

**COGOLUEGNES Charles**

**DUHAMEL Alexy**

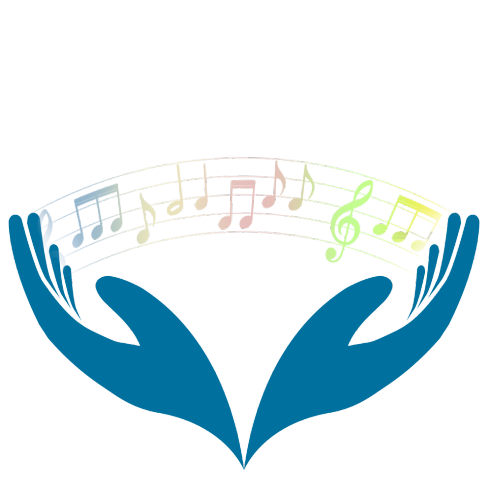
**HAMON Maxime**

**YVINEC Romann**

Année universitaire 2019-2020

**MASTER 1 INFORMATIQUE**

Rapport de projet



**CONDUCT**

Responsable d’UE : RAIMBAULT Frederic

Tutrice : GIBET Sylvie

Remerciements

Nous tenons tout d’abord à remercier notre tutrice, Sylvie Gibet, pour la proposition et le suivi de notre projet. Merci à Frederic Raimbault pour ses retours pertinents sur notre avancement. Nous sommes reconnaissants envers Tiago Brizolara, doctorant dans l’équipe expression d’Irisa, pour nous avoir donné un cours sur le logiciel PureData. Nous tenons à remercier Axel Jacomme pour avoir été présent au Virtual Lab. Merci à Charlotte Pelletier pour son aide sur les algorithmes de reconnaissance des gestes. Pour finir, nous remercions Mansour Tchenegon, actuellement étudiant en M2 Informatique, pour nous avoir aidé à démarrer le projet, en s’appuyant sur celui de l’an passé (SONIC 2).

Résumé

Notre projet s’intitule CONDUCT. Son but est de permettre de contrôler un son ou une mélodie avec les gestes des mains. Les différentes caractéristiques contrôlables sont le volume, le tempo, la fréquence et le vibrato. Le suivi des gestes est possible grâce à l’appareil LeapMotion, ainsi que le logiciel Unity pour le retour visuel et la calibration. Pour finir, la synthèse de son est effectué avec le logiciel PureData et l’instrumentalisation se fait avec le langage Faust.

**MOT CLÉS : Reconnaissance de geste, synthèse de son, réalité virtuelle.**

Summary

Our project is called CONDUCT. It aims to control a sound or a melody with hands gesture. The features we can control are the volume, the tempo, the frequency and the vibrato. The tracking of movements is possible thanks to a device named LeapMotion, and the software Unity for the visual and calibration feedback. Finally, the sound synthesis is made with the software PureData and the instrumentalization is made with the Faust language.

**KEYWORDS: Gesture recognition, sound synthesis, virtual reality.**

Table des matières

[1. Contexte 1](#_Toc40199327)

[1.1. Équipe de recherche 1](#_Toc40199328)

[1.2. Thématique 1](#_Toc40199329)

[1.2.1. Analyse synthèse et reconnaissance de gestes expressifs 1](#_Toc40199330)

[1.2.2. Synthèse de parole expressive 1](#_Toc40199331)

[1.2.3. Expressivité dans les textes, fouilles de textes 2](#_Toc40199332)

[2. Travail demandé 3](#_Toc40199333)

[2.1. Cahier des charges 3](#_Toc40199334)

[2.2. Spécification du sujet 3](#_Toc40199335)

[3. Travail réalisé 4](#_Toc40199336)

[3.1. Problèmes rencontrés 4](#_Toc40199337)

[3.1.1. Reconnaissance de configurations manuelles : Unity et Leap Motion 4](#_Toc40199338)

[3.1.2. Synthèse sonore sous PureData 4](#_Toc40199339)

[3.1.3. Interface visuelle 6](#_Toc40199340)

[3.2. Solutions retenues 6](#_Toc40199341)

[3.3. Choix effectués 6](#_Toc40199342)

[3.4. Produit réalisé 7](#_Toc40199343)

[4. Glossaire 8](#_Toc40199344)

[5. Table des figures 9](#_Toc40199345)

[6. Bibliographie 10](#_Toc40199346)

# Contexte

## Équipe de recherche

L'Équipe de recherche *Expression* est basée en Bretagne, située dans les villes de Vannes, Lannion et Lorient. Elle est composée d’une cinquantaine de membres dont **Sylvie Gibet**, notre tutrice de projet.

Figure - Logo équipe Expression, laboratoire Irisa

Cette équipe a pour grande objectif d’étudier le langage humain, que ça soit le geste, la parole ou le texte. Les grandes thématiques que l’équipe aborde sont « l’analyse synthèse et reconnaissance de gestes expressifs », la « synthèse de parole expressive » et l'Expressivité dans les textes, fouilles de textes », nous allons maintenant détailler ces thématiques.

## Thématique

### Analyse synthèse et reconnaissance de gestes expressifs

Cette thématique vise à tirer des caractéristiques de hauts niveaux des gestes, à partir de caractéristiques de bas niveau. On cherche à déterminer la signification d’un geste à partir de données brutes capturées grâce à des capteurs utilisant des technologies de pointes. On pourra utiliser cette technologie dans les domaines de l’art, du jeu ou encore de la santé.

### Synthèse de parole expressive

Dans ce cas, nous cherchons à régler le problème de la synthèse de parole classique, le manque d’expression de ces synthèses. L'équipe cherche donc à extraire d’un texte « l’expression » afin de l’utiliser pour synthétiser une voix expressive. On va ensuite chercher à évaluer ces synthèses. Les champs d’applications de la synthèse de parole expressive sont les médias, l’assistance ou encore des cas de la vie de tous les jours, tel l’interaction avec des objets connectés.

### Expressivité dans les textes, fouilles de textes

On va ici chercher à extraire le sens dans les textes, afin de pouvoir mieux les comprendre, les appréhender, rendre l’implicite explicite. L’objectif de tout cela est de rendre la compréhension des textes plus aisées, de prendre des décisions plus rapidement.

Trois parties sont à distinguer dans cet objectif, le premier est d’acquérir les textes qui vont nous servir de sources, tout en le filtrant afin d’éliminer le bruit, ensuite nous allons extraire les caractéristiques de ces textes, extraire les informations dont nous avons besoin, puis pour finir, nous voulons représenter les résultats afin de les mettre en lumière, et les rendre utilisables.

# Travail demandé

## Cahier des charges

[Attirez l’attention du lecteur avec une citation du document ou utilisez cet espace pour mettre en valeur un point clé. Pour placer cette zone de texte n’importe où sur la page, faites-la simplement glisser.]

## Spécification du sujet

[Attirez l’attention du lecteur avec une citation du document ou utilisez cet espace pour mettre en valeur un point clé. Pour placer cette zone de texte n’importe où sur la page, faites-la simplement glisser.]

# Travail réalisé

## Problèmes rencontrés

### Reconnaissance de configurations manuelles : Unity et Leap Motion

### Synthèse sonore sous PureData

Le premier obstacle rencontré était de pouvoir jouer une mélodie et de générer du son à l’aide de Pure Data. La première option était de jouer d’un instrument synthétique et d’envoyer les notes jouées à Pure Data. Bien qu’il aurait été possible de le faire, cela aurait nécessité d’avoir deux utilisateurs : un pour jouer de l’instrument, et l’autre pour modifier les paramètres à l’aide du LeapMotion. De plus, il aurait été nécessaire de disposer d’un synthétiseur ou d’un logiciel permettant de jouer d’un instrument. La seconde option était de faire lire un fichier à Pure Data afin qu’il joue la mélodie, solution que nous avons retenue. Le format que nous avons retenu était le format MIDI. D’une part parce que ce format est prévu pour être utilisé dans le domaine musical, et d’autre part car la librairie “Cyclone” nous mettait à disposition tous les outils dont nous avions besoin pour pouvoir utiliser les fichiers MIDI au sein de Pure Data.

La seconde difficulté rencontrée était de modifier le tempo. Ce paramètre était probablement celui qui était le plus difficile à modifier. Nous avions 3 approches à notre disposition. La première était de relancer la lecture du fichier MIDI à chaque changement de tempo. Bien qu’il s’agît de la solution la plus simple, c’était aussi celle qui correspondait le moins à nos attentes. En effet, nous cherchions à changer le tempo en cours de lecture, et non de redémarrer la lecture à chaque changement. La seconde solution était de stocker toutes les notes dans un tableau, et d’ensuite lire ce tableau à la vitesse voulu à l’aide d’un métronome. Bien que cela nous aurait permis de changer le tempo en cours de lecture, cette approche était loin d’être satisfaisante. La première lecture du fichier ne permettait pas d'émettre du son. Bien que ce problème ne soit que peu dérangeant pour des mélodies courtes, il devient bien plus handicapant pour des mélodies plus longue, vu que l’utilisateur n’aurait aucun retour lors de cette première lecture. De plus, cette méthode aurait drastiquement changé la mélodie, vu qu’il nous aurait été impossible de conserver, par exemple, les accords, ou le temps entre deux notes. La troisième approche, qui est celle que nous avons retenu, était de prétraiter l’information reçu puis d’envoyer cette information à l’objet « Seq » de la librairie « Cyclone ». En effet, cet objet, que nous utilisions pour la lecture de fichier MIDI, permet de changer le tempo en cours d’exécution. Bien que cette méthode fût la plus complexe à mettre en place, dû à la façon dont cet objet permet de modifier le tempo, elle était aussi la plus satisfaisante car elle permettait de changer le tempo en cours d’exécution tout en ne modifiant pas la mélodie.

Une autre difficulté était de choisir les paramètres à modifier. Tout d’abord, nous devions limiter leur nombre. Sachant que nous devions assigner un geste à chaque paramètre, un nombre trop important de paramètre aurait forcé l’utilisateur à se rappeler de trop de geste. Nous nous sommes donc arrêtés à 4 paramètres, car cela offre à l’utilisateur un contrôle suffisant sur la musique tout en limitant le nombre de geste qu’il doit retenir.

Après avoir décidé du nombre de paramètres, nous devions encore choisir les paramètres en eux-mêmes. Avec un nombre aussi limité, il était nécessaire de sélectionner les paramètres les plus pertinents, donc ceux ayant le plus d’effets sur le son. Les paramètres sont les suivants :

* Le volume
* Le tempo
* L’attaque
* Le vibrato (ce paramètre n’a un impact que sur le violon)

Au cours du développement, nous avions remarqués que l’attaque (qui était un paramètre des instruments généré avec Faust) n'avait aucun effet sur le son, peu importe sa valeur. Nous avons par conséquent décidé de le remplacer par le pitch (ou, fréquence).

### Interface visuelle

## Solutions retenues

[Attirez l’attention du lecteur avec une citation du document ou utilisez cet espace pour mettre en valeur un point clé. Pour placer cette zone de texte n’importe où sur la page, faites-la simplement glisser.]

## Choix effectués

Figure - Fonctionnement de l'application

## Produit réalisé

[Attirez l’attention du lecteur avec une citation du document ou utilisez cet espace pour mettre en valeur un point clé. Pour placer cette zone de texte n’importe où sur la page, faites-la simplement glisser.]

# Glossaire

MIDI : une interface de communication entre synthétiseur. Elle est principalement utilisée dans le monde de la musique, permettant l’échange d’information entre différents appareils. Les messages au format MIDI contiennent les informations sur les notes joué, permettant ainsi à un synthétiseur de recréer la mélodie. Les fichiers MIDI, que nous utilisons dans notre projet, sont une suite de messages au format MIDI. [10]

PureData : un logiciel de programmation visuel. Il est principalement utilisé dans le domaine de la musique ou du traitement de sons. [5]

Faust : Faust est un langage de programmation fonctionnel axé sur la synthèse audio. Il permet notamment la création de synthétiseurs et d’instruments. [6]

Algorithme K-NN : un algorithme d’apprentissage supervisé. Nous avons un jeu de données composées de m données de dimension Rd. Afin de déterminer la nature d’une nouvelle donnée (par exemple une position de main). Nous allons calculer la distance entre notre donnée et les données du jeu de données. La classe la plus présente parmi les k plus proches voisins de notre donnée dans l'espace sera la classe de notre donnée.

# Table des figures

[Figure 1 - Logo équipe Expression, laboratoire Irisa 1](file:///C:\Users\Megaport\Desktop\CharlesPerso\M1_INFO\INF2212_ProjetTutore\CONDUCT\Documentation\rapport.docx#_Toc40290174)

[Figure 2 - Fonctionnement de l'application 7](file:///C:\Users\Megaport\Desktop\CharlesPerso\M1_INFO\INF2212_ProjetTutore\CONDUCT\Documentation\rapport.docx#_Toc40290175)

# Bibliographie

[1] A. Bouënard, S. Gibet, M. Wanderley (2012). Hybrid Inverse Motion Control for Virtual Characters Interacting with Sound Synthesis - Application to percussion motion. The Visual Computer 28(4): 357-370, 2012

[2] Lei Chen, Sylvie Gibet. CONDUCT: an expressive conducting gesture corpus for sound control, LREC 2018, Japon, May 2018

[3] *Larousse*[en ligne][consulté le 8 mai 2020] : <https://www.larousse.fr/dictionnaires/francais/synesth%C3%A9sie/76181>

[4] *Valve Developer Community* [en ligne][consulté le 8 mai 2020] :

<https://developer.valvesoftware.com/wiki/Decals>

[5] *Site de Pure Data* [en ligne] [consulté le 9 mai 2020] : <https://puredata.info/>

[6] *Site de Faust* [en ligne] [consulté le 9 mai 2020] : <https://faust.grame.fr/>

[7] *Dépôt GitHub de Cyclone* [en ligne] [consulté le 9 mai 2020] : <https://github.com/porres/pd-cyclone>

[8] *Site de Max* [en ligne] [consulté le 9 mai 2020] : <https://cycling74.com/products/max/>

[9] *Documentation de MSP* [en ligne] [consulté le 9 mai 2020] :

<https://docs.cycling74.com/max5/tutorials/msp-tut/mspaudioio.html>

[10] Jim Heckroth, Crystal Semiconductor Corp : The Complete MIDI 1.0 Detailed Specification,The MIDI Manufacturers Association, Avril 2006, <https://www.midi.org/specifications-old/item/the-midi-1-0-specification>

Logo des mains par vecteezy.com

ANNEXES