d'augmenter la probabilité de récupérer les données, ces marques ont été créées. Elles sont externes et munies d'un mécanisme de libération qui permet à la marque de se détacher du poisson à un moment prédéterminé. Une fois à la surface de l'eau les données peuvent être récupérées via le système ARGOS à bord d'un satellite. La première génération de marques pop-up ne fournissait que la position de remontée à la surface et la température de la mer. Ces marques fournissent la distance parcourue en ligne droite par les poissons à partir du lieu de marquage. Les versions plus récentes des marques pop-up enregistrent la température, la profondeur et la lumière ambiante du jour. Lorsque la marque pop-up est physiquement récupérée, on peut récupérer les données complètes de profondeur et de température qui auront été enregistrées minute après minute.
- Le ping GPS : les marques satellites (généralement attachées à l'animal) sont utilisées pour étudier les déplacements à grande échelle et la physiologie des mammifères marins, des oiseaux et des tortues marines (voir [148]). Ces marques ont été déployées avec succès sur des requins pèlerins [149], mais elles ne sont applicables que sur des espèces qui fréquentent la surface.

Plusieurs problèmes se posent avec l'utilisation de tels marquages. Tout d'abord, le signal GPS ne peut pas être utilisé car les espèces que nous étudions ne remontent pas à la surface. Le suivi par marquage conventionnel ou par marque-archive ne permet pas d'obtenir un niveau de détail suffisant pour l'élaboration d'un modèle comportemental. Le tracking actif ne permet le suivi que de quelques individus à la fois pour des espèces benthos ou à faibles déplacements et seulement un poisson pélagique. De plus, ce suivi nécessite une présence humaine en mer en continu et devient vite couteuse.

Suite à ce constat, à un CCTP infructueux et plusieurs contacts avec des fournisseurs internationaux (par exemple VEMCO, leader mondial du domaine et NEREIDES[14] spécialiste dans les domaines de l'instrumentation navale, l'analyse de la qualité de l'eau et de l'air, l'océanographie, la météorologie et la mécanique marine), nous avons entrepris l'étude d'un système de bouées permettant de capter le signal émis par les émetteurs acoustiques positionnés sur les individus marins. Les données recueillies sont ensuite traitées et transmises (en quasi temps-réel) à un serveur au sein de l'Université. Le but étant, à partir des données

[14] http://www.nereides.fr/

captées par l'ensemble des bouées de trianguler une position précise de chaque individu afin d'obtenir un positionnement précis et en temps réel.

L'idée est de déployer un maillage de bouées-récepteurs où chaque module est autonome et indépendant. Il sera alors possible de déplacer chaque bouée afin que le maillage puisse être adapté aux déplacements des individus suivi. Le schéma ci-dessous détaille les différents composants nécessaires à la création de chaque module.

Figure 30 : description d'une bouée-récepteur

Chaque bouée devra être capable de capter le signal émis par un émetteur acoustique, de convertir ce signal analogique en numérique, d'identifier l'identifiant de l'émetteur source et de transmettre ces données à un serveur distant. Si à un instant *t* le signal émis par une même source est capté par un seul dispositif de réception nous aurons alors des données d'absence/présence de l'individu, si le signal est capté par deux récepteurs nous aurons une zone de présence (plus ou moins précise selon la puissance du signal reçu). Plus le maillage est important et plus il sera possible d'avoir un positionnement précis.

2. Conclusion et perspectives

Dans cette partie, nous avons présenté brièvement la plateforme de recherche marine de l'Université de Corse, STELLA MARE. Au sein de cette unité de nombreux programmes de recherches ont été entrepris dans les domaines de l'écologie, de la biologie des organismes, l'halieutique, l'aquaculture, la parasitologie, la modélisation et simulation informatique des systèmes complexes. Le programme de modélisation comportementale d'espèces halieutiques décrit ici consiste à élaborer un système d'information et d'étude dédié au suivi d'espèces marines du littoral corse et au développement d'un outil de modélisation et de simulation comportementale pour l'étude, la protection et/ou l'exploitation d'espèces ciblées.

C'est dans cette optique que nous avons débuté le suivi d'espèces marines par l'étude de la *Maja squinado*, une araignée de mer présente sur le littoral corse. Nous avons équipé quatre araignées d'émetteurs acoustiques, puis nous les avons relâchées (deux en décembre 2012 et deux en mars 2013) afin de suivre leurs déplacements. Ce suivi nous a permis d'obtenir des premiers résultats intéressants sur le comportement *in situ* de cette espèce et d'acquérir de nouvelles compétences dans le domaine du suivi d'espèces marines. Cette étude a été également l'occasion d'acquérir une première expérience du terrain, un premier contact avec une espèce vivante et des premiers travaux scientifiques effectués en mer. Il a fallu passer plusieurs heures en mer, souvent par mauvais temps et dans des conditions peu agréables. Ces travaux de suivi nous ont amené à réfléchir sur la création d'un nouveau dispositif de monitoring pour le positionnement en temps réel d'espèces halieutiques. Après une étude des technologies existantes et après avoir contacté plusieurs fournisseurs, nous avons décidé de créer un ensemble de module de type bouée-récepteur qui formerait en mer un maillage permettant de capter un signal acoustique et de positionner la source de l'émission.

Pour l'instant, nous allons réaliser les premiers tests sur la réception du signal acoustique et sur l'acquisition de ces données analogiques/numériques sur un système de stockage embarqué. Il reste encore de nombreuses contraintes liées à l'environnement marin qu'il faudra prendre en compte et qui sont autant de verrous scientifiques. Pour n'en citer que quelques uns, il faudra s'intéresser à la thermocline qui a l'effet d'une barrière acoustique. Il faudra prendre en compte dans la captation du signal des paramètres environnementaux comme la température de l'eau, sa salinité, sa turbidité et la pression qui agissent sur la transmission du signal. Il faudra également gérer la consommation énergétique du module. Il

est indispensable que chaque bouée soit autonome en énergie suffisamment longtemps afin de réaliser des campagnes de plusieurs jours.

Ce projet en est donc à ses premiers pas, mais il n'est pas difficile de mesurer l'intérêt de ces travaux dans le domaine de la recherche marine. Ce dispositif, facilement déployable en mer, pourra être utilisé pour de la surveillance du milieu en temps réel comme par exemple des paramètres physicochimiques (température, salinité, oxygène dissous, pH, turbidité), la pollution sonore sous-marine ou la surveillance de conduites ou canalisations immergées.

CHAPITRE VII :

Conclusion générale

L'objectif de nos travaux était d'élaborer une méthode logicielle permettant l'apprentissage de chants polyphoniques. Ce projet, mené en collaboration entre l'Università di Corsica, l'association E Voce di U Cumune et l'IRCAM, s'inscrit dans un projet universitaire de caractérisation et de valorisation du patrimoine culturel corse. Le chant polyphonique corse et plus particulièrement *a paghjella*, sont transmis de génération en génération par voix orale depuis plusieurs siècles. Tout en étant conscient que la complicité entre Hommes a une part importante dans la pratique du chant traditionnel, nous proposons de fournir un outil qui permet à une ou plusieurs personnes de réaliser des chants polyphoniques sans avoir besoin de connaitre la musique et sans maitriser les subtilités de la *paghjella*.

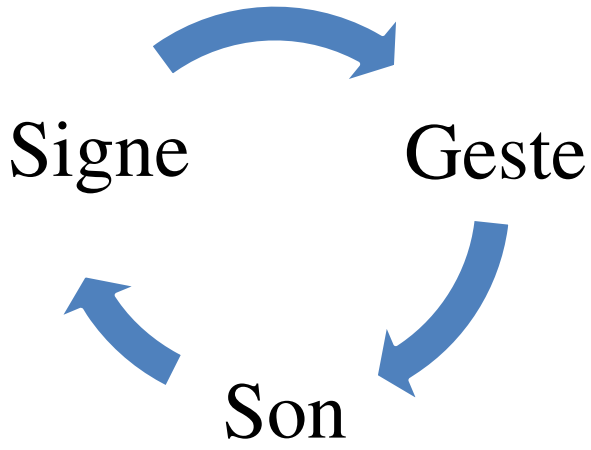


Figure 31 : Processus signe-geste-son

1. Bilan des travaux

La réalisation d'un tel outil, qui suit un processus qui va du signe au son (figure 31), s'est déroulée en plusieurs étapes :

- tout d'abord, nous avons validé le langage mimophonique créé par Nando Aquaviva. Quelques modifications mineures ont été apportées afin d'en faciliter l'apprentissage (nom des intervalles, ordre des signes, etc.). Le détail de ces travaux, auxquels j'ai participé, ne figure pas dans ce manuscrit car il fait parti d'une autre thèse dont la rédaction est en cours[15]. La mimophonie est un langage gestuel qui décrit l'échelle musicale corse. Elle se compose de 24

[15] Thèse en cours de rédaction par Claude Bellagamba

gestes auxquels est associé à chacun un nom, un signe (permettant une notation "en portée musicale") et une valeur en cents qui correspond à l'intervalle de note avec la tonique.

- une fois le langage maitrisé, nous avons cherché un moyen de capter ces gestes. De nombreuses technologies permettent de capter des gestes, mais les gants données semblaient être le meilleur support pour ce projet. Les contraintes de création des gants étaient: un faible coût, transportables, facilement déployables et permettant une grande liberté de mouvements. Nous avons donc créé une paire de gants équipés chacun de cinq capteurs de pression qui mesurent la flexion des doigts et d'un accéléromètre qui donne la position de la main dans un plan vertical. Les gants de données sont connectés à un ordinateur via Bluetooth et numérisent les gestes effectués pour qu'ils soient traités.

- l'utilisation de gants équipés de capteurs de flexion implique que les valeurs retournées pour un même geste peuvent varier selon l'utilisateur et la vitesse d'exécution. Il n'est donc pas possible de décrire chaque geste par un ensemble de valeurs fixes. Nous avons dû intégrer à notre système un module de reconnaissances. L'IRCAM a développé une librairie sous MAX/MSP qui permet de faire ce genre de traitement. En effet, la bibliothèque FTM possède de nombreuses fonctions utiles pour la reconnaissance. Dans notre projet, nous avons utilisé une fonction *mnm.mahalanobis* qui offre une manière de déterminer une similitude entre un ensemble de données connues et inconnues. Pour chaque geste nous avons donc un fichier contenant les six valeurs qui correspondent aux valeurs moyennes de chaque capteur pour ce geste. A chaque exécution d'un geste on va chercher, grâce à la distance de mahalanobis la série de valeurs la plus proche. Cette méthode est rapide et efficace mais nécessite au préalable que chaque geste soit "appris" par la machine. Cette étape peut être fastidieuse si le langage comporte de nombreux gestes. Nous avons également testé un système d'inférence flou. Ce système, basé sur l'utilisation de règles, est très utilisé lorsque des paramètres d'entrée ne sont pas représentés par des valeurs exactes mais par des fonctions. En représentant les valeurs émises par un capteur comme une probabilité de valeurs, nous parvenons à retrouver le geste voulu.

- afin de compléter notre système d'apprentissage de chants polyphoniques, nous devions associer à chaque geste un son. Nous avons intégré une version modifiée du module CHANT de l'IRCAM qui permet de générer des voyelles chantées à partir d'une analyse formantique d'une voix existante. Nous avons donc recréé cinq voyelles : a, è, i, o, ou (qui sont les cinq voyelles de la langue corse) dans le module CHANT qui permet de faire varier plusieurs paramètres de la voix comme notamment la fréquence fondamentale, le vibrato, le timbre, etc. Un autre moyen de produire des sons à partir des gestes est de coupler le module de

reconnaissance de gestes à un sampler contenant toutes les hauteurs de note enregistrées. Cette méthode offre un résultat plus réaliste, plus proche du naturel pour l'apprenant, mais nécessite que l'on enregistre toutes les voix à l'avance et permet de chanter uniquement à la tonalité enregistrée.

Les opportunités professionnelles et mes compétences acquises dans le domaine des capteurs sans fils m'ont permis d'orienter mes travaux dans un nouveau domaine de recherche scientifique. En effet, j'ai été recruté par une entreprise (la SITEC[16]) pour participer à un programme de recherche au sein de l'UMS CNRS 3514 STELLA MARE, laboratoire de recherche dans le milieu marin de l'Università di Corsica. Ces travaux consistent à modéliser le comportement d'espèces halieutiques afin de proposer des outils de gestion durable des ressources de la mer. Nous avons réalisé un suivi des plusieurs semaine sur une espèce d'araignée de mer (la *Maja squinado*), sur des homards et des oursins. Nous avons également entrepris le développement d'un système de suivi d'espèces marines temps-réel qui offrira aux chercheurs un outil adapté à l'étude comportementale.

2. Perspectives de recherche

Nos perspectives sont multiples. Dans le cadre de la mimophonie, nous pensons à court terme améliorer le système. L'aspect esthétique et ergonomique des gants de données n'a pas été réalisé. Nous avons créé des prototypes qui nous ont permis de valider des méthodes. Nous pourrons donc proposer une version adaptée à l'utilisation "publique".

Nous pouvons également améliorer notre système en y ajoutant plusieurs modules. Toujours à court terme, nous envisageons, en utilisant la bibliothèque SuperVP de l'IRCAM [150] [151], de permettre à l'utilisateur de réaliser des modifications sur le timbre et sur la fréquence fondamentale.

A moyen terme, il serait intéressant de faire évoluer le système de production sonore afin qu'il puisse modifier les voyelles chantées. Dans notre projet, le changement de voyelles peut se faire dans le module de synthèse vocale CHANT. Il est envisagé de réaliser un module de reconnaissance vocale qui repère la voyelle prononcée par le chanteur et choisisse la voyelle de synthèse équivalente. De nombreux travaux ont déjà été réalisés dans ce domaine [152]

[16] http://www.sitec.fr

[153], bien que les plus connus soient *Dragon NaturallySpeaking* de Nuance[17] et *Siri* de Apple[18].

Dans l'optique de développer des outils d'apprentissage de chants traditionnels, nous avons pensé développer un clavier numérique basé sur l'échelle musique de la mimophonie. A chaque touche de ce synthétiseur serait attribué un signe de la mimophonie auquel on associerait un intervalle de note par rapport à la tonique choisie. Cet outil permettra d'une part de se familiariser avec l'échelle musicale corse, ses micro-intervalles et ses harmoniques, d'autre part, il sera un bon outil de création pour des compositeurs. Il pourra également être utilisé pour lire des partitions écrites en utilisant ce langage, comme par exemple l'*Agnus Dei* d'Olmi Capella écrit par Nando Acquaviva (Figure 32).

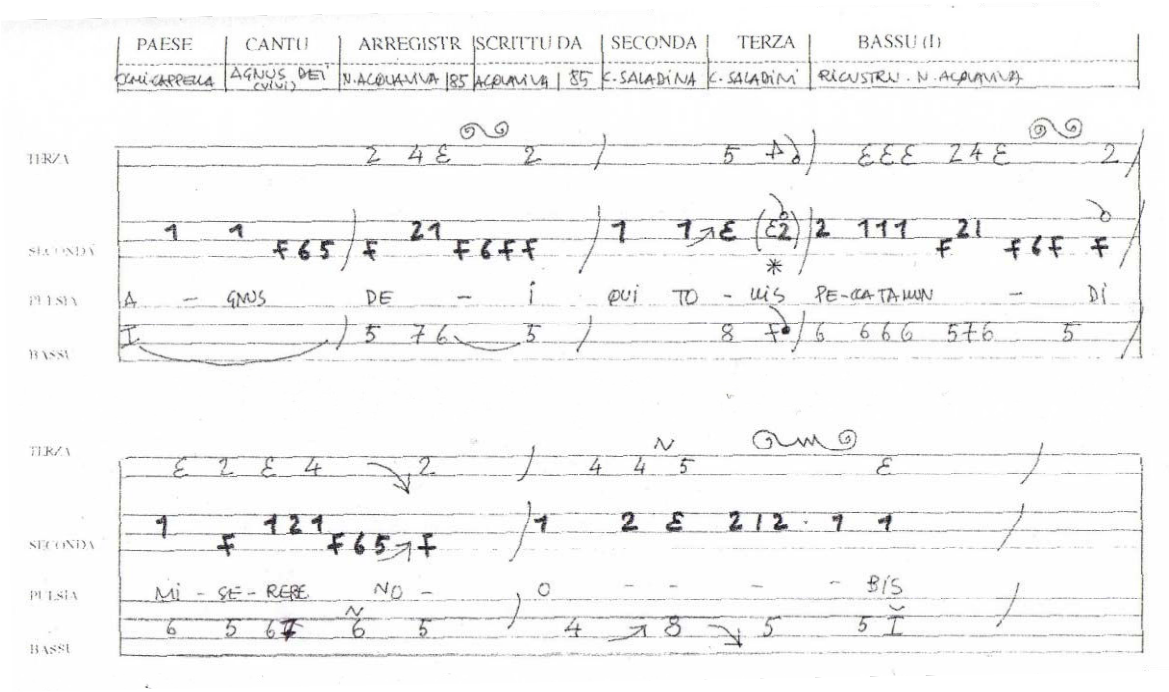


Figure 32 : Agnus Dei d'Olmi Capella

Dans le cadre des travaux menés à la plateforme STELLA MARE sur le programme de modélisation comportementale d'espèces halieutiques, nous envisageons à court terme de débuter le suivi de différentes espèces. Nous allons mener en parallèle le suivi de deux espèces : le homard commun et l'oursin violet.

Nous allons relâcher six homards équipés d'émetteurs acoustiques afin de pouvoir suivre leurs déplacements. Nous aurons au préalable déployé un maillage d'hydrophones

[17] http://www.nuance.fr/index.htm

[18] http://www.apple.com/fr/ios/siri/

immergés qui permettront d'enregistrer en continu les données émises par les émetteurs. Nous effectuerons également un tracking actif afin de suivre au mieux les individus marqués.

Pour le suivi des oursins, le marquage par tags acoustiques ne semble pas être une solution appropriée. En effet, il semblerait que les distances parcourues par une telle espèce soient de quelques centimètres par heures. Nous envisageons donc un suivi par capture d'images (vidéos et photos). Les oursins seront marqués d'un repère visuel qui permettra de les identifier et l'activité sera mesurée par rapport à un point fixe. Les premiers tests seront réalisés en bassin au sein de la plateforme.

Le suivi d'espèces halieutiques en temps-réel est un projet auquel nous attachons un intérêt plus particulier. En effet, le développement d'un tel outil nous permettra d'effectuer nos suivis dans de meilleures conditions et d'obtenir des résultats plus précis. Ce projet se déroule en plusieurs étapes :

- Le déploiement des bouées, celles-ci doivent être capables de capter le signal acoustique puis le traiter afin de récupérer la fréquence d'émission et la puissance du signal émis.
- Le transfert des données, qui s'effectuera par liaison GSM à partir de modules embarqués dans la bouée.
- Le traitement serveur qui se fera à STELLA MARE. Nous devrons à partir des données récoltées par les bouées recalculer une position. Ceci sera possible si une émission est captée par plusieurs bouées.
- La gestion de l'énergie. Les structures larguées en mer seront autonomes en énergie. Il faudra donc déterminer quelle est la solution qui offrira une couverture énergétique optimale (batteries, panneaux solaires, éoliennes, etc.).

Bibliographie

[1] J.-S. Gualtieri, J.-F. Santucci et D. Federici, Singing voices synthesis and gestural control for mimophony, NISP'07, 2007.

[2] D. Cantella, Contributions aux recherches sur le chant corse, I polyphonies vocales et orgues, Centre d'Ethnologie Française, Musée des Arts et Traditions Populaires, E voce di u Cumune, 1992, p. p.51.

[3] C. Cadoz, Instrumental gesture and musical composition, Proceedings of the International Computer Music Conference, 1988.

[4] M. Wanderley, B. Vines, N. Middleton, C. McKay et W. Hatch, The musical significance of clarinetists' ancillary gestures : An exploration of the field, vol. 34(1), The Journal of New Music Research, 2005, pp. 97-113.

[5] S. Dahl et A. Friberg, Expressiveness of Musician's Body Movements in Performances on Marimba, vol. volume 2915, S. Verlag., Éd., Lecture Notes in Computer Science, 2004, pp. pages 479-486.

[6] S. Dahl, F. Bevilacqua, R. Bresin, M. Clayton, K. Hansen, L. Leante, I. Poggi, N. Rasamimanana et M. Wanderley, Musical Gestures. Sound, Movement, and meaning, Musical Gestures. Sound, Movement, and meaning éd., Leman, M. and Godøy, R. I., (Eds.), 2009, p. 36–68.

[7] C. Cadoz et M. Wanderley, Gesture-music, Trends in Gestural Control of Music, Ircam, Éd., Paris, 2000, p. 1–55.

[8] D. P. Wanderley M., Controle gestuel de la synthese sonore, Interfaces Homme-Machine et Création Musicale, Paris: Hermès Science Publications, 1999, p. 145–164.

[9] I. Choi, Gestural primitives and the context for computational processing in an interactive performance system, Trends in Gestural Control of Music, Ircam, 2000.

[10] I. Choi, A component model of gestural primitive throughput, 2003.

[11] F. Delalande, La gestique de glenngould, Louise Courteau Editrice, 1988, p. 84–111.

[12] S. Canazza, G. De Poli, S. Rinaldin et A. Vidolin, Sonological analysis of clarinet expressivity, Leman, 1997, p. 431–440.

[13] S. Canazza, G. De Poli et A. Vidolin, Perceptual analysis of the musical expressive intention in a clarinet performance, Leman, 1997, p. 441–450.

[14] M. Waiswisz, The hands, a set of remote midi-controllers, In Proc. of the International Computer Music Conference, 1985.

[15] S. Jordà, M. Kaltenbrunner, G. Geiger et R. Bencina, The reactable, In Proceedings of the International Computer Music Conference, 2005.

[16] A. Tanaka et B. Knapp, Multimodal interaction in music using the electromyogram and relative position sensing, In Proceedings of the International Computer Music Conference, 2002.

[17] S. Goto, The aesthetics and technological aspects of virtual musical instruments : The case of the superpolm midi violin, Leonardo Music Journal, 9, 1999, pp. 115-120.

[18] R. Koehly, M. Wanderley et D. Curtil, Paper fsrs and latex/fabric traction sensors : methods for the development of home-made touch sensors., In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)., 2006.

[19] A. Hunt et M. Wanderley, Mapping performer parameters to synthesis engines, Organized Sound, 7(2), 2002.

[20] A. Hunt, M. Wanderley et M. Paradis, The importance of parameter mapping in electronic instrument design, The Journal of New Music Research, 32(4), 2003.

[21] D. Arfib, J. Couturier, L. Kessous et V. Verfaille, Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces, vol. 7(2), Organized Sound, 2002, pp. 127-144.

[22] V. Verfaille, M. Wanderley et P. Depalle, Mapping strategies for gestural and adaptive control of digital audio effects, vol. 35(1), The Journal of New Music Research, 2006, pp. 71-93.

[23] D. Van Nort, M. Wanderley et P. Depalle, On the choice of mappings based on geometric properties, In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), 2004.

[24] C. Dobrian et F. Bevilacqua, Gestural control of music using the vicon 8 motin capture system, In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME), 2003.

[25] E. Rank, A player model for midi control of synthetic bowed strings, Diderot Forum on

Mathematics and Music, 1999.

[26] S. Serafin, R. Dudas, M. Wanderley et X. Rodet, Gestural control of a real-time physical model of a bowed string instrument, In Proceedings of the International Computer Music Conference, 1999.

[27] N. Orio, N. Schnell et M. Wanderley, Evaluation of input devices for musical expression : Borrowing tools from hci, vol. 26(3), The Computer Music Journal, 2001, pp. 62-76.

[28] J. Malloch, D. Birnbaum, E. Sinyor et M. Wanderley, Towards a new conceptual framework for digital musical instruments, In Proceedings of the Digital Audio Effects Conference (DAFx), 2006.

[29] E. Schoonderwaldt, K. Guettler et A. Askenfelt, Effect of the width of the bow hair on the violin string spectrum, In Proceedings of the Stockholm Music Acoustics Conference (SMAC), 2003.

[30] E. Schoonderwaldt, K. Guettler et A. Askenfelt, An empirical investigation of bow-force limits in the schelleng diagram., 2007.

[31] K. Guettler, E. Schoonderwaldt et A. Askenfelt, Bow speed or bowing position - which one inuences spectrum the most ?, In Proceedings of the Stockholm Music Acoustics Conference (SMAC)., 2003.

[32] A. Askenfelt, Observations on the transient components of the piano tone, vol. 79, In Proceedings of the Stockholm Music Acoustics Conference (SMAC), 1993, pp. 297-301.

[33] M. Barthet, R. Kronland-Martinet et S. Ystad, Consistency of timbre patterns in expressive music performance., Proceedings of the Digital Audio Effects Conference (DAFx)., 2006.

[34] S. Dahl et A. Friberg, Visual perception of expressiveness in musicians' body movements, vol. 24(5), Music Perception, 2007, pp. 433-454.

[35] M. Wanderley, B. Vines, N. Middleton, C. McKay et W. Hatch, The musical significance of clarinetists' ancillary gestures : An exploration of the field., vol. 34(1), The Journal of New Music Research, 2005, pp. 97-113.

[36] C. Traube et P. Depalle, Phonetic gestures underlying guitar timbre description, In Proceedings of the International Conference on Music Perception and Cognition, 2004.

[37] A. Camurri, G. De Poli, A. Friberg, M. Leman et G. Volpe, The mega project : analysis and synthesis of multisensory expressive gesture in performing art applications, vol. 34(1), The Journal of New Music Research, 2005, pp. 5-21.

[38] M. Mancini, R. Bresin et C. Pelachaud, A virtual head driven by music expressivity, vol. 15(6), IEEE Transactions on Audio, Speech and Language Processing, 2007.

[39] L. Turner-Stokes et K. Reid, Three-dimensional motion analysis of upper limb movement in the bowing arm of string-playing musicians, Clinical Biomechanics, 1998.

[40] H. Winold, E. Thelen et B. D. Ulrich, Coordination and control in the bow arm movements of highly skilled cellists, vol. 6(1), Ecological Psychology, 1994, pp. 1-31.

[41] A. P. Baader, O. Kazennikov et M. Wiesendanger, Coordination of bowing and fingering in violin playing, Vols. %1 sur %223(2-3), Cognitive brain research, 2005, pp. 436-443.

[42] F. Styns, L. van Noorden et M. Leman, Some basic observations on how people move on music and how they relate music to movement, Proceedings of the International Music and Gesture Conference, 2006.

[43] M. M. Wanderley et P. Depalle, Gestural control of sound synthesis, vol. 92, Proceedings of the IEEE, 2004.

[44] C. Traube, P. Depalle et M. Wanderley, Indirect acquisition of instrumental gesture based on signal, physical and perceptual information, Singapore, National University of Singapore: NIME ’03: Proceedings of the 2003 conference on New interfaces for musical expression, 2003, p. 42–47.

[45] A. Massoni, Les musiques de Corse, Chants, instruments et danse, Tradition Vivante, Alain Piazzola, Bonchamp-Lès-laval, 2006, p. 23.

[46] C. Bellagamba, Fiumarella, fola musicale, libru di u maestru, E Voce di u Cumune, C.R.D.P., 2004, pp. 17-20.

[47] N. Casalonga et A. Massoni, Le chant religieux corse, sous la direction de Marcel Pérès, Les cahiers du CERIMM, éditions CREAPHIS, 1996, p. 83.

[48] T. Baudel et A. Braffort, Reconnaissance de gestes en environnement reel, Montpellier: Proceedings of Montpellier "93, lnterface to Real and Virtual Worlds, 1993.

[49] M. Krueger, Artificial Reality, vol. 2, Addison-Wesley, Reading, 1990.

[50] R. Bolt, The Human Interface, New York: Van Nostrand Reinhold, 1984.

[51] J. Caelen, Interaction multi-modale dans ICPdraw: expérience et perspectives, Annales de l'école IHM multi-modale, GRECO-PRC communication Homme-Machine, 1991.

[52] K. Thorisson, D. Koons et R. Bolt, Multi-Modal Natural Dialogue, CHI'92 Conference Proceedings, ACM Press, 1992, pp. 653-654.

[53] K. Murakami et H. Taguchi, Gesture Recognition Using Recurrent Neural Networks, CHI'91 Conference Proceedings, ACM Press, 1991, pp. 237-242.

[54] D. Sturman, Whole-Hand Input, Massachusetts Institute of Technology: Ph.D. thesis, Media Arts & Sciences, 1992.

[55] H. Morita, S. Hashimoto et S. A. Ohteru, Computer Music System that Follows a Human Conductor, IEEE Computer, 1991, pp. 44-53.

[56] P. Depalle, S. Tassart et W. M., Instruments virtuels, Résonance n° 12, 1997.

[57] L. Turner-Stokes et K. Reid, Three-dimensional motion analysis of upper limb movement in the bowing arm of string-playing musicians, Clinical Biomechanics, 1998.

[58] M. Bonnet, Y. Guiard, J. Requin et S. A., Mécanismes généraux de la motricité, Paris : Presses Universitaires de France: Traité de Psychologie Expérimentale, 1994.

[59] S. Boukir et F. Cheneviere, Conception d'un système de reconnaissance de gestes dansés, vol. 21, G. d. r. e. d. d. t. d. s. e. d. images, Éd., Saint-Martin d'Hères, 2004, pp. 195-203.

[60] D. Berndt et C. J., Using dynamic time warping to find patterns in time series, Workshop on Knowledge Discovery in Databases, 1994, pp. 359-370.

[61] E. Keogh et M. Pazzani, Scaling up dynamic time warping to massive datasets, Prague, Czech Republic: Proceedings of the 3rd European Conference on Principles of Data Mining and Knowledge Discovery (PKDD), 1999, pp. 1-11.

[62] J. Weismann et R. Saloman, Gesture recognition for virtual reality applications using data glove and neural networks, Washington DC: IEEE Int. JointConf. on Neural Networks, 1999, pp. 2043-2046.

[63] M. Yang et N. Ahuja, Recognizing hand gesture using motion trajectories, Ft. Collins: CVPR 99, IEEE Conf. on Computer Vision and Pattern Recognition, 1999, pp. 466-472.

[64] L. Campbell, Invariant features for 3-D gesture recognition, Int. Conf. on Automatic Face and Gesture Recognition, 1996, pp. 157-162.

[65] F. Chenevière, S. Boukir et E. Beets, Compression et reconnaissance de séquences spatio-temporelles issues d'un ballet contemporain, Gérardmer, Éd., ORASIS'2003, 9ème Congrès Francophone en Vision par Ordinateur, 2003, pp. 353-361.

[66] F. Chenevière, S. Boukir et B. Vachon, A HMM-based dance gesture recognition system, Manchester: IWSSIP 02, 9th Int. Workshop on Systems,Signals and Image Processing, 2002, pp. 322-326.

[67] A. Psarrou, S. Gong et M. Walter, Recognition of human gestures and behaviour based on motion trajectories, Image and Vision Computing, 2002, pp. 349-358.

[68] L. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, vol. 77(2), Proceedings of the IEEE, 1989, pp. 257-286.

[69] T. Starner et A. Pentland, Visual recognition of american sign language using hidden markov models, Zurich, Switzerland: Int. Workshop on AutomaticFace and Gesture Recognition, 1995, pp. 189-194.

[70] C. Vogler et D. Metaxas, A framework for recognizing the simultaneous aspects of american sign language, vol. 81(3), Computer Vision and Image Understanding, 2001, pp. 358-384.

[71] H. Yoon, J. Soh, Y. Bae et H. Yang, Hand gesture recognitionusing combined features of location, angle and velocity, vol. 34, PatternRecognition, 2001, pp. 1491-1501.

[72] N. Szilas, Les réseaux récurrents supervisés : une revue critique, Institut National Polytechnique de Grenoble: Rapport de recherche 972-I, 1997.

[73] L. Campbell et A. Bobick, Recognition of human body motion using phase space constraints, Cambridge MA: ICCV 95, 5th Int. Conf. on Computer Vision, 1995, pp. 624-630.

[74] J.-S. Gualtieri, -P. Bisgambiglia, J.-F. Santucci et D. Federici, Reconnaissance de gestes par Systèmes d'Inférences Flous, vol. 30, TSI, Hermes-Lavoisier, 2011, pp. 571-590.

[75] P. Glorennec, Algorithmes d'apprentissage pour systèmes d'inférence floue, Hermès Science, 1999.

[76] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, vol. 65, Psychological Review, 1958, pp. 386-408.

[77] E. Mamdani et S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, vol. 7, International Journal of Man-Machine Studies, 1975, pp. 1-13.

[78] T. Takagi et M. Sugeno, Fuzzy indentification of systems and its applications to modeling and control, IEEE Trans. on S.M.C., 1985.

[79] L. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, parts 1 and 2, 1975, p. 199–249 and 301–357.

[80] L. Zadeh, The role of fuzzy logic in the management of uncertainty in expert systems, Fuzzy Sets and Systems, 1983, p. 199–227.

[81] B. Bouchon-Meunier, Logique floue et ses applications, Addison-Wesley, 1995.

[82] L. Zadeh, Outline of a new approach to the analysis of complex systems and decision process, IEEE Trans. Systems, Man and Cybernetics, 1973, p. 28–44.

[83] L. Zadeh, G. Klir et B. Yuan, Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems, Advances in Fuzzy Systems, vol. 6, Application and Theory, 1995.

[84] L. Zadeh, Fuzzy sets, vol. 8, Information and Control, 1965, pp. 338-353.

[85] A. Kaufmann, Introduction à la théorie des sous-ensemble flous, Masson, 1973.

[86] Pena-Reyes, Fuzzy modeling, Ph.D. thesis, 2002.

[87] M. Sugeno, Industrial applications of fuzzy control, Elsevier Science Pub. Co., 1985.

[88] M. Puckette, Combining Event and Signal Processing in the MAX Graphical Programming Environment, CMJ, 1991.

[89] N. Orio, S. Lemouton, D. Schwarz et N. Schnell, Score Following: State of the Art and New Developments, Montreal: NIME, 2003.

[90] F. Bevilacqua, R. Muller et N. Schnell., MnM: a Max/MSP Mapping Toolbox, Vancouver: NIME, 2005.

[91] F. Bevilacqua, F. Guédy, N. Schnell, E. Fléty et N. Leroy, Wireless sensor interface and gesture-follower for music pedagogy, New York: Proceedings of the International Conference of New Interfaces for Musical Expression (NIME 07), 2007, pp. 124-129.

[92] J. Yang et Y. Xu., Hidden markov model for gesture recognition, 1994.

[93] A. Wilson et A. Bobick, Learning visual behavior for gesture analysis, IEEE, 1995.

[94] A. F. Bobick et A. D. Wilson, A state-based approach to the representation and recognition of gesture, IEEE Trans. Pattern Analysis and Machine Intelligence, 1997.

[95] A. Wilson et A. Bobick, Parametric hidden markov models for gesture recognition, IEEE Trans. Pattern Analysis and Machine Intelligence, 1999.

[96] A. Wilson et A. Bobick, Nonlinear phmms for the interpretation of parametrized

gesture, 1998.

[97] A. Wilson, Adaptative Models for the Recognition of Human Gesture, PhD thesis, Massachusetts Institute Of Technology, 2000.

[98] X. Ge et P. Smyth, Deformable markov model templates for timeseries pattern matching, University of California, Irvine: Technical report, Departement of Information and Computer Science, 2000.

[99] X. Ge et S. P., Hidden markov models for endpoint detection in plasma etch processes, University of California, Irvine: Technical report, Departement of Information and Computer Science, 2001.

[100] C. Decoux, Estimation de modèles de semi-chaînes de markov cachées par echantillonnage de gibbs, Rev. Statistique Appliquée, 1997.

[101] H. Lei et V. Govindaraju, Regression time warping for similarity measure of sequence, IEEE, 2004.

[102] Y. Yanagisawa et T. Saoh, Clustering multidimensional trajectories based on shape and velocity, IEEE, 2006.

[103] D. Chudova, S. Gaffney et P. Smyth, Probabilistic models for joint clustering and time-warping of multidimensional curves.

[104] H. Kaprykowsky et H. Rodet, Globally optimal short-time dynamic time warping application to score to audio alignment, IEEE, 2006.

[105] F. Bevilacqua, R. Muller et N. Schnell, MnM: a Max/MSP mapping toolbox, Proceedings of the conference on New interfaces for musical expression, 2005.

[106] N. Schnell, R. Borghesi et D. Schwarz, FTM complex data structures for max, Barcelona: Proc. of the ICMC, 2005.

[107] N. Schnell et D. Schwarz, Gabor, multi-representation real-time analysis/synthesis, COST-G6 Conference on Digital Audio Effects (DAFx), 2005.

[108] X. Rodet, Y. Potard et J. Barrière, CHANT - de la synthèse de la voix chantée à la synthèse en général, Ircam - Centre Georges-Pompidou: Rapport Ircam 35/85, 1985 .

[109] A. Klimley, F. Voegeli, S. Beavers et B. Le Boeuf, Automated listening stations for tagged marine fishes, MTS J. 32, 1998, p. 94–101.

[110] F. Voegeli, M. Smale, D. Webber, Y. Andrade et R. O'Dor, Ultrasonic telemetry, tracking and automated monitoring technology for sharks, Environ. Biol. Fish. 60,

2001, p. 267–281.

[111] C. Lowe et R. Bray, Fish movement and activity patterns, University of California Press, California: Allen, L.G., Horn, M.H., Pondella, D.J. (Eds.), The Ecology of California Marine Fishes, 2006, p. 524–553.

[112] C. Lowe, D. Topping, D. Cartamil et Y. Papastamatiou, Movement patterns, home range, and habitat utilization of adult kelp bass Paralabrax clathratus in a temperate no-take marine reserve, Mar. Ecol. Prog. Ser. 256, 2003, p. 205–216.

[113] P. S. Trefethen, Sonic equipment for tracking individual fish, Washington, DC.: Special Scientific Report – Fisheries Number 179, United States Fish and Wildlife Service, 1956.

[114] S. Clements, D. Jepsen et M. Karnowski, Optimization of an automated telemetry array for detecting transmitter-implanted fish, N. Am. J. Fish. Manage. 25, 2005, p. 429–436.

[115] M. Domeier, Methods for the deployment and maintenance of an acoustic tag tracking array: an example from California's Channel Island, Mar. Technol. Soc. J. 39, 2005, p. 74–80.

[116] M. Heupel, J. Semmens et A. Hobday, Automated acoustic tracking of aquatic animals: scales, design, and deployment of listening station arrays, Mar. Fresh. Res. 57, 2006, p. 1–13.

[117] B. Hartill, M. Morrison, M. Smith, J. Boubée et D. Parsons, Diurnal and tidal movement of snapper (Pagrus auratus, Saparidae) in an estuarine environment, Mar. Fresh. Res. 54, 2003, p. 931–940.

[118] M. Espinoza, T. Farrugia, D. Webber et F. L. C. Smith, Testing a new acoustic telemetry technique to quantify long-term, fine-scale movements of aquatic animals, Fisheries Research 108, 2011, pp. 364-371.

[119] M. Espinoza, T. Farrugia et C. Lowe, Habitat use, movements and site fidelity of the gray smooth-hound shark (Mustelus californicus Gill 1863) in a newly restored southern California estuary, Journal of Experimental Marine Biology & Ecology 401: 6, 2011.

[120] L. A. Comeau, S. E. Campana et M. Castonguay, Automated monitoring of a large-scale cod (Gadus morhua) migration in the open sea, vol. 59 Issue 12, Canadian Journal of Fisheries & Aquatic Sciences, 2002, p. 1845.

[121] L. Comeau, R. Sonier et J. Hanson, Seasonal movements of Atlantic rock crab (Cancer

irroratus Say) transplanted into a mussel aquaculture site, Aquaculture Research, 2011, pp. 1-9.

[122] C. Bernárdez, E. González-Gurriarán, B. García-Calvo, A. Corgos et J. Freire, Movements of juvenile and adult spider crab (Maja squinado) in the Ría da Coruña, A. t. a. a. a. P. o. t. F. C. o. F. T. h. i. Europe, Éd., Ustica, Italia: Spedicato, M.T.; Lembo, G.; Marmulla, G. (eds.), 2005, pp. 9-13.

[123] R. Bradford, B. Bruce, R. McAuley et G. Robinson, An Evaluation of Passive Acoustic Monitoring Using Satellite Communication Technology for Near Real-Time Detection of Tagged Animals in a Marine Setting, The Open Fish Science Journal, 2011, pp. 10-20.

[124] C. Jadot, A. Donnay, M. L. Acolas, Y. Cornet et M. Bégout Anras, Activity patterns, home-range size, and habitat utilization of Sarpa salpa (Teleostei: Sparidae) in the Mediterranean Sea, ICES Journal of Marine Science, 63, 2006, pp. 128-139.

[125] C. Dumont, J. Himmelman et M. Russell, Size-specific movement of green sea urchins Strongylocentrotus droebachiensis on urchin barrens in eastern Canada, vol. 276, Marine Ecology Progress Series, 2004, p. 93–101.

[126] Z. Stevcic, Relations interspécifiques de l'Araignée de mer, Rapp. Comm. int. Mer Médit. 19(2), 1968, pp. 147-149.

[127] G. Kergariou, Régime alimentaire de Maia squinado, CIEM. Comité des crustacés, coquillages et benthos CM 1974/K;36. 6 p., 1974.

[128] D. Le Foll, Biologie et exploitation de l'araignée de mer Maja squinado Herbst en Manche Ouest, Brest: Thèse de Doctorat de l'Université de Bretagne Occidentale, 1993.

[129] G. Olivi, Zoologica Adriatica, Bossano. i-xxxii + 334 p. + 9 PL, 1792.

[130] L. Stalio, Catalogo metodico e descrittivo dei crostacei dell'Adriatico, Parte I, Atti dell'Ist. Veneto 5 (3), 1877, pp. 355-385.

[131] R. Bourdon, Inventaire de la faune marine de Roscoff. Décapodes - Stomatopodes, Trav. Stat. Biol.Roscoff16, 1965, pp. 1-45.

[132] Z. Stevcic, A short outline of the biology of the spinous spider crab, Bull. Sci. Conseil Acad. RSF Yougoslavie, Section A 12(11-12), 1967, pp. 313-314.

[133] Z. Stevcic, L'Araignée de mer et le substrat, Rapp. Comm. int. Mer Médit. 19(2), 1968, pp. 151-153.

[134] G. Kergariou, L'araignée de mer, Maia squinado L., sur le littoral de Bretagne, Science et Pêche, Bull. Inst. Pêches marit. 205, 1971, pp. 11-19.

[135] P. Camus, Résultats d'une opération de marquage d'araignée de mer (Maia squinado, Herbst) adulte en baie d'Audierne (Bretagne sud), CIEM. Comité des mollusques et crustacés CM 1983/K:29.11 p., 1983.

[136] L. Holthuis et E. Gottlieb, An annotated list of the Decapoda Crustacea of Mediterranean coast of Israël, with an appendix listing the Decapoda of the Eastern Mediterranean, Bull. Res. Counc. of Israël 7B(l-2), 1958, pp. 1-126.

[137] Z. Stevcic, Odnos rakovice prema dubini. I rapporti tra la grancevola e la profondita. Spiny spider crab in relation to the depth, Thalassia Jugoslavica 5, 1969, pp. 352-360.

[138] L. Holthuis, Vrais crabes, Rome: Fischer, W. M.-L. Bauchot et M. Schneider (rédacteurs), Fiches FAO d'identification des espèces pour les besoins de la pêche. (Révision 1). Méditerranée et Mer Noire. Zone dépêche 31. Vol. I. Végétaux et Invertébrés, 1987, pp. 321-368.

[139] E. González-Gurriarán, J. Freire et C. Bernárdez, Migratory patterns of female spider crabs Maja squinado detected using electronic tags and telemetry, Journal of Crustacean Biology 22 (1), 2002, pp. 91-97 .

[140] A. Jensen, Marking and tagging fishes, Fishery Leaflet 534, 1962.

[141] H. YUEN, Behavior of skipjack tuna, Katsuwonus pelamis, as determined by tracking with ultrasonic devices, J. Fish. Res. Board Can., 27(11): 2071-9, 1970.

[142] F. CAREY et K. Lawson, Temperature regulation in free-swimming bluefin tuna, Comp. Biochem. Physiol. (A Comp. Physiol.), 44(2): 375-92., 1973.

[143] R. Laurs, H. Yuen et J. H. Johnson, Small-scale movements of albacore Thunnus alalunga in relation to ocean features as indicated by ultrasonic tracking and oceanographic sampling, Serv. NMFS, 75(2):347-55., 1977.

[144] F. Carey et B. Robinson, Daily patterns in the activities of swordfish, Xiphias gladius, observed by acoustic telemetry, Fish. Bull. NOAA/NMFS, 79(2):277-92., 1981.

[145] R. Brill, D. Holts, R. Chang, S. Sullivan, H. Dewar et F. Carey, Vertical and horizontal movements of striped marlin (Tetrapturus audax) near the Hawaiian Islands, determined by ultrasonic telemetry, with simultaneous measurement of oceanic currents, Mar. Bio. 117:567–574., 1993.

[146] R. Hill, Theory of geolocation by light levels, Berkley, CA: Le Bouef, B.J., and Laws. R.M. (eds.), Elephant Seals: Population Ecology, Behavior, and Physiology, University of California Press, 1994, pp. 227-236.

[147] J. Gunn, Smart archival tag comes up trumps for tuna, Australian Fisheries. 53: 10-11, 1994.

[148] J. Metcalfe, Summary report of a workshop on daylight measurements for geolocation in animal telemetry, Dordrecht, The Netherlands: “Electronic Tagging and Tracking in Marine Fisheries” Reviews: Methods and Technologies in Fish Biology and Fisheries, 2001, pp. 331-342.

[149] B. Block, H. Dewar, C. Farwell et E. Prince, A new satellite technology for tracking the movements of Atlantic bluefin tuna, Proceedings National Academy Sciences USA, 95, 1998, pp. 9384-9389.

[150] I. Priede, A basking shark (Cetorhinus maximus) tracked by satellite together with simultaneous remote- sensing, Fish. Res. 2, 1984, pp. 201-216.

[151] A. Roebel et X. Rodet, Efficient Spectral Envelope Estimation and its application to pitch shifting and envelope preservation, Madrid: International Conference on Digital Audio Effects., 2005, pp. 30-35.

[152] A. Roebel, F. Villavicencio et X. Rodet, On Cepstral and All-Pole based Spectral Envelope Modeling with unknown Model order, vol. 28, Pattern Recognition Letters, 2007, pp. 1343-1350.

[153] M. Fanty, E. Barnard et A. Cole, Alphabet recognition, Handbook of Neural Computation, 1995.

[154] S. Furui, Cepstral analysis technique for automatic speaker verification, IEEE Transactions on Acoustics, Speech and Signal Processing, 1981.

Liste des figures

FIGURE 1 : PROCESSUS "DU GESTE AU SON" 10
FIGURE 2 : ECHELLE MUSICALE CORSE 18
FIGURE 3 : CORRESPONDANCE GESTE-SIGNE 20
FIGURE 4 : UN DES DEUX GANTS DE DONNÉES 25
FIGURE 5 : TRAME DES CAPTEURS NUMÉRISÉE ET ENVOYÉE SUR LE PORT SÉRIE 27
FIGURE 6 : BOITIER DE RÉCEPTION 29
FIGURE 7 : EXEMPLE DE SOUS ENSEMBLE FLOU 37
FIGURE 8 : EXEMPLE DE SIF 38
FIGURE 9 : D'UN DES SOUS ENSEMBLES FLOUS D'ENTRÉE 40
FIGURE 10 : SOUS ENSEMBLES FLOUS POUR L'INDEX 41
FIGURE 11 : SEFS REPRÉSENTANT DES SORTIES 41
FIGURE 12 : ENSEMBLE DES RÈGLES POUR LE CHANTEUR DE SECONDE 42
FIGURE 13 : EXEMPLE DE RÉSULTATS DES TESTS DE RECONNAISSANCE SUR HUIT GESTES 43
FIGURE 14 : SYNTHÈSE DE VOIX BASÉE SUR FOF 51
FIGURE 15 : PATCH PRINCIPAL DU MODULE CHANT 52
FIGURE 16 : DESCRIPTION FORMANTIQUE POUR LA SYNTHÈSE DE LA VOYELLE 'I' 53
FIGURE 17 : LES FORMANTS QUI APPARAISSENT SUR LE SPECTRE D'UNE VOYELLE CHANTÉE 54
FIGURE 18 : PROCESSUS DU GESTE AU SON 57
FIGURE 19 : MODULE DE RÉCEPTION DES DONNÉES ISSUES DES GANTS 58
FIGURE 20 : MODULE DE RECONNAISSANCE DES GESTES PROVENANT DES GANTS DE DONNÉES 59
FIGURE 21 : LA QUALITÉ DE LA RECONNAISSANCE D'UN GESTE EST LIÉE AU NOMBRE D'ENREGISTREMENTS RÉALISÉS 60
FIGURE 22 : MODULE DE SAMPLER 61
FIGURE 23 : MODULE DE LECTURE D'UN SON 61
FIGURE 24 : CARACTÉRISTIQUES D'ÉMETTEURS VEMCO UTILISÉS POUR L'ÉTUDE 68
FIGURE 25 : CARACTÉRISTIQUES DES ARAIGNÉES CAPTURÉES POUR LE SUIVI 69
FIGURE 26 : MAJA 3 ÉQUIPÉE D'UN ÉMETTEUR ACOUSTIQUE AVANT LARGAGE 70
FIGURE 27 : ENSEMBLE DES POINTS DE SUIVI DES ARAIGNÉES DE MER 71
FIGURE 28 : TRAJECTOIRE DE MAJA 3 PENDANT LA CAMPAGNE DE SUIVI 72
FIGURE 29 : COMPARAISON ENTRE LES SIGNAUX ENREGISTRÉS LE JOUR ET LA NUIT 73
FIGURE 30 : DESCRIPTION D'UNE BOUÉE-RÉCEPTEUR 76
FIGURE 31 : PROCESSUS SIGNE-GESTE-SON 80
FIGURE 32 : AGNUS DEI D'OLMI CAPELLA 83