



HAL
open science

Architecture logicielle et matérielle d'un système de détection des émotions utilisant les signaux physiologiques. Application à la mnémothérapie musicale

Chaka Koné

► **To cite this version:**

Chaka Koné. Architecture logicielle et matérielle d'un système de détection des émotions utilisant les signaux physiologiques. Application à la mnémothérapie musicale. Electronique. Université Côte d'Azur, 2018. Français. NNT : 2018AZUR4042 . tel-01936711

HAL Id: tel-01936711

<https://theses.hal.science/tel-01936711>

Submitted on 27 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

Architecture Logicielle et Matérielle d'un Système de Détection des Emotions utilisant les Signaux Physiologiques. Application à la Mnémothérapie Musicale

Chaka KONÉ

Laboratoire d'Electronique Antennes et Télécommunications (LEAT) et le laboratoire
d'Informatique, Signaux et Systèmes à Sophia antipolis (I3S)

**Présentée en vue de l'obtention
du grade de docteur en** Electronique
d'Université Côte d'Azur

Dirigée par : Cécile Belleudy
et co-dirigée par Nhan Le-Thanh

Soutenue le : 01 juin 2018

Devant le jury, composé de :

Nacer Abouchi, Professeur , CPE Lyon

Ladjel Bellatreche, Professeur, ISAE-ENSMA

Cécile Belleudy, Maître de conférences HDR, Université Côte d'Azur

Mireille Blay, Professeur, Université Côte d'Azur

Johann Laurent, Maître de conférences HDR, Université de Bretagne

Nhan Le-Thanh, Professeur, Université Côte d'Azur

Résumé

Ce travail de thèse s'inscrit dans le domaine de l'informatique affective et plus précisément de l'intelligence artificielle et du domaine de l'exploration d'architecture. L'objectif de ce travail est de concevoir un système complet de détection des émotions en utilisant des capteurs de signaux physiologiques. Ce travail se place donc à l'intersection de l'informatique pour la définition d'algorithme de détection des émotions et de l'électronique pour l'élaboration d'une méthodologie d'exploration d'architecture pour la conception de nœuds de capteurs

Dans un premier temps, ce travail a permis de définir de nouveaux algorithmes de détection, multimodale et instantanée des émotions. Dans l'approche monomodale initialement développée, nous avons démontré que certains capteurs sont plus adaptés à détecter certaines émotions que d'autres. Aussi, dans l'approche multimodale, l'approche retenue consiste à utiliser plusieurs signaux physiologiques distincts et à les coupler avec les compétences de chaque type de signal à reconnaître une activité émotionnelle spécifique. Nous avons ainsi proposé une nouvelle méthode de fusion au niveau des décisions. Dans notre système, deux algorithmes de classification à savoir le k-nearest neighbors (kNN : k plus proches voisins) et le support vector machine (SVM : machine à vecteurs de support) ont été implantés et chacun d'entre eux permet d'obtenir un taux de reconnaissance des émotions supérieurs à 80%.

Afin de concevoir un tel système alimenté sur pile, nous avons étudié différentes architectures et proposé un modèle analytique d'estimation de la consommation d'énergie des systèmes d'objets communicants à un niveau haut d'abstraction. La précision de ce modèle a été prouvée en comparant l'autonomie estimée avec notre modèle et celle obtenue par des mesures sur une plateforme réelle. Afin de tenir compte des contraintes applicatives et utilisateurs, nous avons développé un environnement, en SystemC TLM, d'exploration d'architecture et de ces configurations. Ce dernier offre aux concepteurs d'objet communicant pour la santé un outil de conception et de simulation d'architecture. Il permet ainsi d'évaluer les performances et l'autonomie de ces systèmes avant leur conception. L'exploration d'architecture est basée sur une librairie de microcontrôleurs, de capteurs de différents types et de modules radio qui peuvent représenter différents protocoles de communication.

Pour atteindre l'objectif visé par cette thèse de Doctorat, nous avons utilisé cet outil d'exploration et de simulation pour proposer une architecture matérielle composée d'un ensemble de nœud capteur. Cette architecture permet de faire l'acquisition et la transmission des données physiologiques tout en satisfaisant les contraintes applicatives et utilisateurs. A l'aide de ces nœuds capteur, nous avons conçu une base de données de signaux physiologiques incluant 5 états émotionnels en utilisant l'approche de stimulation appelée Mnémothérapie musicale. Deux architectures logicielles ont été proposées (sur mobile et PC). Elles permettent de faire l'acquisition des données sur les nœuds conçus en garantissant l'ordonnancement de l'acquisition. Enfin, les deux architectures matérielle/logicielle ont été combinées pour donner le système Emotica (EMOTION CAPTURE). Emotica est donc un système complet de détection des émotions en utilisant des signaux physiologiques. Les résultats montrent que le système permet d'avoir d'excellent taux de reconnaissance des émotions tout en présentant une autonomie acceptable.

Mots clés : détection automatique des émotions, signaux physiologiques, fusion des signaux, approche multimodale, estimation de la consommation énergie, exploration d'architecture, nœuds de capteurs.

Abstract

Our work fits into the field of affective computing taking part of artificial intelligence and the field of design space exploration. The aim of this work is to design a complete emotions detection system using sensors of physiological signals. This work therefore stands at the intersection of computing for the definition of emotions detection algorithms and electronics for first the development of a methodology of architecture exploration and energy consumption estimation and then for the design of sensors nodes.

Initially, this work helped to define algorithms of multimodal and instantaneous of emotions detection. In the initial monomodal approach, we demonstrated that some sensors allow a better detection of certain emotions rather than others. In the multimodal approach, the novelty is to use several distinct physiological signals and coupled with the skills of each type of signal to recognize one emotional activity from others. Also We propose a new fusion method to the different levels of decisions. In our system, two classification algorithms namely the k-nearest neighbors (kNN) and support vector machine (SVM) have been implemented and allows everyone to recognize the emotions with recognition rates of above 80%.

To design such a battery powered system, we studied different architecture and proposed an analytical model for energy consumption estimation of communicating objects systems at high level of abstraction. The accuracy of this model has been proven by comparing the autonomy estimated with our developed model with measurements on a real platform. Then, we have developed a framework providing to healthcare communicating objects designers, a simulation environment for complete architecture allowing them to evaluate performance and autonomy of these systems before the real hardware design. It also offers a tool for architecture exploration for these kind communicating objects systems thanks to the fact that the proposed tool contains a library of different microcontrollers, sensors, radio modules.

Then, to achieve the objective of this doctoral thesis, we used this tool of exploration and simulation to propose a hardware architecture consisting of three nodes to make the acquisition and transmission of physiological signals by satisfying architectural and application constraints. Then, we designed a database of physiological signals including 5 emotional states by using the stimulation approach called musical Mnemotherapy. Two software architectures have been proposed (on mobile and PC). They allow to

acquire data on the nodes designed by ensuring the acquisition scheduling. Finally, both hardware/software architectures have been combined to give the Emotica (EMOTION CAPTURE) system. Emotica is a complete system of detection of emotions using physiological signals. The results show that the system allows to have excellent recognition rates by having a large autonomy.

Keywords : Automatic detection of emotions, physiological signals, signals fusion, multimodal approach, energy consumption estimation, design space exploration, sensor nodes.

Liste des publications

Conférences internationales

- KONÉ, Chaka, BELLEUDY, Cécile, Jean Claude BROUTART, et LE THANH, Nhan. *Emotion Recognition using Musical Mnemotherapy with three Physiological Sensors*. In : International Conference ICT, Society, and Human Beings 2017.
- KONÉ, Chaka, BELLEUDY, Cécile, et LE THANH, Nhan. *Energy Modeling and Architecture Exploration for Emotion Detection Systems*. In : Euromicro Conference on Digital System Design (DSD). 2016. p. Advanced Systems in Healthcare, Wellness and Personal Assistance Track.
- KONÉ, Chaka, MEFTAH, Imen Tayari, LE THANH, Nhan, BELLEUDY, Cécile. *Detecting and Classification of the Basic Emotions Using a Multimodal Approach for Emotions Detection*. In : The Fourth International Conference on Global Health Challenges, GLOBAL HEALTH 2015. 2015.
- KONÉ, Chaka, TAYARI, Imen Meftah, LE-THANH, Nhan, BELLEUDY, Cécile. *Multimodal Recognition of Emotions Using Physiological Signals with the Method of Decision-Level Fusion for Healthcare Applications*. In : International Conference on Smart Homes and Health Telematics. Springer International Publishing, 2015. p. 301-306.

Journal international

- Koné, C., Le Thanh, N., Flamary, R., & Belleudy, C. (2018). *Performance Comparison of the KNN and SVM Classification Algorithms in the Emotion Detection System EMOTICA*. International Journal of Sensor Networks and Data Communications, à paraître.

Colloque

KONÉ, Chaka, BELLEUDY, Cécile, et LE-THANH, Nhan. *Système de Détection des émotions : architecture et autonomie*. In : Colloque GDR SoC-SiP (System On Chip-System In Package). 2016.

Communications sans publication

- Demo : Chaka Koné, *Real Life Emotion Recognition Using Physiological Signals*, University Nice-Sophia Antipolis, France. In : Embedded Software And Micro-Electronics ESAME 2015. **Best University Booth prize**
- Poster : Chaka Koné, Imen Meftah Tayari, Nhan Le-Thanh and Cécile Belleudy : *How to Detect Human Emotions Using Physiological Signals ? A Multimodal Approach*. In : Sensors, Energy harvesting, wireless Network and Smart Objects conference. Gardanne 2015, october 14-16

Remerciements

Mes premiers mots de remerciements s'adressent à mes directeurs de thèse, Cécile Belleudy et Nhan Le-Thanh. Je leur remercie pour m'avoir confié ce sujet de thèse très passionnant, pour m'avoir fait bénéficier de leurs savoirs, de leurs précieux conseils et enfin de leurs disponibilités.

Je tiens également à remercier des personnes qui ont beaucoup contribué à mon travail de thèse. Je pense notamment à Jean Claude Broutard, pour m'avoir expliqué et initié à la mnémothérapie musicale. Il a su partager avec moi sa passion de ce domaine. Nous avons énormément travaillé ensemble pour définir le protocole de stimulation utilisé dans cette thèse. Je pense également à Rémi Flamary qui a consacré de son temps pour m'aider à mieux comprendre les techniques de validation des hyper-paramètres des classifieurs. Je pense également à Daniel Gaffé pour avoir relu ma thèse dont les commentaires, critiques et suggestions m'ont été d'une grande aide.

Je souhaite ensuite remercier tous les membres du laboratoire LEAT. Je remercie particulièrement les camarades doctorants, post-doctorants et stagiaires que j'ai eu à fréquenter pendant ces dernières années. Je les remercie pour leurs gentillesse, leurs conseils et surtout pour leurs amitiés. Je suis heureux d'avoir pu partager avec eux de nombreux moments de convivialité, de débats enrichissants et d'activités diverses notamment les soirées aux bars et restaurants, les sorties à la plage, à la montagne pour le ski ou la randonnée sans oublier les pauses café.

Mes remerciements s'adressent aussi à tous les membres de l'Association des Étudiants et Stagiaires Maliens de Nice pour tous les bons moments que nous avons passés ensemble à l'occasion des différents "grins" ou lors de nos rencontres festives. Merci aussi à tous les amis et camarades que j'ai rencontrés durant ces huit dernières années. Tous et toutes, m'ont apporté d'énormes choses.

Merci également à mes parents, à mes frères et sœurs, ainsi qu'à tous mes amis qui sont au Mali pour leurs soutiens inconditionnels, leurs conseils. Je reste si loin mais si proche d'eux.

Je ne saurais oublier de remercier ma cousine, mon cousin et sa femme ainsi que mes nièces et mon neveu pour toute leurs gentillesse, leurs soutiens tout au long de ces cinq dernières années. Merci à eux pour les nombreux conseils qui m'ont certainement

aidés dans mes choix de projets professionnels, pour leurs sympathies, et pour les bons plats traditionnels qui à chaque fois garantissait un bon appétit.

"L'écriture est la photographie du savoir mais elle n'est pas le savoir lui-même. Le savoir est une lumière qui est en l'homme. Il est l'héritage de tout ce que les ancêtres ont pu connaître et qu'ils nous ont transmis en germe, tout comme le baobab est contenu en puissance dans sa graine". Amadou Hampâthé Bâ

Sommaire

Résumé	ii
Abstract	iv
Liste des publications	vi
Remerciements	viii
Table des matières	xiii
Liste des tableaux	xiv
Liste des figures	xvi
Introduction générale	1
1 Etat de l'art	8
1.1 Détection des émotions	9
1.1.1 Vers une définition de l'émotion	9
1.1.1.1 Définitions	9
1.1.1.2 Rôle des émotions dans les interactions sociales	16
1.1.2 Reconnaissance des émotions	17
1.1.2.1 Reconnaissance des émotions dans les expressions faciales	17
1.1.2.2 Reconnaissance des émotions dans les expressions vocales	20
1.1.2.3 Reconnaissance des émotions dans les signaux physiolo-	
giques	22
1.1.3 Synthèse	28
1.1.4 Conclusion	30
1.2 Exploration d'architecture d'objets communicants pour la santé	30
1.2.1 Objets communicants pour la santé	31
1.2.2 Techniques d'exploration d'architecture	34

1.3	Consommation énergétique dans les systèmes embarqués médicaux . . .	34
1.3.1	Quelles contraintes sur les systèmes d'acquisition de signaux physiologiques ?	35
1.3.2	Origines de la consommation	36
1.3.2.1	Puissance statique	36
1.3.2.2	Puissance dynamique	38
1.3.2.3	Quelques techniques de réduction de la consommation énergétique	39
1.3.3	Décharge de la batterie	41
1.3.4	Modélisation et estimation de la consommation énergétique . . .	42
1.3.5	Conclusion	44
2	Détection des émotions par des signaux physiologiques	46
2.1	Introduction	46
2.2	Approche monomodale	47
2.2.1	Acquisition des signaux	48
2.2.2	Apprentissage	50
2.2.2.1	Découpage	51
2.2.2.2	Filtrage	52
2.2.2.3	Extraction des vecteurs caractéristiques	53
2.2.2.4	Construction d'une base d'apprentissage	54
2.2.2.5	Validation des paramètres	55
2.2.3	Détection	57
2.2.4	Choix d'un classifieur	60
2.2.4.1	KNN [Larose, 2005]	60
2.2.4.2	SVM [Wang, 2005]	61
2.2.4.3	Résultats et comparaisons des performances des deux classifieurs	63
2.3	Approche multimodale	66
2.3.1	les différentes techniques de fusion des signaux	67
2.3.1.1	Fusion au niveau signal	67
2.3.1.2	Fusion au niveau vecteurs caractéristiques	67
2.3.1.3	Fusion au niveau des décisions monomodales	68
2.3.2	Résultats et discussions	69
2.4	Apport par rapport à l'existant	72
2.5	Conclusion	72
3	Architecture du système d'acquisition et autonomie	74
3.1	Architecture de collecte de donnée pour la détection des émotions	74
3.2	Modèle de consommation	76
3.2.1	Quelques protocoles de communications	77
3.2.2	Modèle analytique d'estimation de la consommation et de l'autonomie	79

3.2.3	Caractérisation des paramètres	83
3.2.4	Test et résultats	85
3.3	Étude de l'influence des choix architecturaux sur la qualité de la détection	91
3.4	Conclusion	92
4	Vers un framework couplant autonomie et performance	93
4.1	Conception et implémentation d'un framework de simulation et d'exploration	94
4.1.1	Pourquoi un nouveau framework d'exploration et de simulation ?	95
4.1.2	Généralités sur la modélisation SystemC	97
4.1.3	Principes de la modélisation au niveau transactionnel (TLM)	99
4.2	Développement du framework de simulation/exploration	100
4.2.1	Architecture générale du framework	100
4.2.2	Modélisation du module de traitement	101
4.2.3	Modélisation du module radio	103
4.2.4	Modélisation du module sensors	104
4.2.5	Modélisation du module batterie	105
4.3	Résultats et discussions	106
4.3.1	Modélisation du système d'acquisition avec un noeud capteurs	106
4.3.2	Étude de l'influence du microcontrôleur sur la consommation	108
4.3.3	Étude de l'influence de la fréquence d'échantillonnage sur la consommation	111
4.3.4	Étude de l'influence du protocole de communication sur la consommation	113
4.3.5	Modélisation du système d'acquisition avec trois nœuds capteur : étude de l'autonomie	116
4.4	Conclusion	118
5	Le système Emotica	120
5.1	Architecture de collecte de données d'Emotica	122
5.2	Architecture de traitement des données dans le système Emotica	124
5.2.1	Application Matlab	124
5.2.2	Application mobile	127
5.3	Application visée et conception de notre propre base d'apprentissage de signaux physiologiques	129
5.4	Résultats et discussions	134
5.5	Conclusion	136
	Conclusion générale et perspectives	138
	A Glossaire	142
	B Conception de la base de données	143
B.1	Programmation du microcontrôleur	152

C	Classification des émotions et Big five	154
C.0.1	Modèle de combinaisons des émotions	154
C.0.2	Comment classifier les émotions en trois classes suivant leurs valences	155
C.0.3	Interprétation du Big five	155
D	Exemple d'un ordonnancement implémenté dans le système Emotica	156
E	Comparaison de différents microcontrôleurs	157
F	Personnalisation de la détection des émotions	159
	Bibliographie	162

Liste des tableaux

1.1	Définition et différenciation des différents états affectifs. Tableau extrait de [Scherer and Zentner, 2001]	11
1.2	Ensemble des émotions de base selon les différents auteurs : Tableau extrait et adapté de [Tato, 1999] [Tayari, 2013]	15
1.3	Quelques études sur les systèmes de détection des émotions utilisant les expressions vocales : Tableau extrait de [Devillers et al., 2005]	22
1.4	Quelques études sur les systèmes de classifications des états affectifs "affects" utilisant les signaux physiologiques : Tableau extrait de [van den Broek et al., 2011]	28
2.1	Images utilisées dans le protocole du MIT pour provoquer les différents émotions chez l'acteur.	50
2.2	Comparaisons des différentes méthodes de détection uni-modale instantanée des émotions	65
2.3	Taux d'erreur suivant les valences émotionnelles.	71
2.4	Comparaisons des différentes méthodes de détection multimodale instantanée	71
3.1	Récapitulatif des protocoles de communication utilisés dans notre travail	77
3.2	Récapitulatif des paramètres mesurés sur Bitalino	83
3.3	Autonomie selon différents protocoles et de chip RF : $F_s=10$ Hz, $I_{mcu}=11.05$ mA	87
3.4	Autonomie en fonction du module radio : fréquence du mcu 1 MHz, $F_s=10$ Hz, $I_{mcu}=6.77$ mA	88
3.5	Autonomie en fonction de la fréquence d'échantillonnage avec $I_{mcu}=11.05$ mA, et une batterie de 500 mAh	88
3.6	Autonomie en fonction de la fréquence d'échantillonnage avec $I_{mcu}=11.05$ mA et une batterie de 1300 mAh	89
3.7	Comparaison entre autonomie mesurée et calculée	90
3.8	Taux de reconnaissance des émotions en fonction de la fréquence d'échantillonnage	91

4.1	Quelques simulateurs de réseaux de capteurs sans fils	96
4.2	Taille de la mémoire et courants consommés par les MCUs à 8 MHz . .	103
4.3	Comparaison des deux modèles d'estimation développés	116
5.1	Comparaison des deux bases de données	136
B.1	Différences entre Thread et tâches asynchrones	148
E.1	Taille de la mémoire et courants consommés par les MCUs à 8 M Hz .	157
E.2	comparaison des microcontrôleurs	158

Liste des figures

1.1	Le modèle tri-dimensionnelle de Wundt. Chaque émotion peut être comprise comme un pattern d'activation de ressentis émotionnels pouvant chacun être situé dans cet espace.	12
1.2	Modèle bi-dimensionnelle des émotions de Russell avec la dimension horizontale de valence et la dimension verticale d'activation. Figure extraite de Feldman-Barrett & Russell (1998).	13
1.3	Cône des émotions de Plutchik. Figure extraite de [Plutchik, 1980] . . .	15
1.4	Le Facial Action Coding System composé de 44 unités d'action.	18
1.5	Les différentes étapes d'un système de détection des émotions utilisant les expressions faciales.	19
1.6	les différentes étapes d'un système de détection des émotions en utilisant les expressions vocales.	21
1.7	Capteur EMG utilisant un Raspberry Pi.	24
1.8	Capteur GSR sans fil fourni par Shimmer. Il permet de mesurer les caractéristiques électrique ou la conductance de la peau.	25
1.9	Capteur sans fil EEG à 14 canaux proposé par EMOTIV Epop+. . . .	25
1.10	Capteur ECG envoyant directement le signal à un dispositif mobile [Jurik et al., 2009].	26
1.11	Capteur BVP	26
1.12	Capteur mobile de respiration permettant de capter la respiration ventrale ainsi que thoracique [Baud and Sèdes, 2009].	27
1.13	Processus de conception d'un système. Cette figure est extraite et adaptée de [Carnot, 2011]	31
1.14	smart body analyser ws-50 de Withings	32
1.15	Schéma de l'implant cochléaire DIGISONIC [Ménard, 2008]	33
1.16	Système d'aide des paraplégiques basé sur trois appareils pour la rééducation	34
1.17	Inverseur réalisé à l'aide de transistors CMOS	37

1.18	Les principaux niveaux d'abstraction. Sur cette figure, on ne recense que les 4 principaux niveaux car les niveaux algorithmique et système sont fusionnés. Cette figure est tirée du cours de Bertrand Granado en master ESA à l'université de Cergy Pontoise	40
2.1	Méthodologie générale de notre méthode de détection des émotions.	48
2.2	Synoptique de la phase d'apprentissage.	51
2.3	Représentation du signal brut EMG sur une journée d'acquisition avec les huit émotions.	52
2.4	Les techniques de séparation des données en base d'apprentissage et de test. Figure extraite de [Flamary et al., 2015]	56
2.5	Découpage de la banque de signaux disponible en base d'apprentissage, validation et test.	57
2.6	Synoptique de la phase de détection	58
2.7	Modèle de combinaison des émotions proposé par Plutchik [Plutchik, 1980].	59
2.8	Taux de reconnaissance mono-modal des émotions en utilisant l'algorithme du KNN.	64
2.9	Taux de reconnaissance mono-modal des émotions en utilisant l'algorithme du SVM.	65
2.10	Fusion au niveau signal	67
2.11	Fusion au niveau des caractéristiques.	67
2.12	Fusion au niveau des décisions.	68
2.13	Taux de reconnaissance multi-modal des émotions en utilisant l'algorithme du KNN.	70
2.14	Taux de reconnaissance multi-modal des émotions en utilisant l'algorithme du SVM.	70
3.1	Chaîne d'acquisition et de traitement pour la détection des émotions	75
3.2	Positionnement des 4 protocoles de communication suivant le débit et la sensibilité [Semtech, 2017]	78
3.3	Plate-forme Bitalino sous ses 3 formats à savoir Board, Freestyle et Plugged	85
3.4	Autonomie en fonction de la fréquence d'échantillonnage et du protocole de communication	89
4.1	Environnement pour l'exploration d'architecture	94
4.2	Architecture d'un noeud capteur	101
4.3	Machine d'états finis associée au microcontrôleur	102
4.4	Machine d'états du module radio	104
4.5	Représentation d'un signal ECG	107
4.6	Résultat de l'exploration d'architecture pour un noeud composé d'un capteur et fonctionnant à 100 Hz. A gauche, la consommation en courant et à droite, l'occupation de la mémoire	109
4.7	Résultat de simulation pour 2 capteurs à 100 Hz	110

LISTE DES FIGURES

4.8	Résultat de simulation pour 3 capteurs à 100 Hz	111
4.9	Résultats de simulation pour 3 capteurs à 10 Hz	112
4.10	Résultats de simulation pour 3 capteurs à 1000Hz	113
4.11	Résultat de simulation de consommation en courant pour les protocoles Bluetooth et Zigbee	114
4.12	Résultat de simulation de consommation en courant pour le protocole Lora avec la modulation Lora à gauche et la modulation GFSK à droite	115
4.13	Architecture matérielle proposée pour un noeud de capteur	117
5.1	Fonctionnement global du système Emotica	121
5.2	Les trois nœuds d’Emotica	122
5.3	Dispositif servant à l’indexage des signaux physiologiques	124
5.4	Fenêtre principale et celle de configuration du GUI Matlab d’Emotica .	125
5.5	Fenêtre d’acquisition du GUI Matlab d’Emotica	126
5.6	Fenêtres principale et de configuration de la liste des émotions sur Android	128
5.7	Fenêtre Android permettant de créer un nouveau profil et de lancer une recherche	129
5.8	Taux de reconnaissance multi-modal des émotions en utilisant l’algorithme du KNN sur notre propre base	134
5.9	Taux de reconnaissance multi-modal des émotions en utilisant l’algorithme du SVM sur notre propre base.	135
B.1	Diagramme de classes de la base de données	145
B.2	Web service	146
B.3	Processus réel derrière le procédé des tâches asynchrones	148
B.4	Méthode de création d’un spinner	149
B.5	Machine d’états d’un noeud de capteur	152
B.6	Arbre des dépendances du code MCU	153
D.1	Exemple d’un ordonnancement sur l’architecture d’Emotica	156
F.1	Modèle de personnalisation de la détection des émotions	159

Introduction générale

Fondements

Dans le domaine de la psychologie, la communication des personnes est définie comme étant le processus par lequel une personne (ou un groupe de personne) émet un message et le transmet à une autre (ou un groupe). Le message reçu présente généralement une marge d'erreur. Au niveau émetteur, ces erreurs sont dûes au codage de la langue parlée ou écrite, au langage gestuel ou d'autres signes et symboles. Au niveau récepteur, elles sont liées au décodage du message. Enfin, il ne faut pas négliger l'influence du canal de communication lors de leur génération. A cause de ces erreurs, des études ont montré que plus de 80% de la communication est non verbale [Tellier, 2008]. C'est à dire que lorsque nous discutons, 80% de ce que nous entendons ou analysons sont : la tonalité de la voix, le débit de la parole, les expressions faciales, les réactions physiologiques ... C'est ainsi qu'un adage du Mali dit que *"la parole est trois choses : ce qu'on veut dire, ce qu'on dit réellement et ce que notre interlocuteur comprend"*. La notion d'émotion intervient car ce que *nous voulons dire* est largement influencé par l'environnement dans lequel nous sommes donc par les émotions que nous ressentons (celles exprimées ou non). Ainsi, notre interlocuteur, aussi, ne comprend que l'information que ses émotions acceptent de laisser passer. Dès lors, l'homme a commencé à étudier les émotions humaines inconsciemment. En discutant les uns avec les autres, nous écoutons non seulement les mots prononcés mais nous essayons aussi de sentir son état émotionnel. Ceci permet de disséquer le vrai du faux, comme par exemple, reconnaître un style de forme comme le sarcasme qui consiste effectivement à dire quelque chose que nous ne pensons guère et c'est à l'interlocuteur de le comprendre par lui même. De la même façon, imaginons une entreprise présentant son nouveau produit devant une assistance non conquise au préalable. Tout au long de sa présentation, l'orateur essaiera de capter une réaction émotionnelle de son assistance. Ainsi, il pourra amener son assistance aux émotions qu'il désire en adaptant efficacement ses propos en conséquence ; c'est ce que nous appelons l'art de la parole. Les politiques sont peut-être ceux qui ont, le mieux compris, l'importance des émotions dans notre société, c'est pourquoi ils suivent des formations adéquates leur permettant de mieux gérer leurs

émotions et ainsi de manipuler peut-être la population, en faisant exactement paraître (ressentir) l'émotion que le peuple attend d'eux au moment opportun.

Longtemps, la recherche sur les émotions relevait des domaines de la psychologie, de l'éthologie, bénéficiant des connaissances de l'anthropologie, de la sociologie, de l'ethnologie et des neurosciences. C'est ainsi que de nombreux travaux existent dans ces domaines pour comprendre la notion d'émotion, leurs impacts dans nos interactions sociales, dans notre processus de développement et, comment et pourquoi, elles se manifestent. Des modèles de la psychologie ont donc émergé pour modéliser les émotions et pour définir leurs combinaisons. Ensuite, grâce à la révolution technologique liée, entre autre, à la miniaturisation des composants, notre société est devenue assujettie aux nouvelles technologies et à l'informatique qui sont présentes dans tous les aspects de notre vie personnelle et professionnelle. En effet, ces nouvelles technologies sont parties intégrantes de nos quotidiens. Elles sont dans nos maisons, nous assistant dans la cuisine, dans nos voitures dans le contrôle de la mécanique et du multimédia, dans nos objets de divertissements au quotidien. De ce fait, afin qu'ils paraissent socialement intelligents, ces nouvelles technologies doivent contenir des modèles pour détecter automatiquement les états émotionnels des personnes. C'est ainsi qu'est née l'idée de l'informatique affective par Rosalind Picard en 1995. L'informatique affective fait partie intégrante de l'intelligence artificielle et a pour objectif d'étudier les interactions entre les technologies et les émotions. Elle vise à donner aux machines la capacité de comprendre, de synthétiser nos émotions et même d'exprimer des émotions. Elle permet donc de donner une dimension supplémentaire dans les relations homme-machine. Cette relation homme-machine est explorée dans le film *"I Robot"* avec Will Smith extrait de l'œuvre *Le Robot qui rêvait* d'Isaac Asimov. Cet auteur de Sciences fiction très célèbre a défini les trois lois de la Robotique qui ont été adoptées par les roboticiens eux-mêmes. Les œuvres d'Asimov sur la robotique, se focalisent sur les besoins des robots dans la conquête spatiale, dans l'aide aux personnes dans leurs quotidiens. Ces robots sont pensants, sensitifs et doivent être hors de danger pour l'être humain grâce aux trois lois. Les applications de l'informatique affective sont multiples ; elles vont des applications de jeux interactifs ludiques, à celles de E-learning en passant par des applications d'e-santé comme la prévention de la dépression, la prise en charges de personnes âgées ... et des applications socio-économiques.

Dans cette thèse, nous nous intéressons à la détection des émotions humaines. L'analyse des expressions faciales, vocales ou des signaux physiologiques permet de détecter les émotions. Contrairement aux expressions faciales ou vocales avec lesquelles nous pouvons facilement cacher ou tromper nos émotions, il est très difficile de feindre ses signaux physiologiques. Nous avons donc choisi d'utiliser ces signaux physiologiques issus du système nerveux autonome dans nos travaux. Ils sont sous contrôle involontaire du sujet et sont acquis grâce à l'utilisation de capteurs plus ou moins invasifs. Heureusement, nous observons une véritable évolution des technologies de conception de circuits intégrés, ce qui permet d'avoir des systèmes électroniques de plus en plus performants et complexes composés parfois de milliards de transistors, de plusieurs mémoires, d'unité de traitement, de circuit de connectivité, de capteur, ... Les systèmes

ainsi conçus ont des exigences sur la taille et le poids, et demandent une maîtrise de la consommation énergétique. Pour déterminer les émotions, nous avons besoin en premier lieu de plateformes capables de faire l'acquisition des signaux physiologiques. De ce fait, afin de fournir le système le plus efficace possible, nous devons satisfaire les contraintes utilisateurs qui sont liées aux dimensions de cette plateforme, à son autonomie et à faible prix.

De plus, d'après l'organisation mondiale de la santé (OMS), en 2018, la dépression serait au 2^{ème} rang des maladies les plus sévères en terme de coût global, après les maladies cardio-vasculaires. En 2030, elle représentera la 1^{ère} cause d'invalidité [OMS, 2017]. Les maladies neurodégénératives, comme la maladie d'Alzheimer où environ 40% des patients souffrent de dépression, sont caractérisées par la présence de perturbations cognitives et comportementales, entraînant progressivement une perte d'autonomie pour la réalisation des actes de la vie quotidienne. Ces perturbations fluctuent, rendant leurs évaluations parfois difficiles avec les outils cliniques habituels réalisés au cours d'une consultation. Le recours aux nouvelles Technologies de l'Information et de la Communication permet d'améliorer à la fois les techniques d'évaluation clinique, mais également les techniques thérapeutiques non médicamenteuses. Un suivi plus adéquat des patients passe, entre autre, par le fait de pouvoir évaluer son état émotionnel. Les informations produites constituent une aide au diagnostic et à l'applicabilité des traitements médicaux. Pour ces raisons, nous avons choisi d'axer nos travaux sur la détection des émotions pour des applications médicales.

Contexte applicatif et axes de recherche

Les domaines d'utilisation d'un détecteur d'émotion sont multiples. De nombreuses applications peuvent profiter d'un tel système pour modifier dynamiquement son comportement ou pour récolter des informations sur l'état d'un utilisateur. Par exemple, l'utilisation d'un tel système dans les jeux vidéo permet de modifier l'environnement virtuel du jeu en fonction des émotions (modification des scènes, du comportement des avatars, niveau de difficulté, ...). D'autres contextes applicatifs peuvent être imaginés comme dans l'apprentissage. Nous pouvons imaginer l'emploi d'un tel détecteur afin d'améliorer la productivité de l'apprentissage et obtenir des informations sur l'état d'écoute de l'auditoire. Un autre contexte applicatif concerne le domaine médical pour identifier la dépression ou des troubles liés au stress ou le suivi de l'état émotionnel des personnes âgées ou des personnes atteintes de la maladie d'Alzheimer. Des contraintes spécifiques existent pour chacun des domaines applicatifs que nous venons de citer. Bien que la détection des émotions reste le sujet, les contraintes liées aux choix des capteurs, aux choix des émotions à détecter, à l'environnement des utilisateurs et les informations pertinentes à identifier sont caractéristiques de chaque domaine d'application. Cependant, notre but était de rechercher un système adaptable à ces différentes contraintes en proposant une modélisation commune, mais une implémentation adaptée au contexte applicatif.

Dans le cadre de nos travaux, nous avons choisi un domaine d'application lié au milieu médical. Cette application concerne le suivi de l'état émotionnel et comportemental des personnes atteintes de la maladie d'Alzheimer. La maladie d'Alzheimer est une maladie neurodégénérative (perte progressive de neurones) incurable qui entraîne la perte progressive et irréversible des fonctions mentales et notamment de la mémoire. Elle est la cause la plus fréquente de démence chez l'être humain. Elle fut initialement décrite par le médecin Allemand Alois Alzheimer en 1906. La perte progressive de la mémoire s'explique par un dysfonctionnement de l'hippocampe. Cette dernière est responsable de la mémoire volontaire (c'est cet organe qui recherche une information dans notre mémoire et qui joue le rôle de bibliothécaire du cerveau). Chez les personnes atteintes de la maladie d'Alzheimer, la mémoire volontaire est non fonctionnelle alors que la mémoire involontaire fonctionne très bien. Grâce à une musique, à une odeur, à une image . . . , il est possible d'activer la mémoire involontaire. L'activation de cette mémoire permet au patient atteint de la maladie d'Alzheimer de se remémorer certains événements de sa vie évoqués, par une musique, une odeur, une image . . . Notre objectif est donc de faire sortir les patients de leurs états apathiques pour les amener à un état de reviviscence (revivre un épisode de sa vie). Le détecteur d'émotions, que nous voulons concevoir permet d'une part de réduire la plasticité cérébrale et par la même occasion d'apporter chaque jour un peu de joie au patient Alzheimer grâce à la stimulation cognitive. D'autre part, il permet le réapprentissage du bonheur et l'amélioration de l'engagement du patient à travers la mémoire inconsciente.

Pour répondre à cette problématique, nous avons utilisé une stimulation cognitive appelée Mnémothérapie musicale. La mnémothérapie musicale est une thérapie non médicamenteuse pour la maladie d'Alzheimer. Elle consiste à stimuler un patient atteint de la maladie d'Alzheimer avec des musiques connues et susceptibles d'être liées à des épisodes marquant de sa vie afin de l'amener dans l'état de reviviscence. Avec la Mnémothérapie musicale comme protocole d'induction des états émotionnels, nous avons sélectionné, pour la reconnaissance, les émotions suivantes : désagréable, neutre, plaisir, joie, reviviscence. Pour ces cinq émotions, l'objectif est de constituer une base de données de signaux physiologiques représentatifs et cette base de donnée doit contenir, entre autre, la base d'apprentissage. Il était alors essentiel d'étudier et de développer une architecture logicielle mais aussi matérielle capable de réaliser l'acquisition et le traitement des signaux physiologiques et par conséquent la détection des émotions. L'architecture matérielle a pour contrainte de satisfaire les contraintes applicatives et utilisateur. La conception d'un tel système nécessite, afin d'en évaluer ses caractéristiques, de se doter d'outil de simulation et d'exploration d'architecture pour ce type d'objets communicants. Cet outil doit être suffisamment générique pour permettre de définir un système d'acquisition à un haut niveau d'abstraction.

Pour répondre aux objectifs précités, les travaux, réalisés dans le cadre de cette thèse, s'ouvrent donc sur trois grands axes à savoir :

- La modélisation et la détection des émotions humaines en utilisant des signaux physiologiques : l'objectif de cette thèse est d'étudier et de développer un système complet de collecte de signaux physiologiques, de traitement et

d'interprétation de ces signaux. Le système développé permet de déterminer l'état émotionnel d'une personne et de réaliser ainsi le suivi de ses émotions et de son comportement. L'objectif est de définir un algorithme performant de reconnaissance des émotions. Pour y parvenir, différentes solutions ont été explorées : approche monomodale en utilisant différents classifieurs et les paramètres de configuration associés puis une approche multimodale. L'approche multimodale consiste à considérer des signaux de plusieurs capteurs différents et à fusionner ces signaux pour détecter un seul état émotionnel permettant d'avoir une vision plus globale de l'émotion ressentie par la personne. La pertinence de ces algorithmes a été validée sur une base de donnée du MIT puis la base de donnée construite.

- Modélisation de la consommation et de l'autonomie d'objets communicants à un haut niveau d'abstraction : une première étape a consisté à concevoir un système d'acquisition, de signaux physiologiques. Afin de répondre aux contraintes utilisateur, ce système doit posséder une autonomie acceptable. Pour répondre à cette contrainte, un nouveau modèle de consommation a été défini. Ce modèle est générique pour ce type de système et intègre des choix architecturaux et des paramètres de configuration. Afin de maintenir un taux de reconnaissance des émotions de qualité, ce dernier a été estimé en fonction des paramètres de l'architecture afin de rechercher le meilleur compromis performance/autonomie.
- Conception d'un environnement d'exploration d'architecture en prenant en compte les contraintes architecturales, applicatives et celles de l'utilisateur : Afin de concevoir un système répondant à ce cahier des charges, une exploration d'architecture était nécessaire et nous a conduit à développer un nouvel environnement de simulation pour évaluer de l'impact de différents choix architecturaux (unité de traitement, configuration et type de capteurs, protocole de communication) sur les contraintes. De plus, la mise en œuvre de l'architecture retenue a permis de créer une base de données sur des sujets dans le cadre de la mnémothérapie musicale.

Structure du manuscrit

Après une introduction générale fixant les problématiques et les axes de recherches de cette thèse, ce manuscrit se compose de cinq chapitres et est structuré comme suit :

Le premier chapitre débute par une introduction à la problématique de la détection des émotions et présente les bases des travaux réalisés. Il décrit la complexité de la définition de l'émotion, les différentes approches de représentation des émotions ainsi que leurs rôles dans nos interactions sociales. Un état de l'art est aussi fourni afin de recenser les grands travaux dans ce domaine en utilisant trois types de signaux (signaux physiologiques, expressions vocales et faciales). La deuxième partie de ce chapitre traite de la problématique de la conception du système d'acquisition et de traitement de ces signaux. Différentes techniques d'exploration d'architecture d'objets communicants

pour la santé sont décrites. Cette partie présente également la problématique de la modélisation de ces architectures et des paramètres associés.

Le deuxième chapitre présente notre méthodologie de reconnaissance des émotions à partir de signaux physiologiques. Dans un premier temps, les différentes étapes de ce traitement c'est à dire l'apprentissage, le filtrage, l'extraction de caractéristiques et enfin la détection propre des émotions sont introduites. Différents classifieurs lors de la détection ont été étudiés ainsi que la valeur des paramètres de configuration associés. L'objectif est d'aboutir à un taux de reconnaissance le plus élevé possible. Les premiers résultats obtenus à l'aide d'une approche monomodale n'étant pas satisfaisants, une approche multimodale de reconnaissance des émotions a été proposée. Enfin, dans un souci de validation, une comparaison est effectuée entre la méthode de détection définie et d'autres méthodes présentes dans la littérature sur la base de données partagée par le MIT.

Le troisième chapitre présente l'architecture matérielle du système de détection des émotions avec un intérêt particulier pour le système d'acquisition des signaux physiologiques. L'acceptation d'un tel dispositif ne peut être possible que si sa taille, son poids mais aussi son autonomie sont satisfaisants. Dans le cadre de cette thèse, nous nous sommes intéressés à la problématique de l'autonomie de ces systèmes. Ainsi un modèle d'estimation, à haut niveau d'abstraction, de la consommation énergétique et de l'autonomie pour des objets communicants a été proposé. Ce modèle s'appuie sur une analyse modulaire du système et par l'identification de paramètres adéquats traduisant leurs fonctionnement. L'évaluation de ces paramètres a été réalisée sur une plateforme d'acquisition de signaux physiologiques. Différentes expérimentations nous ont permis, d'une part d'analyser l'impact du protocole de communication, et d'autre part, de valider le modèle défini. Afin de proposer une architecture présentant une autonomie acceptable mais aussi une performance satisfaisante, ce modèle a été couplé avec le taux de reconnaissance des émotions.

L'objectif du quatrième chapitre est de proposer un environnement de simulation adapté à notre problématique et permettant de réaliser une exploration d'architectures pour la détection des émotions. Il débute par un bref état de l'art sur les simulateurs de réseaux de capteurs et conclut sur le besoin de développer notre propre simulateur. L'environnement de simulation/d'exploration comporte, une librairie de composants (unités de traitement, capteurs, module radio) qui sont modélisés par des machines d'états en SystemC TLM. Cet environnement nous a permis d'étudier à la fois la consommation de courant et l'autonomie en fonction du nombre de capteurs et de la fréquence d'échantillonnage, de l'unité de traitement et du protocole de communication. De plus selon la configuration de cette architecture, il est possible d'associer le taux de reconnaissance des émotions afin de rechercher le meilleur compromis.

Dans le cinquième chapitre, l'objectif est d'effectuer une validation de l'ensemble des travaux réalisés et, pour y parvenir, de se placer dans un contexte réel d'utilisation de la détection des émotions. Le contexte applicatif choisi est la Mnémothérapie musicale. Cette étude a été réalisée en collaboration avec le centre d'accueil de jour, des patients atteints de la maladie d'Alzheimer, de la fondation Noisiez à Sophia Antipolis. Dans

une première partie, l'architecture matérielle finale proposée ainsi que l'architecture logicielle sont décrites. L'acquisition des signaux physiologiques a nécessité de définir un nouveau protocole expérimental de capture de ces signaux sur un ensemble de sujet et a abouti à la création de notre propre base de données émotionnelle. Il faut noter que la base de données fournie par le MIT a été conçue sur des acteurs qui jouaient les émotions alors que dans nos expérimentations, les émotions étaient induites sur des sujets par des stimulus musicaux. Ainsi sur ce contexte applicatif, la validation de l'ensemble du système défini a été réalisée.

Ce mémoire s'achève sur une conclusion générale comportant une synthèse de notre contribution, tout en confirmant l'originalité de notre travail et précise les nombreuses perspectives de recherche envisageables en vue d'améliorer le système ainsi développé et de déboucher sur de nombreuses applications.

Chapitre 1

Etat de l'art

Sommaire

1.1	Détection des émotions	9
1.1.1	Vers une définition de l'émotion	9
1.1.2	Reconnaissance des émotions	17
1.1.3	Synthèse	28
1.1.4	Conclusion	30
1.2	Exploration d'architecture d'objets communicants pour la santé	30
1.2.1	Objets communicants pour la santé	31
1.2.2	Techniques d'exploration d'architecture	34
1.3	Consommation énergétique dans les systèmes embarqués médicaux	34
1.3.1	Quelles contraintes sur les systèmes d'acquisition de signaux physiologiques ?	35
1.3.2	Origines de la consommation	36
1.3.3	Décharge de la batterie	41
1.3.4	Modélisation et estimation de la consommation énergétique	42
1.3.5	Conclusion	44

Notre travail s'articule autour de deux grands axes :

- la détection des émotions et donc des algorithmes de détection sous-adjacents,
- l'exploration architecturale d'objets communicants pour la santé avec un accent particulier sur l'autonomie de ces systèmes.

L'objectif de ce travail est de proposer un système complet de reconnaissance des émotions satisfaisant toutes les contraintes applicatives et utilisateur qui se déclinent comme l'autonomie, la taille et le positionnement des capteurs ...

Pour atteindre cet objectif, il est nécessaire de comprendre la notion d'émotions, de consommation énergétique et d'autonomie d'un objet. Pour cela, nous commençons ce manuscrit par une présentation de quelques travaux sur la problématique de la définition

1.1 Détection des émotions

du mot *émotion*, son rôle dans les interactions sociales avant d'énoncer brièvement les modèles de représentations des émotions. Cette partie est suivie par un état de l'art sur les systèmes de reconnaissance des émotions et plus particulièrement sur les principes mis en œuvre. Après avoir conclu sur le type de détecteur d'émotion retenu, un état de l'art sur les systèmes d'acquisition et de traitement des signaux correspondant est présenté. Dans le but de concevoir ce type de système et de faire de l'exploration d'architecture, différentes techniques de modélisation sont résumées avec un intérêt particulier pour la consommation énergétique.

En raison de la complexité de la notion d'émotions, qui s'appuie sur plusieurs disciplines, cet état de l'art non exhaustif débute par une sélection de quelques travaux majeurs relatifs à cette problématique.

1.1 Détection des émotions

Étymologiquement le mot *émotion* vient du latin *emovere*, *emotum* (enlever, secouer) et de *movere* (se mouvoir) signifiant mouvement vers l'extérieur. Donc, l'émotion consiste en des mouvements et se résout par l'action. J'ai peur et je fuis, je convoite et je prends, je suis en colère et je déchire. Ceci permet de lier l'émotion à l'action mais elle montre toute la complexité de la définition de l'émotion. En effet, dès qu'on commence à parler des émotions tout semble confus et plusieurs questions nous viennent à l'esprit.

- Qu'est ce qu'une émotion ?
- Quelles différences existent entre émotion, sentiment et humeur ?
- Quelle est l'importance de l'émotion dans le processus de développement d'une personne ?
- Est-il vraiment possible de modéliser les émotions scientifiquement ?

Les réponses à ces questions se trouvent dans notre vie de tous les jours, dans les travaux des psychologues, sociologues, scientifiques, etc. Dans ce mémoire, nous tentons d'apporter des réponses à ces questions. Ces réponses nous permettront de comprendre la notion d'émotion et la problématique de sa modélisation d'un point de vue informatique.

1.1.1 Vers une définition de l'émotion

1.1.1.1 Définitions

D'après la définition étymologique de l'émotion, cette dernière est une notion remplie d'action, elle est donc liée au mouvement.

Chaque personne connaît la notion d'émotion, jusqu'au moment où on lui demande d'en donner une définition [Fehr and Russell, 1984]. Il est relativement aisé de dire à une personne qu'elle est sous l'emprise d'une réaction émotionnelle. En effet, nous remarquons/distinguons les états émotionnels de cette personne grâce à ces expressions faciales, l'intonation de sa voix, sa transpiration, ses expressions comportementales,

mais aussi et surtout grâce à ce qu'elle dit ou ne dit pas, à ce qu'elle fait ou ne fait pas [Piolat and Bannour, 2008].

L'émotion est une notion très complexe, cette complexité la rend difficile à comprendre donc à définir, elle diffère d'un individu à l'autre. Par exemple, si un groupe de personnes regarde un film d'horreur ou d'action, ils n'auront pas les mêmes expressions émotives. De même, si ces mêmes personnes participent à un concours qu'ils réussissent, l'expression de leur joie sera différente. De ce fait, il existe différentes définitions de l'émotion qui relèvent des différentes approches. Charles Darwin [Darwin, 1872], fondateur de la théorie de l'évolution, qualifie l'émotion comme une qualité innée, universelle et communicative dépendant de la faculté d'adaptation et de survie de notre organisme. De cette définition, on peut dire que l'émotion est un phénomène ayant trois composantes : une composante physiologique, psychologique et comportementale. Paul Ekman [Ekman, 1992] [Ekman, 1999], propose une approche attestant que le phénomène émotionnel se déroule selon 11 critères dont les 4 principaux critères communs sont :

- . **L'universalité** des signaux des différents états émotionnels et de leurs expressions : [Ekman et al., 1972] et d'autres auteurs [Izard and Malatesta, 1980] ont montré en faisant évaluer les expressions faciales sur diverses populations que l'expression des émotions de base est reconnue de façon universelle avec quelques petites nuances selon l'émotion et la culture.
- . Chaque émotion s'appuie sur un **contexte physiologique spécifique** : Ekman et ses collaborateurs ont montré qu'il existe des configurations spécifiques de mise en action du système nerveux autonome permettant à l'homme de se préparer à l'action.
- . Le **déclenchement** d'un état émotionnel et le **changement** d'un état à un autre est **rapide** ;
- . **La durée d'une émotion est limitée**. Les manifestations objectives des émotions de base durent quelques secondes et leurs expériences subjectives peuvent aller jusqu'à quelques minutes.

Les termes utilisés pour caractériser les sentiments et l'humeur créent encore plus de confusion quant à la différenciation de ces notions avec celle de l'émotion. Afin de palier à cette confusion, les critères définis par Ekman permettent de différencier les émotions (joie, tristesse, etc) des sentiments qui ont une durée prolongée et des déclencheurs moins précis (par exemple l'amour, l'amitié, etc) et des humeurs qui sont de durées variables et de causes peu précises (dépression, euphorie, etc). Ces notions sont toutes caractérisées par des états affectifs particuliers. Ainsi l'émotion "joie", le sentiment "amitié" et l'humeur "euphorie" sont vécus avec des affects spécifiques permettant de déterminer l'existence de chacun de ces états et de les nommer [Cosnier, 2003]. Scherer et Zentner [Scherer and Zentner, 2001] ont établi un ensemble de caractéristiques regroupés dans la Table 1.1. Ce tableau permet de caractériser chacun des états affectifs et des les différencier.

1.1 Détection des émotions

Caractéristique	Intensité	Durée	Synchronisation	Événement	Incitation d'évaluation	Rapidité du changement	Impact comportemental
Type d'état affectif : définition brève (exemples)							
Préférences : Jugements évaluatifs de stimuli dans le sens d'aimer ou de ne pas aimer, ou de préférer ou non un autre stimulus (aime, n'aime pas, positif, négatif)	F	M	TF	TH	H	TF	M
Émotions : Des épisodes relativement courts de réponse synchronisée de tous ou de la plupart des sous-systèmes organiques en réponse à l'évaluation d'un événement externe ou interne comme étant d'importance majeure (colère, triste, joyeux, craintif, honteux, fier, exalté, désespéré)	H	F	TH	TH	TH	TH	TH
Humeur : Diffuse l'état d'atteinte. Il est plus prononcé en tant que changement de sensation subjective, de faible intensité mais de durée relativement longue, souvent sans cause apparente (gai, sombre, irritable, apathique, déprimé)	M	H	F	F	F	H	H
Les positions interpersonnelles : Position affective prise envers une autre personne dans une interaction spécifique, colorant l'échange interpersonnel dans cette situation (lointain, froid, chaleureux, de soutien, méprisant)	M	M	F	H	F	TH	H
Attitudes : Relativement durable. Croyances et prédispositions de couleur affective envers des objets ou des personnes (aimer, aimer, haïr, valoriser, désirer)	M	H	TF	TF	F	F	F
Traits de personnalité : Emotionnellement chargé. Stabilité des dispositions de la personnalité et des tendances comportementales, typiques d'une personne (nerveuse, anxieuse, imprudente, morose, hostile, envieuse, jalouse)	F	TH	TF	TF	TF	TF	F

TF : Très Faible ; F : Faible ; M : Moyen ; H : Haut ; TH : Très Haut

TABLE 1.1 – Définition et différenciation des différents états affectifs. Tableau extrait de [Scherer and Zentner, 2001]

Dans la littérature, il existe plusieurs approches pour décrire l'ensemble des phénomènes émotionnels. Les deux approches les plus connues sont l'approche dimensionnelle et catégorielle.

Approche dimensionnelle : Cette approche se propose de décomposer les états émotionnels en plusieurs facteurs primaires dont l'objectif est d'établir une analogie entre la structure et les expressions des émotions. Wundt [Wundt, 1896] fut le premier à élaborer un modèle dimensionnel des émotions. Dans ces travaux, la sphère de la conscience liée aux émotions et aux sentiments est décrite en fonction de trois dimensions : plaisir-déplaisir, relaxation-tension, calme-excitation. La figure 1.1 illustre ce modèle de Wundt. Contrairement à Wundt qui a proposé un modèle plutôt circulaire, Woodworth [Woodworth and Schlosberg, 1938] a proposé un modèle linéaire de six émotions fondamentales, dans l'ordre suivant : bonheur, surprise, peur, colère, dégoût, mépris. Schlosberg inspiré du modèle linéaire de Woodworth a proposé un modèle circulaire à deux dimensions (bonheur-colère, surprise-peur). Quelques années plus tard, il ajouta une troisième dimension à son modèle après être influencé par les travaux de Duffy et nomma cette dimension (sommeil-tension) [Schlosberg, 1954]. Schlosberg, tout comme Wundt proposa donc un modèle circulaire à trois dimensions mais ne considéra pas les mêmes émotions que Wundt. Russel, dans son modèle circulaire [Russell, 1980]

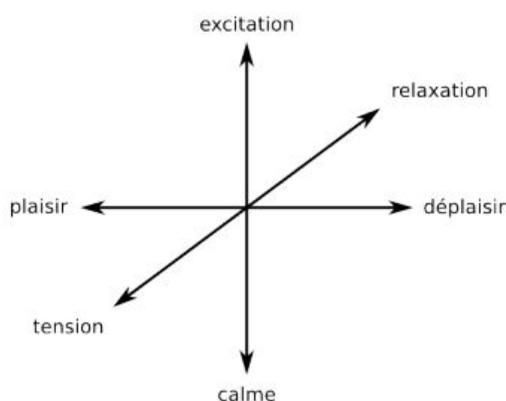


FIGURE 1.1 – Le modèle tri-dimensionnelle de Wundt. Chaque émotion peut être comprise comme un pattern d'activation de ressentis émotionnels pouvant chacun être situé dans cet espace.

propose de représenter les émotions dans un espace bi-dimensionnel (voir figure 1.2) *valence et activation* où se croisent quatre oppositions bipolaires. La valence permet de distinguer les émotions positives comme la joie, des émotions négatives comme la colère. L'activation représente le niveau d'excitation corporelle, qui transparait par le nombre de réactions physiologiques [Tayari, 2013], comme l'accélération du rythme cardiaque, la transpiration, etc. La combinaison des différentes dimensions permettrait de reconnaître des expressions faciales.

En 1997 [Russell, 1997], jugeant son modèle bidimensionnel insuffisant, il ajouta une dimension supplémentaire (*contrôle ou domination*) correspondant à l'effort du locuteur pour contrôler son émotion et permet de distinguer les émotions provoquées par le sujet lui-même ou par l'environnement. Cette nouvelle dimension permet de mieux distinguer

1.1 Détection des émotions

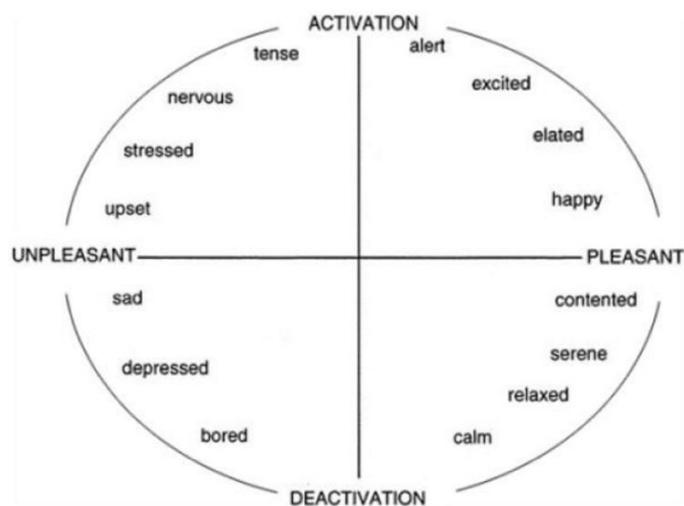


FIGURE 1.2 – Modèle bi-dimensionnelle des émotions de Russell avec la dimension horizontale de valence et la dimension verticale d'activation. Figure extraite de Feldman-Barrett & Russell (1998).

les émotions négatives les unes des autres (comme la distinction entre la colère qui pousse plutôt quelqu'un à prendre une posture de combat et la peur amenant souvent la personne à fuir).

D'autres au contraire, comme Duffy [Duffy, 1941], pensent que l'ensemble des comportements affectifs est réduit seulement au facteur **tension**. Duffy définit ce facteur comme une tension générale de l'organisme ou comme un niveau d'éveil.

Approche catégorielle : Cette approche basée sur les émotions dites "fondamentales" consiste à considérer les émotions comme des caractéristiques épisodiques et universelles [Darwin, 1872]. Les émotions fondamentales sont innées, automatiques, inconscientes et sont généralement exprimées dans le but de se protéger en présence d'un danger ou d'une urgence. Elles se trouvent dans toutes les sociétés humaines et même chez l'animal. Elles peuvent être définies comme étant la base de développement des autres émotions dites émotions secondaires qui se développent tout au long de la vie en interaction avec l'environnement dans lequel nous vivons.

Plusieurs auteurs ont travaillé sur cette approche catégorielle de représentation des émotions. Leurs travaux diffèrent selon le nombre d'émotions fondamentales qu'ils considèrent. Dans les paragraphes qui suivent, nous présentons quelques travaux.

Descartes pose les racines des théories modernes des émotions basiques. Ainsi en 1649, [Descartes, 1990], il décrit un modèle de fonctionnement des émotions humaines reposant sur 6 primitives émotionnelles : *l'émerveillement, l'amour, le désir, la joie, la haine et la tristesse*. Selon lui, ces émotions primaires sont ressenties par l'homme simultanément sous forme d'états émotionnels complexes et non de façon isolée. Plusieurs auteurs s'inscrivent sur la théorie évolutionniste de Darwin dont le caractère

universel des émotions permet de définir un nombre restreint d'émotions basiques. C'est ainsi que bien après Descartes, Tomkins [Tomkins, 1962] définit un modèle composé de 9 émotions primaires *la joie, l'intérêt, la surprise, la colère, le dégoût, l'anxiété, la peur, la honte et le mépris*. Par la suite, Izard [Izard, 1971] ajoute 3 autres émotions primaires à celles proposées par Tomkins *la tristesse, la culpabilité et la timidité* et enlève le *mépris* de la liste. Arnold [Arnold, 1960] propose onze émotions primaires : *la colère, l'aversion, le courage, le dégoût, le désir, la peur, la haine, l'espoir, l'amour, la tristesse, le désespoir*. Woodworth [Woodworth and Schlosberg, 1938] a proposé une schématisation linéaire de six émotions fondamentales, dans l'ordre suivant : *bonheur, surprise, peur, colère, dégoût, mépris*. Panksepp, en 1982, [Panksepp, 1982] proposa un modèle à quatre émotions primaires *l'espoir, la peur, la rage et la panique* puis remplaça cette liste en 1989 par *la colère, le dégoût, la tristesse et l'intérêt*. En 1993, McLean a proposé un modèle intégrant six émotions fondamentales : *le désir, la colère, la peur, la tristesse, la joie et l'affection*. Ces six sont respectivement liés à 6 comportements adaptatifs : la recherche, l'agression, la protection, l'abattement, le triomphe et la caresse.

Dans la littérature, les modèles précités sont peu utilisés en informatique. En effet, ces modèles ne proposent pas une combinaison sur plusieurs dimensions des émotions permettant d'aboutir à des émotions complexes. Ainsi, le dictionnaire émotionnel pouvant provenir de l'utilisation de ces modèles ne contiendra qu'une faible proportion des émotions humaines. Le modèle proposé par Plutchik [Plutchik, 1980] permet de remédier à cette lacune. Plutchik établit que les 8 émotions basiques sont construites de 4 paires opposées : *joie-tristesse ; acceptation-dégoût ; peur-colère et enfin surprise-anticipation*. Sur un cône, sont disposés les huit dimensions des émotions fondamentales. Au sommet de ce cône sont placés les termes qui désignent chaque émotion à son intensité maximale. La Figure 1.3 montre le cône des émotions de Plutchik. Les émotions élémentaires sont placées dans une roue correspondant ainsi à la première dimension émotionnelle. Les émotions secondaires correspondent à des mélanges d'émotions primaires et les émotions tertiaires correspondent à des mélanges d'émotions secondaires. Le modèle circulaire peut être transformée en cône afin de représenter les différents degrés d'intensité des émotions primaires et secondaires (voir figure 1.3).

Malgré la diversité des propositions, six émotions apparaissent quasiment dans toutes les propositions. Ces émotions sont communément connues sous le nom de "Big Six" et sont issues des travaux de Ekman, Friesen et Ellsworth [Ekman et al., 1982] (*la colère, le dégoût, la peur, la joie, la tristesse et la surprise*). Leur caractère universel et inné fait qu'elles sont rattachées respectivement à six comportements *"la ré-intégration, la destruction, la reproduction, la réjection, la protection et l'attention"*. Les théories modernes sur les émotions basiques divergent quant au nombre ou à la liste exhaustive des émotions de base. Le tableau 1.2 résume la liste des émotions de base selon quelques auteurs.

Les travaux présentés dans ce mémoire sont une continuité des travaux de thèse d'Imen Tayari [Tayari et al., 2013] qui s'appuient sur l'approche proposée par Plutchik

1.1 Détection des émotions

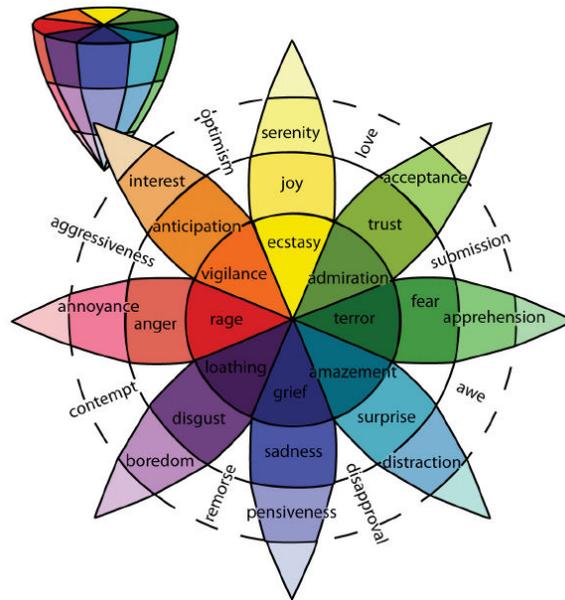


FIGURE 1.3 – Cône des émotions de Plutchik. Figure extraite de [Plutchik, 1980]

Auteurs	Émotions basiques
Ekman et al.	colère, dégoût, peur, joie, tristesse, surprise
Izard	colère, mépris, dégoût, détresse, peur, culpabilité, intérêt, joie, honte, surprise
Woodworth	bonheur, surprise, peur, colère, dégoût, mépris
Plutchik	acceptation, colère, anticipation, dégoût, peur, joie, tristesse, surprise
Tomkins	colère, intérêt, mépris, dégoût, détresse, peur, joie, honte, surprise
Arnold	colère, aversion, courage, découragement, désir, désespoir, peur, haine, espoir, amour, tristesse
Fridja	désir, bonheur, intérêt, surprise, émerveillement, peine
Scott	peur, colère, solitude, plaisir, amour, anxiété, curiosité
Arieti	peur, rage, satisfaction, tension, appétit
Epstein	peur, colère, tristesse, amour, joie
Osgood	peur, colère, anxiété-tristesse, joie, quiétude, plaisir, intérêt/attente, stupeur, inquiétude, dégoût
Gray	rage, terreur, anxiété, joie
Mower	douleur, plaisir
James McDougall	peur, chagrin, amour, rage, colère, dégoût, exaltation, peur, soumission, émotion tendre, émerveillement
Weiner and Graham	Bonheur, tristesse
Panksepp	espérance, peur, rage, panique

TABLE 1.2 – Ensemble des émotions de base selon les différents auteurs : Tableau extrait et adapté de [Tato, 1999] [Tayari, 2013]

[Plutchik, 1980] pour définir un modèle multidimensionnel de représentation des émotions dont chaque émotion pourrait être considérée comme un mélange des huit

émotions fondamentales définies par Plutchik. Dans nos travaux, nous allons utiliser le modèle de Plutchik. Ce modèle permet d'une part d'avoir un dictionnaire de plus de 70 émotions [Tayari, 2013]. D'autre part, il permet de proposer une approche vectorielle des différents états émotionnels où chaque émotion complexe pourrait être décrite comme une combinaison des émotions fondamentales proposées par Plutchik.

1.1.1.2 Rôle des émotions dans les interactions sociales

Les êtres humains sont par nature des êtres sociaux, de ce fait, les émotions ont un rôle social évident. Les émotions sont prépondérantes dans tout le processus d'adaptation d'un individu. En effet, plusieurs travaux en psychologie [Campos, 1983]; [Campos et al., 1989]; [Klunnert et al., 1983] mettent en évidence que les expressions émotionnelles sont essentielles aux relations avec l'environnement physique et social. De ce fait, dans le but de vivre en harmonie, nous essayons sans cesse de contrôler nos émotions ainsi que celle des personnes autour de nous. Elles entrent dans le processus de la communication non-verbale et nous permettent non seulement de communiquer avec notre environnement mais aussi d'innover, de réussir et de prospérer. En effet, les émotions peuvent nous subjuguier en étant source de douleur et de plaisir mais aussi d'informations dans certaines situations.

Rivière et Godet [Rivière and Godet, 2003] attestent que les émotions sont nécessaires au bon fonctionnement de nombre de nos facultés, comme la mémoire, le raisonnement, la prise de décision ou encore l'adaptation sociale. En effet, nombre de problème en entreprise sont d'ordre relationnel entre "employé-employé" ou "employeur-employé".

Les émotions négatives (anxiété, peur, colère, haine, chagrin ...) peuvent affecter la santé physique et la qualité de vie. En effet, les personnes souffrant de dépression connaissent des sentiments sévères et prolongés d'émotions négatives comme la tristesse, la colère, le dégoût et la peur. La dépression majeure est une condition extrêmement éreintante pour les personnes souffrant de cette maladie. Les troubles dépressifs sont parmi les raisons les plus importantes d'invalidité dans le monde entier [OMS, 2017]. Aux États-Unis d'Amérique, environ 6,7% de la population (totalisant près de 15 millions de personnes) sont affectées, chaque année, par une maladie mentale grave et c'est la principale cause de handicap pour les Américains âgés de 15 à 44 ans [National Health, 2017]. Ces troubles de santé sont sources de problèmes pour les politiques de santé publique et sont la cause également de pertes considérables d'un point de vue économique.

Concernant les maladies neurodégénératives, les patients, par exemple, atteints de la maladie d'Alzheimer, perdent progressivement leurs autonomies tout au long de leur maladie. Ainsi, dans l'incapacité d'accomplir certaines tâches quotidiennes, ils se frustreront, et entrent souvent dans un état dépressif. Le traitement de cette maladie consiste d'une part à trouver des activités ludiques afin d'améliorer leurs capacités, leurs niveaux d'intérêt et leur humeur. D'autre part, il est nécessaire de stimuler les

1.1 Détection des émotions

cerveaux de ces patients afin de retarder la progression de cette maladie et les aider à se souvenir.

Les maladies neurodégénératives comme celle d'Alzheimer sont caractérisées par la présence de troubles cognitifs et comportementaux. Ces derniers fluctuent au fil du temps, rendant parfois l'évaluation difficile avec les outils cliniques standards. Pour cela, il est envisageable d'exploiter une approche non médicamenteuse des soins quotidiens. En effet, les approches cliniques prennent beaucoup de temps et impliquent un coût financier important, car elles nécessitent un entourage professionnel qualifié. Dans ce contexte, l'utilisation des nouvelles technologies de l'information et de la communication (TIC) sont un atout majeur. Elles peuvent améliorer les techniques d'évaluation clinique, mais aussi les traitements et techniques thérapeutiques non médicamenteux. Pour cela, il est nécessaire d'évaluer l'état émotionnel des patients afin d'aider à un diagnostic précis et à l'applicabilité des traitements médicaux. Nos travaux s'inscrivent sur cette problématique pour définir un système complet de collecte et d'analyse des signaux physiologiques d'une personne afin de déterminer son état émotionnel mais aussi de concevoir une thérapie basée sur la mnémothérapie musicale.

1.1.2 Reconnaissance des émotions

L'émotion occupe une très grande importance dans le domaine des sciences cognitives, de la philosophie ainsi que dans la psychanalyse. Depuis quelques années, elle est très utilisée dans le domaine de l'informatique. Son importance ne cesse d'être démontrée grâce à des applications très attractives. Parmi ces applications, nous pouvons citer : la surveillance des personnes âgées, la prévention de la dépression, les jeux interactifs, le E-learning, et le E-motion. Ce dernier est un projet d'une filiale de l'entreprise japonaise NCR, en collaboration avec l'Université de Californie du Sud. Ce projet vise à mettre en place des distributeurs de billets ayant la capacité de reconnaître le Big six (les six émotions fondamentales). D'un point de vue informatique, la reconnaissance des émotions passe par l'utilisation d'une ou plusieurs des modalités à savoir les expressions faciales, les expressions vocales ou sur la base des signaux physiologiques. Dans la suite de ce chapitre, nous présentons la détection des émotions à partir de chacune de ces modalités en mettant un accent particulier sur celle utilisée au cours de nos travaux à savoir la modalité des signaux physiologiques.

1.1.2.1 Reconnaissance des émotions dans les expressions faciales

Dans le processus de communication, les expressions faciales sont une importante source d'informations. Dans plusieurs sociétés comme les sociétés africaines, un enfant, par un simple regard des expressions du visage de ses parents sait exactement ce que ces derniers veulent lui dire et ajuste en conséquence ses paroles et ses actions. Ainsi, tout comme les parents de cet enfant, nous utilisons ces expressions du visage pour donner des informations aux gens qui nous entourent, afin qu'ils puissent convenablement ajuster ou modifier leurs actions ou mots.

L'électromyographie consiste à mesurer directement l'activité électrique des muscles faciaux avec des électrodes appliquées sur la surface cutanée qui donne accès aux modifications latentes de l'activité faciale non visible. Ainsi, le capteur électromyographe peut être utilisé pour capturer les expressions faciales. Ekman et Friesen à la fin des années 1970 [Ekman and Friesen, 1978] ont créé le Facial Action Coding System (FACS) distinguant 44 unités d'actions (UA) qui correspondent chacune à un changement d'apparence produit par la contraction d'un ou plusieurs muscles faciaux. Ce FACS (voir la figure 1.4) permet le codage de toutes les unités d'actions visibles sur un visage. Chaque UA peut correspondre à la contraction ou à la détente d'un ou plusieurs muscles qui se traduit par le mouvement d'une partie donnée du visage [Tayari, 2013].

Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
					
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
					
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
					
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
					
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Puckerer	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
					
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

FIGURE 1.4 – Le Facial Action Coding System composé de 44 unités d'action.

Sur la figure 1.4, les 30 unités d'action sont montrées. Elles sont automatiquement liées à la contraction d'un ensemble spécifique de muscles du visage et les 14 unités d'action restantes dont la base anatomique est indéterminée sont désignées comme des actions diverses. Ces 14 unités d'action ne sont donc pas représentées sur la figure 1.4.

Dans un système d'analyse des expressions faciales, trois étapes sont nécessaires.

1.1 Détection des émotions

- La première étape consiste à détecter le visage afin de limiter la zone de recherche. Deux approches de recherche existent dans la littérature [Mercier, 2003]. La première consiste à identifier indépendamment les composantes du visage les unes après les autres (nez, les yeux, la bouche...). L'analyse du visage par cette méthode consiste donc à appliquer un ensemble d'opérateurs spécialisés dans la détection et la mesure des caractéristiques d'une composante en tenant compte des connaissances sur ces différentes composantes. La seconde méthode considère le visage comme un tout et détecte donc le visage par comparaison avec des modèles de visages numériques préalablement modélisés. Cette seconde méthode est essentiellement basée sur la reconnaissance de formes ou de mise en correspondance de modèles.
- Après avoir détecté le contour du visage dans la première étape, la deuxième étape consiste à extraire des caractéristiques importantes pour une meilleure description des expressions faciales. Elle identifie les traits du visage comme l'écartement des yeux, l'ouverture de la bouche, le gonflement des joues . . .
- Enfin, sur la base de ces deux premières étapes, l'image extraite d'une catégorie d'expression est classifiée par des méthodes comme le séparateur à vecteurs support [Wang, 2005]. La Figure 1.5 illustre les différentes étapes d'un tel système d'analyse.

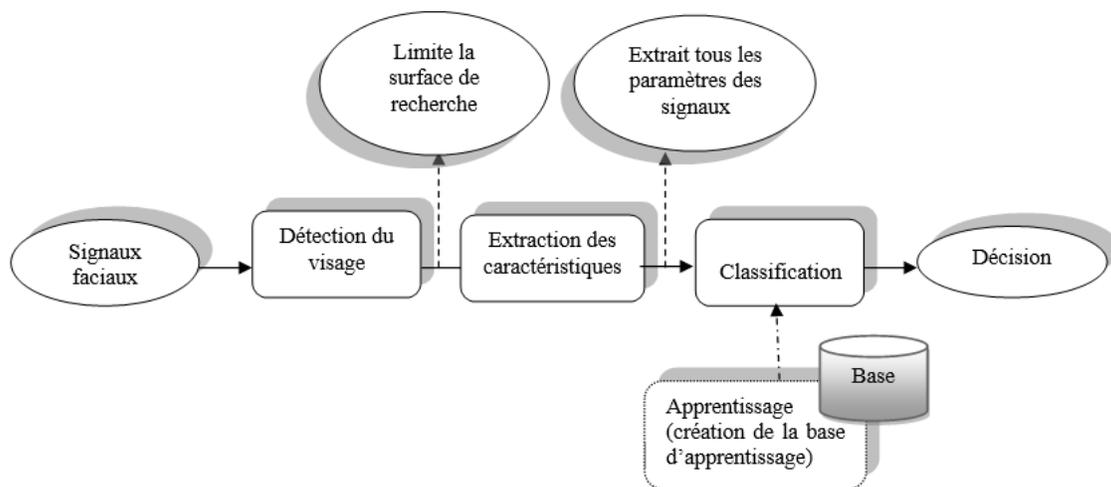


FIGURE 1.5 – Les différentes étapes d'un système de détection des émotions utilisant les expressions faciales.

Plusieurs travaux sur la détection des émotions ont été menés en utilisant les expressions faciales. Gao [Gao et al., 2003] propose une méthode de classification de trois expressions : neutre, sourire et colère.

Cependant, ce système présente des inconvénients parmi lesquels, nous pouvons citer :

1. Les problèmes de mobilité : L'acquisition étant faite à l'aide d'une caméra, le sujet, dont les expressions faciales sont acquises, devra toujours regarder dans l'axe de la caméra. Cette contrainte sur l'utilisateur limite considérablement la mobilité du système d'acquisition.
2. L'autonomie du système : L'analyse des signaux faciaux acquis avec la caméra engendre des coûts élevés de traitement. La plupart des systèmes actuels d'analyse des signaux faciaux ne sont pas souvent alimentés sur des piles en raison de leur consommation. Cependant, de nouvelles caméras commencent à voir le jour permettant d'enregistrer les différences entre 2 images successives. Cette avancée permettra théoriquement de diminuer la consommation de tels systèmes en vue d'augmenter leur autonomie.
3. Le respect de la sécurité de la vie privée : Dans le cadre d'une utilisation réelle de ce système, la caméra pourra enregistrer des scènes intimes lorsqu'on oublie de l'éteindre. C'est d'ailleurs le problème auquel Samsung s'est confronté lorsque le groupe a voulu fournir des téléviseurs haute définition offrant des caractéristiques pour le suivi du visage et pour la reconnaissance vocale. Ces caractéristiques peuvent permettre à des pirates de voir et d'entendre ce qui se passe à côté de ces téléviseurs.

1.1.2.2 Reconnaissance des émotions dans les expressions vocales

À l'instar des signaux faciaux, les expressions vocales sont aussi une véritable source d'information pour les émotions ressenties par une personne. En effet, ces expressions vocales constituent tellement un excellent canal de communication des émotions que les hommes politiques suivent des formations afin de transmettre les émotions à travers leurs discours. L'analyse automatique de ces expressions, dont les émotions sont une information sémantique importante, ont de nombreuses applications telles que l'interaction humain-ordinateur, la gestion des appels téléphoniques automatique et les centres de surveillance. Les expressions vocales sont habituellement mesurées à l'aide de microphones, mais peuvent aussi être mesurées à l'aide d'une caméra comme les expressions du visage. La Figure 1.6 montre les différentes étapes requises dans un système d'analyse des expressions vocales pour la détection des émotions.

Plusieurs travaux se sont intéressés à la reconnaissance d'émotions en utilisant les expressions vocales. Nous citons par exemple les travaux menés par (Devillers et al.) [Devillers et al., 2005] et ceux de (Lee et al.) [Lee et al., 2002b] sur l'adaptation de la stratégie dialogique dans les centres d'appels en déterminant l'état émotionnel de l'utilisateur. Cette stratégie permet de diriger, l'appel de l'utilisateur initialement mis en contact avec le répondeur, vers un humain lorsque des signes de frustration ou de nervosité sont détectés dans sa voix. Istrate [Istrate, 2003], au cours de ses travaux s'est intéressé à la reconnaissance des émotions pour aider les personnes âgées et hospitalisées. Nous pouvons citer un autre exemple dans le cadre du projet (Search and Rescue). Des robots collaboratifs intervenants dans une situation d'urgence ont été développés pour prendre en charge les émotions des victimes et des sauveteurs

1.1 Détection des émotions

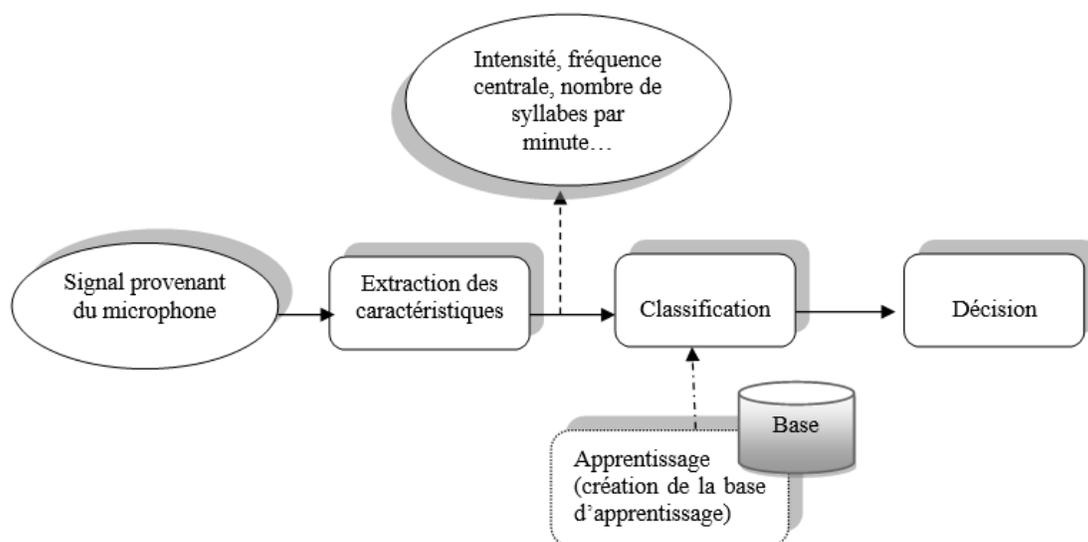


FIGURE 1.6 – les différentes étapes d'un système de détection des émotions en utilisant les expressions vocales.

[Looije et al., 2007]. (Clavel et al.) [Clavel et al., 2008], dans leurs travaux, ont intégré les expressions vocales dans les systèmes de surveillances (visuel) afin de permettre à la machine de diagnostiquer les situations anormales afin d'assister l'homme dans sa tâche de surveillance.

Le tableau 1.3 recense quelques travaux sur les systèmes de détection des émotions utilisant les expressions vocales. Dans tous ces systèmes de détection des émotions, une étape importante appelée "classification" est présente. Elle est basée sur des techniques de classification supervisée ou non. Ces techniques de classification opèrent des regroupements en classes homogènes d'un ensemble d'individus. Dans la classification supervisée, les classes (regroupements homogènes) sont connues et l'on dispose d'exemples de chaque classe. Dans la classification non supervisée, le nombre de classes et la caractéristique de chaque classe sont inconnus, il n'y a donc pas de connaissance a priori (pas d'exemples). Les algorithmes de classification supervisée consiste à labelliser des informations donc à classifier les données futures. L'extraction des données dans cette approche est donc prédictive. Les algorithmes de classification non supervisée quant à eux, cherchent à trouver des partitions de modèle par eux-mêmes. L'extraction des données est donc descriptive.

La plupart des études sur la détection des émotions dans la voix s'est focalisée sur un ensemble minimal d'émotions : positives et négatives, émotion vs. pas d'émotion, colère vs. amusé. Peu de système de détection des émotions utilisant cette modalité vont au-delà de la classification de 4 à 5 émotions de base (très différentes). Les taux de détection décroissent de façon significative sur des émotions nuancées et lorsque le nombre d'émotion devient un peu conséquent comme nous montre le tableau 1.3. Or l'étude de voix spontanées montre que les émotions sont souvent mélangées ou masquées

1.1 Détection des émotions

Authors reference	Style of Corpora	Corpora size	Emotion labels	Type of features	Machine learning	Detection rate
[Dellaert et al., 1996]	Acted	1000 turns (five actors)	Joy, fear, anger, neutral	Prosodic (pitch contour)	MLB, KR, kNN	60–65% (acted, four classes)
[Petrushin, 1999]	Acted	700 turns	Anger, sadness, fear, happiness, neutral	Prosodic spectral (F1, F2, F3)	NNs	70% (acted, five classes)
[Batliner et al., 2000]	Acted (non-professional)	56 calls (15–90 s)	Calm, agitation	Prosodic spectral part-of-speech dialogic	NNs	77% (acted, two classes) 1
	Acted	96 turns with emotion	Emotion, non emotion			95% (acted, two classes)
	Read-speech	50 turns (with emotion), 50 turns (NE)				79 % (read, two classes)
	WOZ vermobil	2395 turns (20 dial.)				73% (Woz, two classes)
[Ang et al., 2002]	DHM communicator	21 k turns (\approx 3500 turns with emotion)	Frustration, annoyance, other	Prosodic	CART decision tree	75% (DHM, two classes)
[Lee et al., 2002a]	DHM call centers (real-life) Speech-Works	7200 turns	Negative, non negative	Prosodic	LDC kNN	75% (DHM, two classes)
[Narayanan, 2002]	DHM call centers (real life)	7200 turns	Negative, non negative	Prosodic language dialogic	LDC kNN	80–90% (DHM, two classes)
[Shafran et al., 2003]	DHM AT&T how may I help you	5147 turns	Negative, non negative	Spectral (MFCC) F0	HMM	76% (DHM, two classes)
				Lexical	SVM	81% (DHM, two classes)
[Forbes and Litman, 2004]	DHM machine mediated	385 Utt. (90 Neg, 15 Pos, 280 Neu)	Positive, negative, neutral	Prosodic lexical dialogic	AdaBoost	84% (DHH-M, three classes)
[Steidl et al., 2005]	DHM AIBO	\approx 6000 words (51 children)	Anger, motherese, emphatic, neutral	Prosodic part-of speech	NNs	60% (DHM, four classes)

TABLE 1.3 – Quelques études sur les systèmes de détection des émotions utilisant les expressions vocales : Tableau extrait de [Devilleers et al., 2005]

MLB : maximum likelihood Bayes classifier ; **KR** : kernel regression ; **LDC** : linear discriminant classifier ; **kNN** : k nearest-neighbors ; **SVM** : support vector machine ; **HMM** : hidden Markov model ; **NNs** : neural networks.

dans les interactions naturelles. Tout ceci montre, à l’heure actuelle, les limitations quant à l’utilisation de tels systèmes.

1.1.2.3 Reconnaissance des émotions dans les signaux physiologiques

Les signaux physiologiques comme le rythme cardiaque instantané, la variabilité de la fréquence cardiaque et la sudation sont des paramètres importants dans le

1.1 Détection des émotions

monitoring de l'état de santé d'une personne pour avoir une information sur la présence d'un dysfonctionnement cardiovasculaire [Acharya et al., 2006], [Neto et al., 2003]. Ces marqueurs traduisent l'état du système nerveux autonome (SNA) et sont observés pour mettre en lumière une augmentation de l'activité sympathique responsable du contrôle de certaines activités inconscientes de l'organisme (comme le rythme cardiaque ou la contraction des muscles lisses) ou une réduction de l'activité vagale [Bousefsaf, 2014].

Le système limbique responsable de la régulation des émotions, des comportements faisant appel à la mémoire à long terme regroupe les structures cérébrales à savoir l'amygdale, l'hypothalamus et le cortex limbique. Les deux derniers à savoir l'hypothalamus et le cortex limbique ont des fonctions relatives à la mémoire et au contrôle de l'orientation dans l'espace. L'amygdale cérébrale dont le noyau est situé dans le lobe temporal est chargé de l'analyse du degré de menace et de la signification émotionnelle de toutes les informations à la fois internes et externes. Elle répond aux sensorialités visuelles, sonores, tactiles, gustatives et olfactives et permet le stockage des traumatismes émotionnels sans contrôle conscient et est liée aux phobies [Jacques et al., 2011].

Le Système Nerveux Autonome est une partie du système nerveux responsable des fonctions qui sont sous contrôle involontaire. Il est impliqué dans des fonctions de l'organisme telles que la thermorégulation ou encore la régulation de la pression sanguine par le contrôle du centre vasomoteur. Ce dernier contrôle le diamètre des vaisseaux sanguins pour pouvoir maintenir une pression artérielle à un niveau tolérable. Les signaux physiologiques dont l'acquisition se fait en utilisant des capteurs de moins en moins invasifs, permettent d'avoir une information sur l'état du système nerveux autonome. L'acquisition de tels signaux permet donc le monitoring de l'état émotionnel et comportemental de l'être humain. Les signaux physiologiques les plus utilisés à cet effet sont :

- la Réponse ÉlectroDermale RED ;
- le volume sanguin impulsionnel ;
- le signal du Volume Respiratoire (VR) ;
- l'activité ÉlectroMyoGraphique (EMG) ;
- la fréquence cardiaque (Fc) ;
- le rythme ÉlectroEncéphaloGramme (EEG).

Jennifer Healey [Healey, 2000], lors de ses travaux thèse, a utilisé les quatre premiers signaux physiologiques pour mettre en place une base de données de signaux physiologiques pour huit émotions. Cette base de données sert de référence pour une majorité des chercheurs travaillant sur les systèmes de détection d'émotions utilisant la modalité des signaux physiologiques. Dans la suite de ce manuscrit, nous présentons en détail l'ensemble de ces signaux.

Activité électromyographique (EMG) : Toute contraction des muscles est accompagnée par la création de potentiels électriques. Ces potentiels électriques peuvent être mesurés par électromyographie (EMG). Le signal électromyographie représente donc l'enregistrement d'une série événements électriques (potentiel membranaire du

muscle) produits par la fibre musculaire lorsque les muscles se contractent. Le tonus émotionnel est une contraction involontaire, permanente et modérée des muscles, entretenue par des influx nerveux. Trois électrodes sont utilisées pour mesurer le signal EMG : deux sont placées le long de l'axe du muscle d'intérêt et une troisième électrode de masse est placée hors de l'axe de ce muscle. Il a été démontré que ce signal se corrèle avec le stress et les émotions à valence négative [Cacioppo and Tassinary, 1990]. L'EMG a été utilisé pour étudier les expressions faciales [Ekman, 2006], les expressions gestuelles [Marrin and Picard, 1998] et le concept de valence émotionnelle [Lang, 1995].

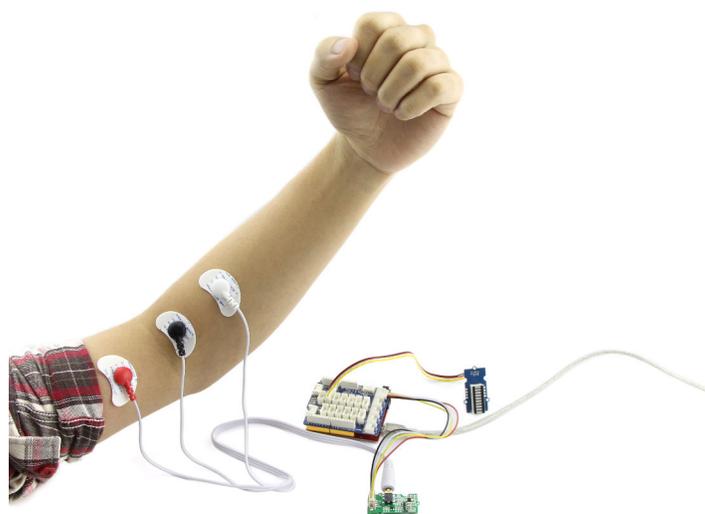


FIGURE 1.7 – Capteur EMG utilisant un Raspberry Pi.

Activité électrodermale (EDA) : La température du corps humain est régulée par le processus de la sécrétion des glandes sudoripares. Cette sécrétion des glandes sudoripares est aussi associée à l'activité du système nerveux sympathique. Chaque activité émotionnelle est partiellement traduite dans la production de la sueur ou l'inhibition des glandes sur les mains, les paumes et les pieds. Cela change la résistance de la peau. Le signal de l'activité électrodermale (EDA) permet de traduire ces changements de résistance de la peau en valeurs numériques. Ces données reflètent le niveau de conductivité électrique de la peau, et elles se mesurent à la surface de la peau. Dans les situations impliquant les émotions, cette activité électrique varie très sensiblement et peut fournir une mesure indirecte de l'intensité émotionnelle. Il donne une bonne information sur l'évaluation du stress, de l'anxiété et de l'attention, et sa réponse croît linéairement avec le niveau d'excitation globale d'une personne [Nakasone et al., 2005].

Rythme Électroencéphalographique (EEG) : L'électroencéphalogramme mesure l'activité électrique du cerveau en plaçant des électrodes sur la surface de la tête. Un

1.1 Détection des émotions



FIGURE 1.8 – Capteur GSR sans fil fourni par Shimmer. Il permet de mesurer les caractéristiques électrique ou la conductance de la peau.

électroencéphalogramme complet comprend plus de 128 électrodes, mais des mesures plus simples utilisant moins de 20 canaux sont entrain de voir le jour. L'EEG mesure l'activité électrique du cortex cérébrale et est utilisé pour déterminer les états mentaux. Cette donnée permet d'avoir une information sur la fatigue et l'hypovigilance. La figure 1.9 montre le capteur EMOTIV Epoc+ qui est un capteur EEG sans fil composé de 14 canaux. Il permet l'acquisition des données EEG brutes en utilisant le software fourni.



FIGURE 1.9 – Capteur sans fil EEG à 14 canaux proposé par EMOTIV Epoc+.

Rythme cardiaque (ECG) : Ce capteur mesure les changements électriques qui accompagnent le battement cardiaque. L'ECG fonctionne la plupart du temps en détectant et amplifiant les minuscules changements électriques sur la peau. Ces changements électriques sont causés au cours du cycle du muscle cardiaque (battement de cœur). Des études ont montré que les changements du rythme cardiaque sont, par exemple des indicateurs intéressants pour la peur [Levenson, 1992] et la colère [Levenson, 1992], [Kahneman, 1973].

Pression sanguine volumique (BVP) : L'appareil de mesure du signal BVP émet la plupart du temps de la lumière et mesure la quantité de lumière réfléchi par la

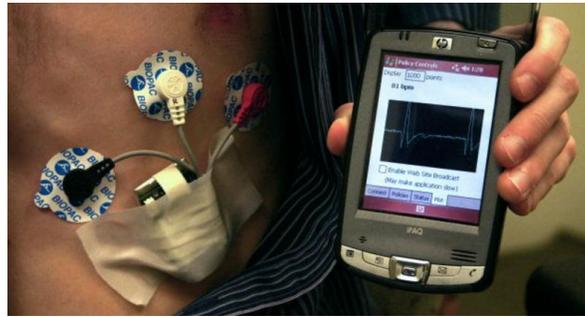


FIGURE 1.10 – Capteur ECG envoyant directement le signal à un dispositif mobile [Jurik et al., 2009].

surface de la peau. Après chaque battement de cœur, le sang est envoyé à travers les vaisseaux sanguins, produisant un engorgement des vaisseaux périphériques sous la source de lumière et modifiant la quantité de lumière réfléchie vers le photo capteur. Par conséquent, la réflectance (coefficient de réflexion) donne une mesure relative de la quantité de sang dans les capillaires à partir de laquelle la fréquence cardiaque et la vasoconstriction peuvent être dérivées. La pression sanguine volumique est donc un indicateur de l'écoulement du sang à travers le corps humain. Il diminue sous l'effort et le stress puisque le sang est détourné vers les muscles qui travaillent afin de les irriguer et les préparer à une action imminente [Abdat, 2010]. Le signal du capteur BVP (voir figure 1.11) se corrèle bien avec les signaux du rythme cardiaque. La fréquence cardiaque peut être déduite du signal BVP. Ce signal est donc, aussi, un bon indicateur de la peur et de la colère.

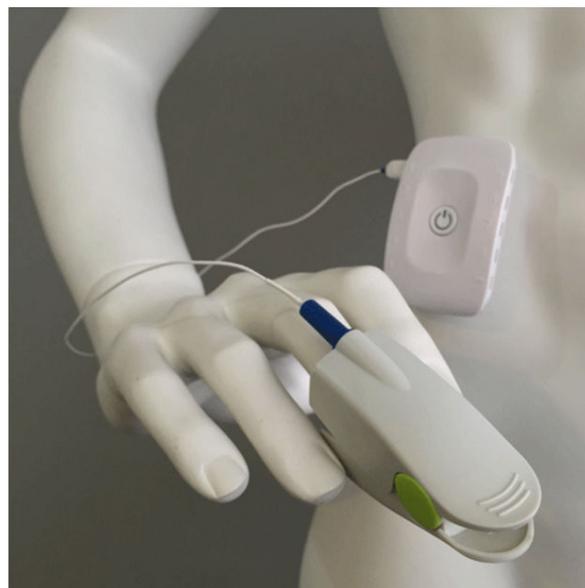


FIGURE 1.11 – Capteur BVP

1.1 Détection des émotions

Volume et rythme respiratoire (RESP) : La mesure de l'échange gazeux des poumons fournit un enregistrement précis du rythme respiratoire. Cette méthode inhibe les activités telles que parler et conduire dans l'environnement naturel. Comme mesure alternative, l'expansion de la cavité thoracique est enregistrée pour capturer l'activité respiratoire. Un capteur à effet Hall composé de deux aimants intégrés à l'intérieur d'un tube élastique est utilisé pour mesurer l'inspiration et l'expiration. L'inspiration étire l'élastique, sépare les aimants et crée un courant et l'expiration permet au capteur de revenir à l'état de base. Ainsi, le rythme respiratoire est donné par l'alternance régulière des mouvements d'inspiration et d'expiration. Cette alternance entraîne ainsi une augmentation ou une diminution du volume de la cage thoracique respectivement à chaque inspiration ou expiration. L'activité physique et l'excitation émotionnelle causent une respiration plus rapide et plus profonde, tandis-que le repos et la relaxation conduisent à une respiration plus lente et moins profonde [Frijda, 1986]. Généralement, des phases de respiration irrégulières sont créées par les émotions à valence négative [Abdat, 2010].



FIGURE 1.12 – Capteur mobile de respiration permettant de capter la respiration ventrale ainsi que thoracique [Baud and Sèdes, 2009].

Le Tableau 1.4 regroupe quelques études sur la détection des émotions en utilisant les signaux physiologiques.

La base de données de référence proposée par Jennifer Healey est souvent utilisée pour comparer différentes techniques de détection des émotions à partir de signaux physiologiques. Cette base de donnée a été conçu avec des acteurs qui jouent des émotions induites par des images, des sons. Le protocole expérimental permettant de construire cette base de données présente des lacunes qui seront expliquées dans le chapitre 5 et qui motivera le fait de créer notre propre base de données. Pour cette nouvelle base de donnée, nous avons utilisé les signaux "EMG, GSR, ECG et EEG".

1.1 Détection des émotions

Authors information	Signals	part.	feat.	sel.	classifiers	target	result (%)
[Fernandez and Picard, 1997]	C ,E	24	5	B-W	HMM,Viterbi	frustration/not	63
[Healey and Picard, 1998]	C ,E ,R ,M	1	11	Fisher	QuadC,LinC	3 emotions anger/peacefulness 2 arousal levels 2 valence levels	87-75 99 84 66
[Healey and Picard, 2000]	C ,E ,R ,M	1	12	SFS	kNN	4 stress levels	87
[Takahashi and Tsukaguchi, 2003]	C ,B	10	12		NN,SVM	2 valence levels	62
[Rani et al., 2003]	C ,E ,M ,S	1	18		FL, RT	3 anxiety levels	59-91
[Herbelin et al., 2005]	C ,E ,R ,M ,S	1	30	reg,LDA	kNN	5 emotions	24
[Takahashi, 2004]	C ,E ,B	12	18		SVM	5 emotions 3 emotions	42 67
[Zhou and Wang, 2005]	C, E, S, and others	32	?		kNN,NN	2 fear levels	92
[Rainville et al., 2006]	C ,R	15	18	ANOVA, PCA	LDA	2 emotions 4 emotions	65 72-83
[Liu et al., 2007a]	C ,E ,M	3	54		SVM	3* 2 levels	85-80-84
[Villon and Lisetti, 2007]	C ,E	40	28		regression model	5 emotions	63-64
[Rani et al., 2007]	C ,E ,M ,S	5	18		FL, RT	anxiety scale?	57-95
[Hönig et al., 2007]	C ,E ,R ,M ,S	24	4*50		LDA,GMM	2 levels of stress	94-89
[Kreibig et al., 2007]	C ,E ,R ,M ,O	34	23		ANOVA PDA	fear, sadness, neutral	69-85
[Liu et al., 2008]	C ,E ,M	6	54		SVM	3*2 levels	81
[Conn et al., 2008]	C ,E ,M ,S	6	?		SVM,QV-learning	3* 2 levels	83
[Cheng and Liu, 2008]	M	1	12	DWT	NN,TM	4 emotions	75
[Benovoy et al., 2008]	C ,E ,R ,S	1	225	SFS,Fisher	LDA,kNN,NN	4 emotions	90

Signals: C : cardiovascular activity; E : electrodermal activity; R : respiration; M : electromyogram; S : skin temperature; O : Expiratory pCO₂.

Classifiers: HMM : Hidden Markov Model; RT : Regression Tree; NN : Artificial Neural Network; SVM : Support Vector Machine; LDA : (Fisher) Linear Discriminant Analysis; kNN : k-Nearest Neighbors; FL : Fuzzy Logic System; TM : Template Matching classifier; QuadC : Quadratic classifier; LinC : Linear classifier; Viterbi : Viterbi decoder

Selection: B-W : Baum-Welch re-estimation algorithm; PCA : Principal Component Analysis; SFS : Sequential Forward Selection; ANOVA : Analysis of Variance; DWT : Discrete Wavelet Transform; Fisher : Fisher projection; PDA : Predictive Discriminant Analysis.

part.: number of participant; **feat.:** number of feature; **result:** classification rate; **sel.:** selection/reduction

TABLE 1.4 – Quelques études sur les systèmes de classifications des états affectifs "affects" utilisant les signaux physiologiques : Tableau extrait de [van den Broek et al., 2011]

1.1.3 Synthèse

L'étude bibliographique sur la problématique de la définition et de la représentation des émotions nous a permis de montrer que la notion d'émotion est complexe à définir. En effet, plusieurs travaux ont été effectués dans ce domaine et la plupart des psychologues donnent des définitions différentes suivant l'approche de représentation qu'ils soutiennent. Nos travaux s'appuient sur l'approche catégorielle de représentation des émotions. Cette approche permet une représentation vectorielle de toute activité émotionnelle humaine. Ainsi comme le soutenait Descartes, les émotions ressenties par un humain sont rarement basiques, cette approche vectorielle développée par Dr. Tayari

1.1 Détection des émotions

[Tayari, 2013] permet de remonter aux émotions complexes qui ne sont que des fusions entre des émotions primaires [Plutchik, 1980].

D'autre part, une activité émotionnelle est caractéristique d'un changement de comportement, d'attitude, etc. Par exemple, sous l'emprise de la colère, involontairement, nous devenons pâles (enfin, pas tous! car cette caractéristique de la colère dépend aussi de l'origine ethnique), nous tremblons des fois, transpirons plus, notre rythme cardiaque s'accélère, le débit de la parole est modifié, etc. De ce fait, les émotions humaines peuvent être détectées en utilisant les expressions du visage, de la voix et des signaux du système nerveux autonome (signaux physiologiques). Dans nos travaux, nous avons choisi d'utiliser la modalité des signaux physiologiques pour la détection des émotions pour trois raisons principales :

1. L'avancée des nouvelles technologies permet de se doter de dispositifs électroniques de plus en plus compactes et intégrés. Cette avancée permet notamment d'avoir des capteurs de signaux physiologiques de moins en moins invasifs tout en garantissant une bonne qualité des signaux.
2. Les signaux physiologiques sont faciles à acquérir contrairement aux signaux des deux autres modalités (signaux faciaux et vocaux). Une fois les capteurs connectés sur une personne, les données sont acquises sans contrainte supplémentaire rendant ainsi possible une mobilité dans l'acquisition de ces signaux. De plus, la gêne occasionnée par les électrodes diminue car les capteurs peuvent être de plus embarqués dans nos montres, nos bracelets, nos chaussures, etc.
3. Les signaux physiologiques appartiennent au système nerveux autonome. Ils ne sont pas sous contrôle volontaire de l'être humain. Grâce à cette caractéristique des signaux physiologiques, la reconnaissance des émotions devient plus fiable qu'avec les expressions vocales ou faciales où nous pouvons facilement leurrer une émotion en jouant sur l'intonation de notre voix ou sur les traits du visage.

David Vandyke [Vandyke, 2013] au-cours de ses travaux a montré l'impact de la dépression dans notre société. En effet, les troubles dépressifs sont parmi les raisons les plus importantes de l'invalidité dans le monde entier. Ils limitent les capacités physiques et morales des patients souffrant de cette maladie affectant ainsi leurs familles et leurs amis. Les pertes pour les entreprises à cause de cette maladie sont considérables et leurs prises en charge sont très coûteuses. Il est donc nécessaire de développer des outils permettant de prévenir un état dépressif. La dépression est caractérisée par le ressenti continu et successif sur plusieurs jours d'émotions négatives. La prévention de cette maladie passe alors par la conception d'outil favorisant la détection des émotions humaines. De même dans le cadre de patients atteints de la maladie d'Alzheimer, la détection des émotions peut servir à la mise en place de thérapie non médicamenteuse comme la mnémothérapie musicale.

1.1.4 Conclusion

Dans cette section, nous avons présenté, dans un premier temps, un état de l'art sur la problématique de la notion d'émotion. En partant de la problématique définitionnelle de l'émotion, aux différentes approches de représentations des émotions jusqu'aux rôles des émotions dans nos interactions sociales, nous avons pu constater toute la complexité de cette notion d'émotions. Puis, dans un second temps, nous avons présenté les travaux effectués sur la reconnaissance des émotions en utilisant les expressions vocales, faciales et les signaux physiologiques. Dans les sections suivantes, nous présentons un état de l'art succinct sur les techniques d'exploration d'architecture d'objets communicants pour la santé. Ensuite, la problématique de la consommation énergétique dans de tels systèmes est abordée. Cette étude vise à définir un système de reconnaissance des émotions ayant une grande autonomie et une grande précision quant à la détection des émotions.

1.2 Exploration d'architecture d'objets communicants pour la santé

Les progrès effectués dans le domaine des nouvelles Technologies de l'Information et de la Communication ont permis de concevoir des systèmes électroniques variés, dotés de capacités de communication (réseaux de transports d'information sans fil, RFID, WSN ...), de détection et/ou d'activation, de traitements diverses. Ces systèmes appelés aussi objets connectés sont de plus en plus déployés et font partie de notre vie de tous les jours. Leur acceptation exigent qu'un certain nombre de contraintes soient satisfaites comme par exemple :

- la taille du système afin qu'il soit le moins encombrant possible ;
- l'autonomie qui doit être adaptée aux applications visées et à l'acceptabilité des utilisateurs. L'idéal serait de disposer de systèmes autonomes.
- la qualité de service qui doit répondre au mieux aux exigences applicatives.

La prise en compte de ces contraintes nécessitent donc de définir de nouvelles approches de conception à un haut niveau d'abstraction et de considérer à la fois les aspects acquisition, transmission et traitements de données.

Les défis scientifiques soulevés par cette croissance des objets connectés peuvent se diviser en deux principaux domaines :

1. **Matériel** : Dans ce domaine, la problématique consiste à développer de nouveaux systèmes microélectroniques communicants et intelligents. De nouvelles approches de conception globale des objets devront être développées avec pour objectif de satisfaire les trois contraintes citées dans le paragraphe précédent.
2. **Modèles et Logiciels** : Dans ce domaine, il faudra développer de nouvelles approches innovantes relatives aux interactions Homme-Machine, à

1.2 Exploration d'architecture d'objets communicants pour la santé

la connectivité de plusieurs objets communicants, à la sécurité et au respect de la vie privée, etc.

La figure 1.13 montre un exemple de procédure de conception d'un tel système. En partant d'un scénario applicatif, le concepteur doit composer avec toutes les contraintes liées aux exigences applicatives et aux besoins utilisateurs en s'appuyant sur des réseaux de capteurs ou des objets connectés adaptés.

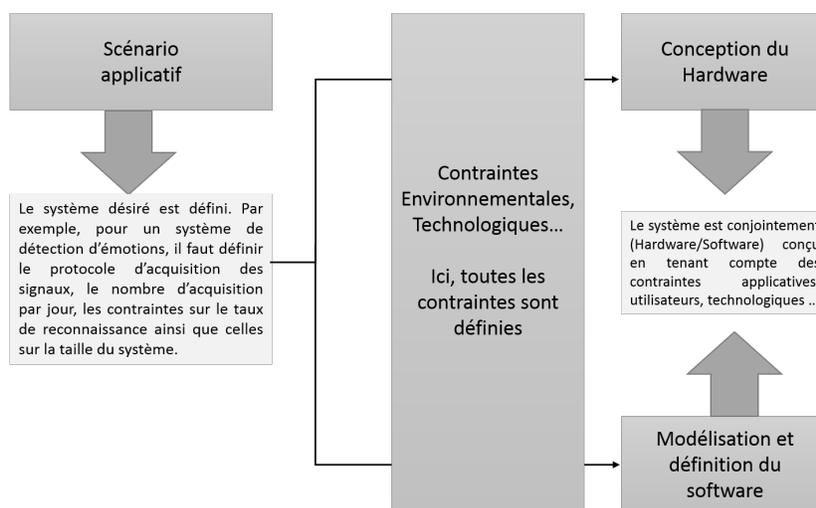


FIGURE 1.13 – Processus de conception d'un système. Cette figure est extraite et adaptée de [Carnot, 2011]

Dans la suite de ce manuscrit, nous décriront quelques travaux utilisant les objets communicants pour la santé, puis nous présenterons quelques techniques d'exploration architecturale.

1.2.1 Objets communicants pour la santé

De nos jours, il suffit de regarder autour de nous pour nous rendre compte de la multitude d'objets communicants qui nous entourent. Diverses applications pour la santé existent et bien d'autres sont en-cours de développement. A titre d'exemple, certaines de ces applications permettent de réaliser des diagnostics à distance (Télémédecine), d'effectuer un suivi médical, de fournir une assistance aux personnes âgées, handicapées moteur ou souffrant de troubles de l'audition ou de la vue. Ces innovations ont pour objectif de couvrir les besoins humains d'assistance, de suivi, de contrôle et de faciliter l'autonomie. En effet, dans le long terme, ces systèmes devraient permettre d'une part aux personnes handicapées et/ou âgées d'améliorer leur autonomie dans leurs activités quotidiennes. D'autre part, elles aideront les médecins dans les diagnostics de certaines maladies/pathologies, dans la prise en charge des bilans de santé et aideront les proches du patient dans la prise en charge. Pour favoriser l'acquisition d'une certaine autonomie pour des activités quotidiennes, ces systèmes

1.2 Exploration d'architecture d'objets communicants pour la santé

d'objets communicants devront être sensitifs, anticipatifs, reconfigurables et simples d'utilisation.

Plusieurs applications innovantes pour la santé ont été développées ou sont en cours de développement. Ces applications seront d'une grande aide pour toute la population (handicapée ou pas), les patients, leurs familles et les médecins. À titre d'exemple, nous avons le **smart body analyser ws-50 de Withings**. Cette balance (voir figure 1.14) est capable de mesurer le poids, le rythme cardiaque, la masse grasseuse, l'état physique de l'utilisateur dans sa globalité ainsi que la qualité de l'air et la température ambiante. En cela, elle permet à l'utilisateur de contrôler l'évolution de son corps.



FIGURE 1.14 – smart body analyser ws-50 de Withings

En 2012, le docteur Jean-Marie Castellucci a mis en place la première application mobile de diagnostics à distance **DocForYou**. À partir d'une douleur (mal de gorge par exemple), l'application recherche par le biais d'un questionnaire interactif, ludique, d'autres symptômes clés qui permettent de déterminer la ou les maladies les plus probables.

Autre exemple, le premier essai d'un implant cochléaire (dispositif servant à donner un certain niveau d'audition aux personnes ayant des déficiences auditives non totales : surdité moyenne, surdité sévère et surdité profonde) sur une personne fut réalisé en 1957 par A. Djourno, professeur de physique médicale et C. Eyriès, otologiste. Depuis, des améliorations technologiques à ces implants ne cessent d'être apportées afin de les rendre plus performants et moins invasifs. Il est placé, chirurgicalement, derrière l'oreille et se compose principalement de deux parties à savoir :

- Une partie externe composée d'un microphone voire d'un processeur et d'un transmetteur.
- Une partie interne composée d'un stimulateur, récepteur et d'un certain nombre d'électrode pouvant atteindre 24. La figure 1.15 montre les différents composants d'un tel système.

Shafiee et ces collègues [Shafiee et al., 2015] ont développé une plateforme de bio-capteurs intégrant une application mobile. Le dispositif est composé d'une bandelette fine, légère et souple élaborée à partir de cellulose et de films de polyester, capable de

1.2 Exploration d'architecture d'objets communicants pour la santé

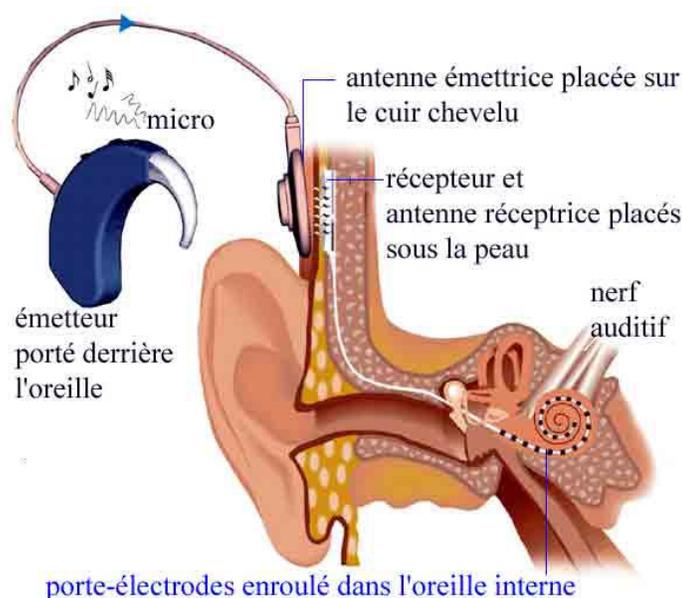


FIGURE 1.15 – Schéma de l'implant cochléaire DIGISONIC [Ménard, 2008]

détecter des agents biologiques dans le sang et la salive. Cette plateforme permet de détecter le **VIH**, la bactérie **E.coli** et le **staphylocoque doré**.

La collaboration de la Fondation Suisse pour les Cyberthèses (FSC) avec l'École Polytechnique Fédérale de Lausanne (EPFL) a permis la mise en œuvre des dispositifs de rééducation motrice et d'assistance à la marche, associant orthèse et électrostimulation transcutanée rétro-contrôlée. L'orthèse est un appareil compensant une fonction absente ou déficiente. Elle assiste une structure articulaire ou musculaire, stabilise une partie du corps pendant une phase de réadaptation ou de repos. La cyberthèse est une attelle mobile constituant un *exosquelette* de soutien et de guidage de la mobilité déficiente d'un membre couplée à une commande cybernétique du système et une stimulation neuromusculaire électrique réglée en boucle fermée CLEMS (Closed-Loop Electrical Muscle Stimulation) [FSC, 2017]. Ces dispositifs de rééducation sont destinés aux personnes ne pouvant plus faire usage de leur membres inférieurs [Métrailler, 2005]. Ce système se compose de trois appareils (MotionMaker - WalkTrainer - WalkMaker) dont chacun est utilisé à une phase spécifique dans le processus de rééducation des membres paralysés. Le premier dispositif, MotionMaker conçu pour les entraînements thérapeutiques est un système de mobilisation des membres inférieurs et de renforcement musculaire. Une fois que le patient atteint une bonne mobilité dans ses membres, le second appareil WalkTrainer est utilisé dans l'objectif de verticaliser le patient paraplégique en vue de la marche. Enfin le troisième est un appareil d'assistance à la marche consistant en une orthèse légère, ne disposant plus de moteurs aidant ainsi le patient à retrouver une marche autonome. La figure 1.16 montre les 3 appareils d'aide pour la rééducation.

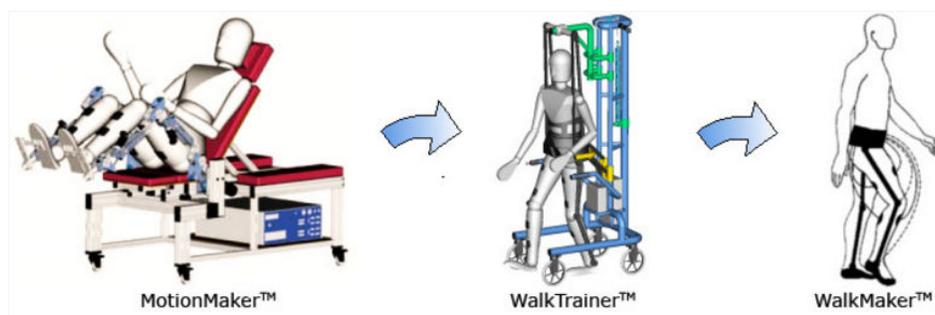


FIGURE 1.16 – Système d’aide des paraplégiques basé sur trois appareils pour la rééducation

1.2.2 Techniques d’exploration d’architecture

De manière générale, la complexité des traitements effectués dans les systèmes actuels ne cessent de croître. À cet effet, la mise en œuvre d’architectures satisfaisant les contraintes sur la gestion de la mémoire, de la communication et de la consommation semble être une évidence. Ce phénomène devient encore plus vrai quand les applications visées sont dans le domaine médical car ce domaine a de grandes exigences sur la taille du système, sur son efficacité ainsi que sur l’autonomie. L’exploration architecturale permet de coupler les contraintes matérielles, applicatives et utilisateurs afin de proposer le système le plus adapté possible. En d’autre terme, nous pouvons donc affirmer que l’exploration architecturale consiste à identifier les architectures potentielles pouvant satisfaire l’ensemble de contraintes données. Cette identification est un processus itératif, conduisant à comparer les performances de l’ensemble des solutions possibles [Majdoub, 2012].

Dans le chapitre 4, nous allons proposer une exploration architecturale et de configuration couplant autonomie et taux de reconnaissance des émotions. Pour notre exploration, nous évaluons quelques architectures en termes de consommation énergétique et de taux de reconnaissance. Ensuite, nous choisissons celle qui remplit au mieux toutes les contraintes à savoir : faible autonomie, dimensions et poids du système d’acquisition réduits, taux de reconnaissance des émotions satisfaisants.

1.3 Consommation énergétique dans les systèmes embarqués médicaux

La consommation des systèmes intégrés est, déjà, depuis quelques années une contrainte importante à prendre en compte lors de la conception. Pour répondre à cette problématique, différents axes de recherche ont été développés qui visent, principalement, à modéliser la consommation énergétique et à proposer des solutions statiques et dynamiques, logicielles/matérielles permettant de réduire cette consommation. En effet, cette problématique est d’autant plus importante

1.3 Consommation énergétique dans les systèmes embarqués médicaux

pour les applications mobiles notamment pour des paramètres comme l'autonomie, l'encombrement de la batterie... que dans les applications non mobiles pour le coût du système. D'un côté, les concepteurs de batteries tentent d'augmenter les capacités des batteries au mieux et de l'autre côté, les besoins en consommation ne cessent d'augmenter. Comme le mentionne le rapport de l'IEA (International Energy Agency), la consommation mondiale sera multipliée par dix en 2030. De ce fait, il est nécessaire de développer et de mettre en place des techniques permettant de garantir des systèmes à basse consommation voire même autonomes.

Dans le cadre de la thèse, l'objectif est de comparer, différents traitements effectués dans un système de détection des émotions utilisant des capteurs de signaux physiologiques, d'un point de vue des performances et de la consommation énergétique. Le but est de coupler l'autonomie et la performance du système afin de proposer un système médical efficace en tout point de vue. De ce fait, il semble primordial de gérer la consommation d'un tel système à tous les niveaux d'abstractions si possible.

Pour proposer notre système de détection des émotions, il est nécessaire d'étudier les contraintes avec lesquelles nous devons composer. La section qui suit présente une liste non exhaustive des contraintes sur les systèmes d'acquisition des signaux physiologiques. Ensuite les différentes sources de consommation dans les circuits, les techniques déjà existantes dans la littérature permettant de réduire cette consommation ainsi que les modèles de décharge de la batterie sont introduites. Enfin, quelques techniques de modélisation de la consommation énergétique sont brièvement décrites. Les sections qui suivent visent à introduire la problématique liée à la modélisation de la consommation d'énergie pour cerner les techniques que nous pourrons par la suite utiliser pour modéliser fidèlement la consommation de notre système.

1.3.1 Quelles contraintes sur les systèmes d'acquisition de signaux physiologiques ?

Le nombre de systèmes d'objets communicants déployés autour de nous et dans nos quotidiens ne cesse d'exploser grâce à la grande avancée des nouvelles Technologies d'Information et de Communication. Cette avancée facilite la création d'un système d'objet communicant pour la détection des émotions.

Les capteurs physiologiques existent en très grands nombres et sont d'une grande diversité. Ils partent des plus invasifs (dispositif lourd, peu autonome, peu ludique) aux moins invasifs se trouvant actuellement dans les smart watches, chaussures, chemises, casquettes, etc. Ces capteurs permettent l'acquisition des signaux ECG, EMG, EEG, GSR, Respiration, etc.

D'autres part, les chercheurs travaillant sur la reconnaissance des émotions utilisant des capteurs physiologiques ou pour d'autres applications liées à la santé (monitoring des signaux, l'e-santé, diagnostics de certaines pathologies, etc) sont confrontés à un épineux problème. En effet, les applications médicales sont très contraignantes car elles se veulent efficaces (signaux pertinents) avec un dispositif non invasif (faibles poids et

1.3 Consommation énergétique dans les systèmes embarqués médicaux

dimensions) et ayant une faible consommation énergétique. Cela nous amène à nous poser ces quelques questions :

Est ce que les capteurs les moins invasifs permettent d'avoir des signaux assez pertinents ?

Sinon quels compromis peut on faire entre un système plus ou moins invasifs et la pertinence des signaux ?

Quelle est l'autonomie de tels systèmes ? et peut on en créer un réseau à faible échelle en vue d'optimisation multicritères qui permet d'améliorer le taux de reconnaissances des émotions ?

Pour l'application que nous visons, notre objectif est donc de créer un réseau de capteurs de signaux physiologiques à faible échelle dont les signaux sont pertinents, le moins invasif possible, ayant une grande autonomie, permettant une acquisition en temps réel, facilitant l'extraction des données acquises, etc. Dans le chapitre 3, nous allons proposer une architecture de systèmes d'acquisition des signaux biologiques remplissant ces contraintes.

1.3.2 Origines de la consommation

Parmi les circuits logiques de base, réalisés par un circuit CMOS, on peut citer l'inverseur. Ce dernier est constitué d'un transistor de type P et d'un autre de type N (voir Figure 1.17). Cette technologie d'agencement de transistors CMOS P et N est utilisée dans plusieurs circuits à cause d'une part de son encombrement très réduit et de sa facilité de mise en œuvre. L'agencement de transistors de type N et P permet ainsi de réaliser des fonctions par exemple logiques, des bascules, éléments de base des systèmes de traitement comme les processeurs. Afin d'analyser les sources de consommation, considérons dans un premier temps un inverseur. Cet inverseur présente un courant dit statique en présence d'inactivité alors que lors de son utilisation, des courants de court-circuit et de commutation apparaissent. L'ensemble de ces courants traduisent la consommation énergétique du circuit. De ce fait on peut distinguer deux types de consommation dans les circuits : **la consommation dynamique et la consommation statique**.

Chacune de ces puissances dissipée est décrite plus en détail dans les paragraphes suivants.

1.3.2.1 Puissance statique

La consommation statique constitue la puissance de repos. Elle est due au fait que même si les transistors sont bloqués, ils ne sont pas complètement fermés. La perte de courant de fuite peut augmenter de façon significative lorsque la tension d'alimentation (tension de seuil) est très basse. Cette consommation statique peut être négligée pour des circuits ayant une très grande consommation dynamique. Elle est donnée par l'équation suivante où V_{dd} représente la tension d'alimentation, I_{leak} définissant le courant de fuite traversant les transistors complémentaires et N_{tr} qui

1.3 Consommation énergétique dans les systèmes embarqués médicaux

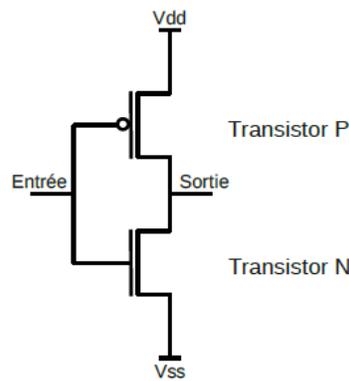


FIGURE 1.17 – Inverseur réalisé à l'aide de transistors CMOS

traduit le nombre de transistor inactifs dans le circuit.

$$P_{Statique} = V_{dd} * N_{tr} * \sum I_{leak} \quad (1.1)$$

Dans cette équation, on remarque une dépendance de la puissance statique à la tension d'alimentation, donc une solution pour réduire cette puissance serait d'abaisser la tension V_{dd} ou de couper l'alimentation si le circuit n'est pas utilisé.

La puissance statique est générée par la présence de trois principaux type de courants de fuite au niveau des transistors inactifs. Les origines et les mécanismes de ces courants de fuite sont expliqués dans [Roy et al., 2003]. Le lecteur intéressé par les mécanismes des courants de fuite pourra se référer à cet article [Roy Kaushik et al.]. Ces courants de fuite sont entre autres :

1. Courants sous la tension de seuil : Ces courants prennent naissance lorsque la tension appliquée entre la grille et la source (V_{GS}) des transistors CMOS est inférieure à une tension de seuil (V_{th} exprimée en volt). Ce courant de fuite dont l'unité est l'ampère (A) est exprimée par l'équation 1.2. Cette équation est expliquée en détail dans [Liu et al., 2007b]. Dans l'équation 1.2 W et L dont les unités sont le mètre correspondent respectivement à la largeur et la longueur du canal (espace entre le drain et la source du transistor), k la constante de Boltzmann, T la température (K), q la constante de charge d'un électron (exprimée en coulomb : C) et n un coefficient dépendant des caractéristiques technologiques du transistor.

$$I_{leak} \approx \frac{W}{L} \left(\frac{kT^2}{q} \right) e^{q \frac{(V_{GS} - V_{th})}{nkT}} \quad (1.2)$$

2. Courants tunnels au niveau de l'oxyde de grille : Des courants de fuite se créent lors du passage des électrons entre le substrat et la grille. Ces courants sont appelés "courants tunnels". A cause de la réduction des épaisseurs du transistor en vue de la miniaturisation de ces derniers, ces courants de fuite ne cessent d'augmenter.

1.3 Consommation énergétique dans les systèmes embarqués médicaux

3. Courants de fuite au niveau des jonction PN des transistors : Ce courant comme son nom l'indique vient de la jonction PN entre le drain et le substrat. Ce courant de fuite est donné par l'équation 1.3 où I_s représente le courant de saturation (dont l'unité est l'Ampère) et V_p la tension de polarisation (exprimée en volt). Cette équation est expliquée dans [Kang et al., 2014].

$$I_{leak} = I_{PN} \approx I_s e^{\left(\frac{qV_p}{kT} - 1\right)} \quad (1.3)$$

1.3.2.2 Puissance dynamique

La consommation dynamique est due au courant de court-circuits (les transistors P et N sont passant simultanément), au courant créé lors de la commutation des transistors, et ceci lors de la charge et la décharge des capacités de charge équivalente. Elle est traduite par l'équation 1.4 où P_{sw} (Watt) représente la puissance dissipée lors des commutations des transistors et P_{cc} (Watt) représente celle dissipée lorsque les transistors sont en court-circuit.

$$P_{dynamique} = P_{sw} + P_{cc} \quad (1.4)$$

En effet, du fait de la latence de commutation des transistors, un courant de court-circuit se crée pendant ce temps de latence. Cette puissance moyenne de court-circuit est exprimée par l'équation 1.5 avec K une constante dépendant des paramètres électriques, τ le temps de montée/descente du signal d'entrée (seconde), F la fréquence d'horloge (Hertz), V_{cc} la tension d'alimentation, V_{th} la tension de seuil. V_{cc} et V_{th} sont exprimées en Volts.

$$P_{cc} \approx k\tau F (V_{cc} - 2V_{th})^3 \quad (1.5)$$

Certaines commutations des transistors sont nécessaires pour le fonctionnement du circuit. D'autres, communément appelés "glitches" surviennent quand le circuit est actif et passe par des états transitoires. Ces commutations involontaires et non requises sont produites lorsque des signaux présents sur les entrées d'un circuit combinatoire possèdent des temps d'arrivée différents, induits notamment par les temps de propagation des portes situées en amont [Lorandel, 2015]. La puissance de commutation est traduite par l'équation 1.6 où V_{dd} (Volts) est la tension d'alimentation du circuit, F la fréquence (Hertz), α le taux de transitions montantes (passage de 0 à 1) par rapport à la fréquence de base et C la totalité des capacités parasites (Coulomb).

$$P_{sw} = V_{dd}^2 * F * \alpha * C \quad (1.6)$$

Cette équation 1.6 nous montre la dépendance quadratique de la puissance dynamique à la tension d'alimentation, de ce fait, pour avoir un système à faible consommation, tous les concepteurs de circuits intégrés jouent en premier sur ce paramètre très important. Toutefois, réduire ce dernier exige un compromis entre

1.3 Consommation énergétique dans les systèmes embarqués médicaux

tension d'alimentation et fréquence de fonctionnement, car plus la tension diminue, plus la commutation des transistors est ralentie donc la fréquence diminue.

Dû au fait que la puissance de commutation est très grande devant la puissance de court-circuit, la puissance dynamique peut être définie par l'équation 1.7.

$$P_{dynamique} \approx V_{dd}^2 * F * \alpha * C \quad (1.7)$$

La puissance totale consommée dans un circuit est la somme de la puissance dynamique et de la puissance statique :

$$P_{Totale} = P_{Dynamique} + P_{Statique} = V_{dd}^2 * F * \alpha * C + V_{dd} * N_{tr} * \sum I_{leak} \quad (1.8)$$

Ainsi on obtient

$$P_{Totale} = (V_{dd} * F * \alpha * C + N_{tr} * \sum I_{leak}) * V_{dd} \quad (1.9)$$

Dans cette section, nous avons vu les différentes équations traduisant la consommation énergétique d'un système qui est due aux puissances dynamiques et statiques. Ces équations nous serviront d'éléments de base dans le chapitre 3, pour modéliser la puissance dynamique et statique des différents modules (capteurs, radio, et de traitement) de la plateforme d'acquisition choisie. Ensuite, sur la base ces modèles, la puissance consommée par l'architecture de collecte de donnée et de traitement pourra être estimée. En s'appuyant sur une bibliothèque de composants, on pourra ainsi rechercher un système présentant une consommation optimale.

1.3.2.3 Quelques techniques de réduction de la consommation énergétique

Les deux équations qui représentent la consommation dynamique et statique, données dans la section précédente, nous permettront, dans une première phase, d'estimer la consommation en identifiant les sources principales, puis de rechercher des réductions de consommation. De nombreuses techniques visant à réduire la consommation sont basées sur l'adaptation de la tension et de la fréquence en fonction des exigences applicatives. En dehors de ces techniques, l'application de ces deux équations peut être très complexe. En effet comme mentionné dans [Zhang et al., 2003], par exemple, la puissance statique dépend de la température qui va elle même être liée à l'activité du circuit. Cette puissance dépend aussi de paramètres statique bas niveau comme le nombre de transistors (N_{tr}) ou le taux de transition (α). Pour la puissance dynamique, des paramètres comme l'activité du signal et la capacité parasite propre à ce signal [Bonamy, 2013] interviennent aussi.

L'estimation de la consommation ainsi que les techniques d'optimisation peuvent se faire sur chacun des cinq niveaux d'abstraction (voir figure 1.18) à savoir le niveau :

- **Technologique** permettant de considérer des paramètres comme la taille des transistors...;
- **Circuit/Logique** en faisant du clock gating, de la logique asynchrone...;
- **Architectural** qui consiste à faire du parallélisme, du pipeline ...;

1.3 Consommation énergétique dans les systèmes embarqués médicaux

- **Algorithmique** qui joue sur la transformation des boucles, la mise en mémoire des données...;
- **Système** qui consiste à mettre en place, des stratégies, au sein du système d'exploitation d'ordonnancement, de partitionnement en fonction de la charge de travail.

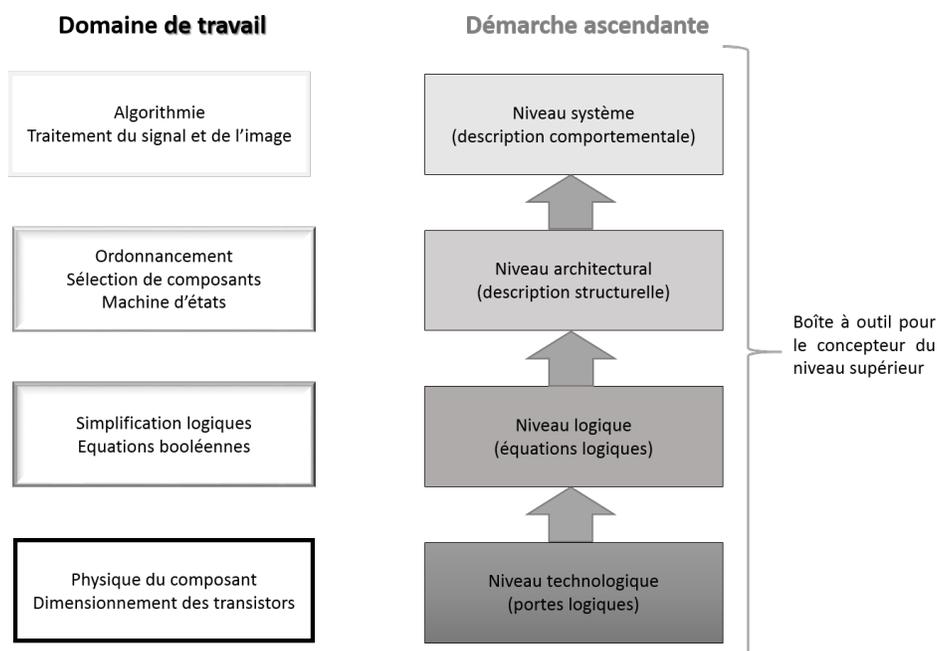


FIGURE 1.18 – Les principaux niveaux d'abstraction. Sur cette figure, on ne recense que les 4 principaux niveaux car les niveaux algorithmique et système sont fusionnés. Cette figure est tirée du cours de Bertrand Granado en master ESA à l'université de Cergy Pontoise

Ci-dessous, nous décrivons brièvement quelques techniques de réduction de la consommation énergétique dans un circuit.

1. **Power gating** : Cette technique consiste à couper l'alimentation d'une partie du circuit lorsqu'elle n'est pas utilisée, ce qui permet d'éliminer les puissances statiques. En effet dans l'équation 1.9, si $V_{dd} = 0$ alors $P_{Totale} = 0$. Cependant, lorsqu'on coupe la tension d'alimentation, il faut sauvegarder toutes les données contenues dans les bascules pour ne pas les perdre. Cette sauvegarde des données entraîne des courants supplémentaires. Pour réduire ces courants supplémentaires, il peut être préférable de baisser la tension d'alimentation (Voltage scaling) au lieu de couper complètement l'alimentation. La réduction de la tension d'alimentation résulte en une amélioration quadratique de la dissipation de puissance [Sentieys, 1997]. Plus V_{dd} diminue, plus le temps de

1.3 Consommation énergétique dans les systèmes embarqués médicaux

propagation augmente de façon exponentielle, il ne faut donc pas que V_{dd} soit trop proche de la tension de seuil.

2. **Clock gating** : Cette technique utilisée pour tous les outils de synthèses consiste à stopper la propagation de l'horloge dans une partie du circuit intégré lorsque celle-ci est inactive. Elle n'agit que sur la consommation dynamique en stoppant toute l'activité d'une partie du dit circuit.
3. **Réduction de la capacité effective** : Afin de réduire cette capacité, une première idée serait de réduire la taille des transistors. La diminution du nombre de Glitches, qui sont des transitions dynamiques parasites, permet de réduire considérablement la consommation de courant.
4. **Frequency scaling (adaptation la fréquence du circuit)** : Cette technique, qui consiste à adapter la fréquence du circuit en fonction de la charge de travail, permet de diminuer les puissances P_{pic} et P_{moyen} . L'énergie E reste inchangée ce qui permet de limiter la dissipation thermique et de maîtriser le pic de puissance.
5. **Utilisation de plusieurs niveaux de tensions** : Ceci consiste à utiliser plusieurs tensions dans le but d'avoir différents modes de fonctionnement (Idle, Repos, fonctionnement . . .). Elle permet de réduire la consommation globale du circuit en choisissant la tension d'alimentation optimale suivant l'utilisation du circuit.

Les travaux [Rethinagiri, 2013], [Mbarek, 2013], [Jadoon, 2013] expliquent plus en détails les différentes techniques d'estimation et ou d'optimisation de la consommation en partant du niveau technologique jusqu'au niveau système.

Certaines de ces techniques de réduction de la consommation sont très bas niveaux. Les plateformes que nous utilisons dans ces travaux sont des plateformes de développement présentes sur le marché. Nous ne pouvons pas avoir une description matérielle à bas niveaux de ces plateformes. A cet effet, nous n'utiliserons pas les bas niveaux d'abstractions. Mais nous pourrons, entre autre, utiliser des techniques "d'utilisation de plusieurs niveaux de tensions" en jouant sur les différents modes de fonctionnement et des techniques de "Frequency scaling" afin d'aboutir à une utilisation optimale de notre système.

1.3.3 Décharge de la batterie

Une batterie est un appareil utilisant un processus électrochimique pour emmagasiner de l'énergie électrique, elle peut recevoir (charge) ou fournir (décharge) de l'électricité. Les concepteurs des batteries sont soumis tous les jours à une contrainte très importante qui consiste à faire un compromis entre la capacité des batteries et leur encombrement (poids, taille). En raison des contraintes liées à notre application, ce compromis est aussi le notre et sera étudié dans la suite du document. La capacité d'une batterie peut être définie comme la quantité d'énergie qu'elle est capable de délivrer à un circuit. Elle est exprimée en Ampère-heure (Ah) ou Milli-ampère-heure

1.3 Consommation énergétique dans les systèmes embarqués médicaux

(mAh) (1 Ah = 3 600 coulombs) ; une batterie ayant une capacité de **X Ah** est capable de délivrer un courant **I = X Ampère pendant une heure**.

Une décharge profonde de la batterie influence beaucoup le mécanisme de décharge de la batterie. En effet ces décharges profondes réduisent la durée de vie d'une batterie car plus la décharge est profonde, moins grande sera la tension de la batterie après recharge. L'énergie totale C fournie par une batterie idéale est donnée par la loi empirique de Peukert [Linden, 1995].

$$C = I^n * T \tag{1.10}$$

où n est une constante propre à la batterie devant être déterminée empiriquement. Cette constante varie généralement entre 1,1 et 1,4. T (Seconde) représente le temps de décharge et I (Ampère) le courant de décharge.

La loi de Peukert comme les autres modèles empiriques proposent une modélisation simplifiée de l'énergie totale de la batterie. Pendant les temps de repos, elle ne décrit pas le comportement de récupération et elle ne prend pas non plus en compte les effets possibles de variations d'intensité. Face à ces problèmes d'imprécision, des modèles cinétiques, abstraits et intermédiaires ont été proposés dans la littérature [Berthier, 2012]. Les modèles cinétiques prennent en compte les caractéristiques physico-chimiques des constituants des batteries et les propriétés de son environnement. Les modèles abstraits et intermédiaires, quant à eux, décrivent le comportement des batteries à partir d'une représentation analytique construite à partir de circuits électriques équivalents. Ces modèles ont l'avantage d'être plus précis que les modèles empiriques mais ont l'inconvénient d'être beaucoup plus complexe.

Cette discussion autour de la modélisation de la décharge d'une batterie nous permet de mieux comprendre les différents aspects liés à l'autonomie d'un système. Il est nécessaire de prendre en compte le poids et la taille de la batterie pour un système non encombrant. Pour améliorer l'autonomie et la durée de vie du système, la prise en compte de la loi de la décharge et de la durée de vie de la batterie est requise. Dans le chapitre 3, nous allons étudier la décharge et le choix optimal de la batterie pour notre contexte applicatif.

1.3.4 Modélisation et estimation de la consommation énergétique

L'estimation de la consommation énergétique est un véritable défi pour tout concepteur de systèmes afin de supprimer ou de diminuer considérablement les coûts de développement de ses systèmes qui visent à satisfaire les contraintes énergétiques. Pour pallier à ce problème, les concepteurs tentent d'estimer la consommation au plus tôt durant le cycle de développement d'un système. En effet, l'estimation commence au niveau système et gagne en précision au fur et à mesure que nous descendons dans les niveaux d'abstraction. La connaissance du système est plus précise aux bas niveaux d'abstraction mais ces niveaux exigent un temps de simulation plus conséquent. Ainsi,

1.3 Consommation énergétique dans les systèmes embarqués médicaux

des techniques et des outils d'estimation de la consommation ont été développés par des académiques et des industriels à différents niveaux d'abstraction [Lorandel, 2015]. La fonctionnalité du système est modélisée au niveau "système" à l'aide d'outils tels que Matlab ou encore avec des langages C/C++. La validation de la fonctionnalité du système peut être faite avec des outils de simulation comme la librairie SystemC utilisant les classes C++ ou Simulink par exemple. Dans ce qui suit, nous présentons brièvement quelques outils/travaux traitant de la problématique d'estimation de la consommation à différents niveaux d'abstraction pour les systèmes numériques.

1. Au **niveau technologique**, le concepteur a une parfaite connaissance du système. Par le biais des mesures directes de la consommation sur le circuit ou par des simulations, une estimation précise de la puissance dissipée/consommée peut être obtenue. Les outils de conception/simulation de circuits comme SPICE, IRSIM [Salz and Horowitz, 1989] [Shah, 2009] permettent la simulation au niveau des composants (transistors, résistances, condensateurs, bobines). Cependant, à cause de la complexité des systèmes actuels (systèmes composés de milliards de transistors) et du volume de données à traiter, l'estimation de la consommation à ce niveau est très coûteuse en temps de simulation. Il est donc nécessaire de changer de niveau d'abstraction et d'aller vers des niveaux d'abstractions plus hauts pour accélérer le temps de simulation. Les modèles d'estimation à ces hauts niveaux sont moins précis que ceux du niveau technologique même s'ils restent quand même assez précis pour avoir une estimation intéressante de la consommation de l'architecture complète.
2. Au **niveau logique**, le circuit est décrit comme un ensemble d'éléments logiques et d'interconnexions entre ces derniers. Dans la littérature, il y a principalement deux techniques d'estimation de la consommation à ce niveau.
 - *Méthodes statistiques* : Cette méthode basée sur les simulations consiste à générer des stimuli aléatoires aux entrées du circuit afin de simuler le circuit un grand nombre de fois. Il est quasiment impossible d'estimer la puissance dissipée pour toutes les combinaisons possibles. Néanmoins, ces méthodes permettent d'obtenir des informations sur l'activité du circuit. Pour le lecteur intéressé, plus d'informations sur cette méthode peuvent être trouvées dans les travaux [Kozhaya and Najm, 1997] [Todorovich et al., 2005] [Burch et al., 1993]
 - *méthodes basées sur les probabilités de transition* [Hassan et al., 2009] [Choy and Wilton, 2006] : L'utilisateur fournit les probabilités de transition des entrées. Ces probabilités sont ensuite propagées jusqu'aux sorties du circuit de manière à déterminer l'activité interne de ce dernier. Cette approche ne nécessitant qu'une seule étape de calcul est très efficace pour déterminer la puissance dissipée. De plus, elle ne nécessite pas une connaissance approfondie des stimuli d'entrées.
3. Au **niveau architectural (RTL : Register Transfer Level)**, le système généralement décrit avec les langages de description de matériel comme le VHDL, est représenté par l'intermédiaire de signaux et d'échanges entre les différents

1.3 Consommation énergétique dans les systèmes embarqués médicaux

blocs le constituant. À ce niveau, des modèles par bloc de fonctionnement du système sont généralement mis en œuvre en vue de l'estimation de la consommation. Des techniques utilisant des modèles analytiques afin d'exprimer la puissance dissipée en fonction de tous les paramètres importants sont souvent développés à ce niveau [Najoua et al., 2012] [Sunwoo et al., 2010].

4. Au **niveau système**, c'est la fonctionnalité du système qui est modélisée [Damaševičius and Štuikys, 2007][Desset et al., 2012][Hsu et al., 2011]. A ce niveau d'abstraction, l'estimation vise plutôt à trouver des tendances liées à la consommation, de manière à prendre des décisions sur le choix d'une architecture. À ce niveau, l'estimation de la consommation n'est pas aussi précise que dans les précédents niveaux. La fonctionnalité du système est souvent décrite sous forme de spécifications ou d'algorithmes. Ces derniers eux-mêmes sont représentés de façon spécifique sous forme de tâches (matérielles/logicielles) ou encore sous forme graphique ou schématique. La fonctionnalité du circuit peut aussi être représentée sous la forme d'un certain nombre d'instructions exécutées séquentiellement [Lorandel, 2015]. À ce niveau, plusieurs outils comme *Matlab/Simulink*, *Ptolemy*, *Vista*, *SystemC* peuvent être utilisés en vue de la modélisation du système et l'estimation de la puissance de ce dernier.

Au cours de nos travaux, nous avons utilisé dans un premier temps une technique d'estimation au niveau architectural en proposant un modèle analytique découpant le système en différents blocs de fonctionnement permettant d'estimer la consommation énergétique du système. Ensuite, dans le souci de proposer un framework permettant de faire de l'exploration architecturale, nous avons développé un second modèle en SystemC donnant une tendance quant à l'évolution de la puissance dissipée. Nous détaillons ces modèles plus loin.

1.3.5 Conclusion

Ce chapitre d'état de l'art nous a tout d'abord permis de comprendre la notion d'émotions dans toute sa complexité en analysant les grandes théories dans le domaine de la psychologie traitant cette problématique. Ensuite, cette étude nous a permis de montrer l'importance des signaux physiologiques dans la détection des émotions. Avec un peu d'exercice, on peut facilement cacher nos émotions ou tromper sur notre état émotionnel avec les expressions faciales ou vocales contrairement aux signaux physiologiques. Ces derniers sont sous un contrôle involontaire c'est à dire, à titre d'exemple, que nous ne pouvons pas empêcher notre rythme cardiaque de s'accélérer quant on prend peur. Outre cet avantage, les signaux physiologiques garantissent une acquisition permanente des signaux quand les capteurs sont branchés sans aucune contrainte supplémentaire pour le sujet. En effet, l'acquisition de signaux faciaux exige le sujet à regarder dans l'axe de la caméra.

Dans la deuxième partie, nous avons étudié une toute autre problématique : la conception d'objets communicants pour la santé. Plus précisément, nous nous

1.3 Consommation énergétique dans les systèmes embarqués médicaux

sommes intéressés à l'aspect architectural et aux contraintes à satisfaire. Nous avons ainsi présenté quelques objets communicants pour la santé en mettant un accent particulier sur les contraintes que ces systèmes doivent satisfaire comme la consommation énergétique et les performances. Les études, menées dans la littérature sur la consommation énergétique de tels systèmes, nous ont permis de cerner les contours de cette problématique mais aussi de considérer l'autonomie sous-jacente. Ainsi, notre but étant de proposer un système complet pour la détection des émotions remplissant les contraintes utilisateurs, applicatives et architecturales, nous avons présenté un résumé de l'état de l'art sur la modélisation de la consommation à différents niveaux d'abstraction. Cette étude nous a ainsi permis de positionner nos travaux. En effet, étant donné que notre étude cible un système complet d'objet communicant, nous avons choisi de modéliser ce dernier sur deux hauts niveaux d'abstractions à savoir "architectural et système". Ces deux niveaux d'abstractions nous permettent de gagner en temps de simulation et sur l'estimation de la consommation énergétique qui devra être le plus proche possible de la réalité. Pour y parvenir, nous devons tenir compte de tous les paramètres de l'architecture de détection des émotions. Ensuite, nous avons présenté différentes techniques de réduction de la puissance dissipée dans un circuit en fonction de ces niveaux d'abstraction. Notre objectif était d'identifier éventuellement des stratégies qui pourraient être appliquées par la suite.

Dans le chapitre qui suit, nous décrivons notre méthodologie de détection des émotions humaines en utilisant les signaux physiologiques.

Détection des émotions par des signaux physiologiques

Sommaire

2.1	Introduction	46
2.2	Approche monomodale	47
2.2.1	Acquisition des signaux	48
2.2.2	Apprentissage	50
2.2.3	Détection	57
2.2.4	Choix d'un classifieur	60
2.3	Approche multimodale	66
2.3.1	les différentes techniques de fusion des signaux	67
2.3.2	Résultats et discussions	69
2.4	Apport par rapport à l'existant	72
2.5	Conclusion	72

2.1 Introduction

Les travaux réalisés dans le cadre de cette thèse sur la détection des émotions sont une continuité des travaux d'Imen Tayari [Tayari, 2013]. Ils s'inscrivent dans la thématique de l'intelligence artificielle et plus particulièrement dans le domaine de l'informatique affective. L'informatique affective est un domaine étudiant la possibilité de donner aux machines, la capacité de reconnaître les émotions et pourquoi pas dans un futur plus ou moins proche d'exprimer elles mêmes des émotions.

Contrairement à l'approche dimensionnelle, dans l'approche catégorielle (voir la section 1.1.1.1), les émotions sont dépendantes les unes des autres, ce qui permet d'un point de vue informatique d'affiner considérablement les calculs d'états émotionnels. Plusieurs modèles ont été proposés comme celui de Plutchik (basé sur huit émotions de

2.2 Approche monomodale

base faites de quatre paires opposées : *joie-tristesse*, *acceptation-dégoût*, *peur-colère*, *surprise-anticipation*). Ce modèle considère les associations des émotions opposées comme conflictuelles, donc ne les prend pas en considération [Plutchik, 1980]. Imen TAYARI s'appuie sur ce modèle dans les travaux de sa thèse en prenant en compte ces états conflictuels qui selon elle ont un sens dans la vie réelle, et elle l'explique ainsi : il se pourrait qu'il y ait une personne qui pleure (*tristesse*) et sourit (*joie*) à la fois qui sont des émotions opposées mais existent et ont un sens dans la vie réelle. Cette association donnerait une émotion complexe (*composite*). Dr. Tayari au cours de ces travaux a donc établi un modèle multidimensionnel vectoriel de représentation des émotions. Ce modèle permet de passer des émotions dites fondamentales aux émotions complexes qui sont des mélanges d'émotions basiques d'après le modèle de Plutchik.

Les algorithmes pour l'approche monomodale ont été principalement développés par Dr. Tayari au cours de sa thèse. Elle a proposé les différentes étapes requises dans l'approche monomodale pour reconnaître les activités émotionnelles dans les signaux physiologiques. Mon travail sur ces algorithmes a consisté dans un premier temps à les optimiser afin de réduire les temps de traitement, et ensuite les améliorer pour augmenter leurs robustesses ainsi que l'efficacité en vue de l'augmentation du taux de reconnaissance. Un accent particulier a été mis sur la validation des paramètres des classifieurs. Cette validation des paramètres est un problème assez complexe dont l'objectif est de définir un algorithme avec un fonctionnement optimal. Elle permet une comparaison plus rigoureuse de différents algorithmes et permet de généraliser l'apprentissage.

Dans ce qui suit, nous présentons notre méthode de détection des émotions avec l'approche monomodale (unimodale) et multimodale.

2.2 Approche monomodale

Comme nous avons vu dans les sections précédentes, on peut utiliser plusieurs signaux physiologiques pour la détection des émotions humaines. L'approche monomodale consiste à prendre les signaux physiologiques issus d'un seul capteur et ensuite, analyser ces derniers par deux phases essentielles de traitement pour la reconnaissance des émotions. La première phase correspond à l'apprentissage. Cette phase a pour objectif de concevoir une base d'apprentissage en appliquant différentes techniques de traitement de signal. La seconde consiste à utiliser la base d'apprentissage construite au cours de la première phase afin de détecter instantanément et automatiquement une activité émotionnelle.

La figure 2.1 illustre notre méthodologie de détection des émotions. En effet, dans notre méthode il y a 2 grandes phases à savoir : la phase *supervisée* et la phase *automatique* (non supervisée). Il est à noter que la phase d'apprentissage et la phase de détection correspondent respectivement à la phase supervisée et automatique. La phase supervisée est la phase en amont de la détection. Elle a pour but d'apprendre à la machine à reconnaître une activité émotionnelle ainsi qu'à discerner une émotion d'une

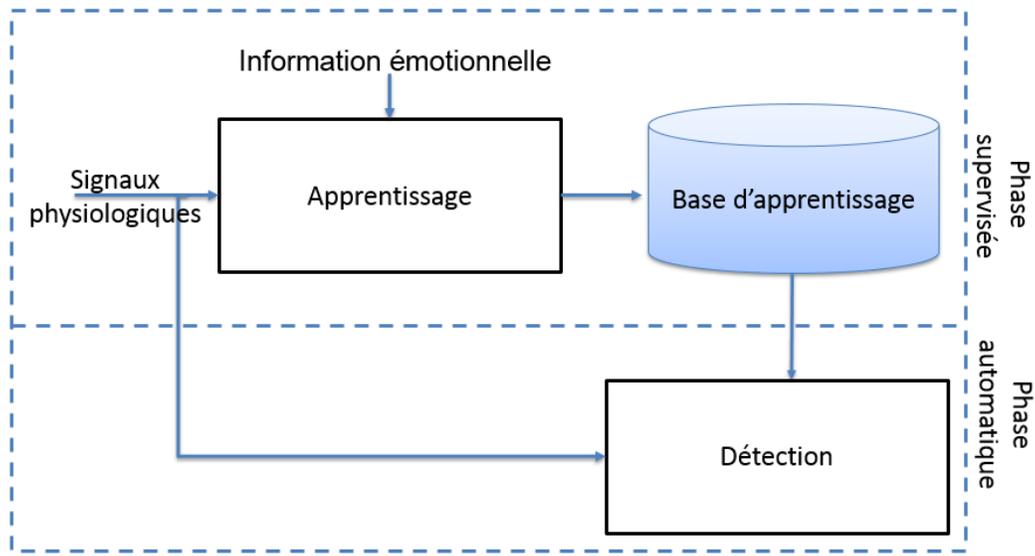


FIGURE 2.1 – Méthodologie générale de notre méthode de détection des émotions.

autre. Une fois que la machine a apprise une méthode pour reconnaître et discerner les activités émotionnelles, la phase automatique devient opérationnelle. Cette seconde phase prend comme entrées un signal physiologique et la base d'apprentissage construite lors de la première phase et consiste à extraire des caractéristiques types dans ce signal afin de détecter une activité émotionnelle. Dans les sections qui suivent, nous présentons les différentes étapes nécessaires aux différentes phases de traitement dans notre méthodologie.

2.2.1 Acquisition des signaux

Outre la difficulté de trouver un algorithme efficace pour la reconnaissance des émotions, l'acquisition des signaux pose en lui-même un problème délicat. Dans la section 1.3.1, nous avons brièvement abordé les contraintes que doit satisfaire le système physique d'acquisition des signaux physiologiques. En plus de ce système physique composé des capteurs de signaux physiologiques, d'un microcontrôleur, d'un module RF pour la transmission ainsi que d'une source d'alimentation (module batterie), le système d'acquisition dont on parle ici contient aussi une partie protocolaire. Par partie protocolaire, il ne faut pas entendre protocole de communication mais plutôt un protocole d'essai pour l'induction des émotions. En effet, un protocole simple d'utilisation et efficace permettant de stimuler la personne sur laquelle l'acquisition est faite est requis. Ce protocole stimule la personne sur laquelle les données sont acquises pour ressentir et exprimer les différentes émotions et donc, le protocole est d'une aide considérable pour la construction d'une base de données de signaux physiologiques. La méthode d'acquisition des signaux est essentielle. Elle influence fortement l'efficacité de la détection. En effet, la qualité des signaux (donc de la détection) n'est pas la

2.2 Approche monomodale

même quand les données acquises sont simulées ou réelles d'où la problématique de la définition d'un protocole adéquat et efficace. Ces protocoles font appels à des méthodes d'induction des émotions afin d'acquérir une base de données dans laquelle l'influence de chaque état émotionnel a été fidèlement reflétée.

Plusieurs protocoles d'induction des états émotionnels existent dans la littérature. A titre d'exemple, nous pouvons citer celui proposé par Clynes Manfred [Clynes, 1988] qui consiste à stimuler le sujet par différentes mélodies musicales pour produire 8 états émotionnels. Cette méthode appelée *Sentics cycle*, est défendue par son initiateur comme étant une méthode pour s'entraîner chez soit à ressentir chacun des états émotionnels considérés. Le système international de l'image affective (international affective picture system IAPS) développé par Lang et al. [Lang, 2005], est un autre protocole d'induction des états émotionnels. Le système IAPS fut adopté par de nombreuses études. Il fut notamment utilisé par Faiza Abat dans ses travaux de thèse [Abdat, 2010] vu la simplicité de sa manipulation. Enfin, un autre protocole basé sur les films est proposé dans la littérature pour l'induction des émotions [film, 2017]. Les réactions émotionnelles provoquées par ce dernier sont souvent trop faibles lorsqu'on parle d'émotions complexes [Abdat, 2010].

Dans ce qui suit, nous présentons la première base de données de signaux physiologiques que nous avons utilisée dans le cadre de nos travaux. Cette base de données conçue en utilisant le protocole *sentics cycle* est présentée dans la section suivante. Ensuite, plus loin, nous présentons notre propre base de données de signaux physiologiques construite pour 5 états émotionnels dans le cadre d'une application de ces travaux de thèse. Pour la conception de notre base de données, nous avons élaboré un protocole d'induction des émotions utilisant la Mnémothérapie musicale.

Base de données de signaux physiologiques fournis par le MIT

Dans nos travaux, nous avons utilisé la base de données de signaux physiologiques établie par Jennifer Healey au cours de sa thèse [Healey, 2000] soutenu au MIT (Massachusetts Institute of Technology). Cette base sert de référence pour la grande majorité des personnes qui travaillent sur la détection des émotions. Elle permet une comparaison de différents algorithmes pour la reconnaissance des émotions sur les mêmes données.

Pendant une durée de 32 jours, un acteur sur lequel étaient connectés 4 capteurs de signaux physiologiques (*EMG*, *RESPIration*, *GSR*, *BVP* échantillonnés tous à une fréquence de 20 Hertz) était stimulé pour exprimer et ressentir huit émotions : *No-emotion*, *Anger*, *Hate*, *Grief*, *Love*, *Romantic Love*, *Joy* and *Reverence*.

La méthode *Sentics cycle* développée par Manfred Clynes [Clynes, 2013] a été utilisée comme protocole de stimulation lors de la conception de la base de données du MIT. Ladite méthode consiste à stimuler l'acteur sur lequel les données sont acquises avec des mélodies ou avec des images pour qu'il ressente et exprime une activité émotionnelle donnée. Chaque jour, l'acteur assis sur une chaise dorsalée et connecté aux 4 capteurs de signaux physiologiques (*EMG*, *RESP*, *BVP*, *GSR*) est stimulé pour exprimer chaque

émotion pendant une durée de 3 à 4 minutes. Étant donné que la base est conçue pour 8 activités émotionnelles, une session d’acquisition durerait en moyenne 25 minutes. Le lecteur intéressé par le protocole d’acquisition utilisé par le MIT est invité à se reporter à ces différents travaux pour plus de détails [Healey, 2000], [Clynes, 1988], [Clynes, 2013]. Le tableau 2.1 regroupe les images utilisées dans le protocole du MIT pour inciter l’acteur à exprimer les huit émotions prises en compte ainsi que la description, l’intensité et la valence de ces différents états émotionnels.

Émotion	Images	Description	Intensité	Valence
Neutre	papier blanc, machine à écrire	ennui, offre d’emploi	basse	neutre
Colère	personnes en colère	désir de se battre	très haute	très négative
Haine	injustice, cruauté	colère passive	basse	négative
Chagrin	enfants souffrant, perte de mère	perte, tristesse	haute	négative
Amour platonique	famille, été	joie, paix	basse	positive
Amour romantique	rencontre romantique	excitation, désir	très haute	positive
Joie	de la musique	beaucoup de joie	moyennement haute	positive
Révérence	église, prière	calme, paix	très basse	neutre

TABLE 2.1 – Images utilisées dans le protocole du MIT pour provoquer les différents émotions chez l’acteur.

Ainsi comme montré dans le tableau 2.1, chaque image est choisie selon l’émotion que l’acteur doit exprimer. A titre d’exemple, des images d’enfants souffrant de famine provoque du chagrin chez l’acteur. Lorsque des images d’église ou de prière lui sont présentés, il ressentirait l’émotion révérence tandis que des images de personnes en colère provoque chez l’acteur la même émotion. Ce tableau est expliqué plus en détail dans [Healey, 2000].

Avec cette base de données du MIT, nous avons obtenu l’élément de base pour commencer notre analyse en vue d’une détection des émotions. Cette analyse comme nous l’avons dit dans les sections précédentes se base sur deux phases, à savoir la phase d’apprentissage et celle de détection. Les sections qui suivent présentent l’ensemble des étapes requises pour ces différentes phases.

2.2.2 Apprentissage

À l’école primaire, par exemple, pour apprendre à un enfant à distinguer les différentes lettres de l’alphabet ; on lui montre les lettres les unes après les autres tout

2.2 Approche monomodale

en lui fournissant une méthode pour différencier ces lettres. Une fois que l'enfant a appris ce processus de reconnaissance des lettres, il le stocke dans sa mémoire et peut alors distinguer les lettres de lui-même sans information complémentaire. Ainsi, de la même manière qu'on apprend à cet enfant à reconnaître les lettres de l'alphabet, la phase d'apprentissage dans notre méthodologie consiste à apprendre à la machine un mécanisme de reconnaissance des différents états émotionnels. Le but de cette étape est de fournir une base d'apprentissage qui sera utilisée dans la phase de détection. La base d'apprentissage est assimilable à ce que l'enfant stocke dans sa mémoire. Dans notre méthodologie, la phase d'apprentissage nécessite 4 étapes (voir figure 2.2) essentielles à savoir : Découpage, Filtrage, extraction des caractéristiques, création de la base d'apprentissage. Dans ce qui suit, nous expliquons chacune de ces étapes.

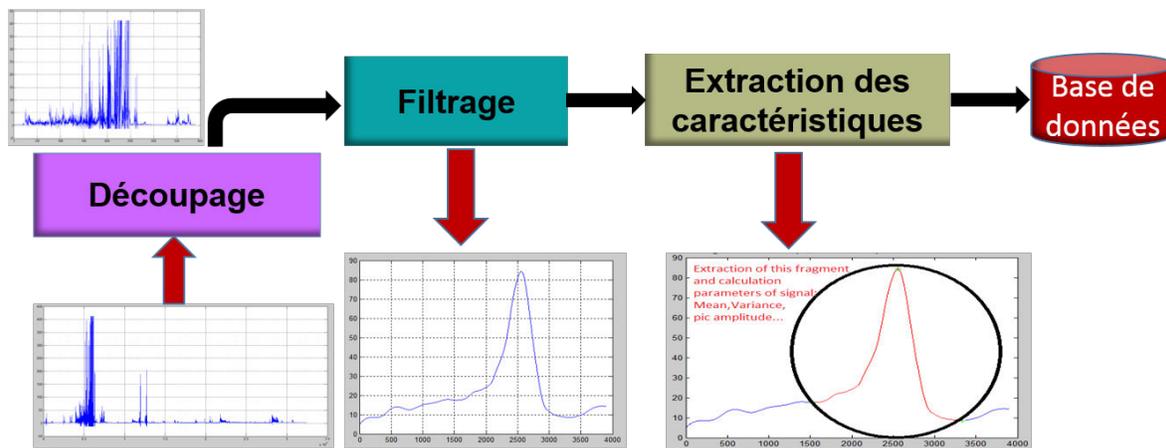


FIGURE 2.2 – Synoptique de la phase d'apprentissage.

2.2.2.1 Découpage

Après avoir acquis les signaux physiologiques (dans un premier temps, nous utilisons les signaux fournis par le MIT), il est nécessaire d'isoler la partie du signal correspondant à une émotion donnée puisque nous avons une information sur la période à laquelle chacune des huit émotions est exprimée. Cette étape consiste donc à diviser chaque signal journalier en 8 portions de signaux correspondant aux 8 émotions comme le montre la figure 2.3. Pour l'apprentissage, les portions sont étudiées les unes après les autres tout comme l'enfant apprend les lettres les unes après les autres pour établir le mécanisme de reconnaissance dont nous avons parlé. La base d'apprentissage est constituée de l'ensemble des caractéristiques extraites dans les signaux. Ces dernières permettent de caractériser les activités émotionnelles et de distinguer les différents états émotionnelles.

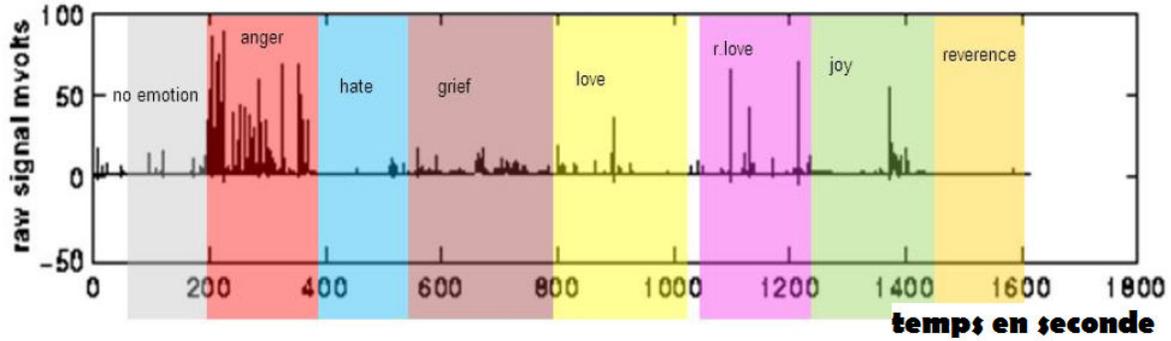


FIGURE 2.3 – Représentation du signal brut EMG sur une journée d'acquisition avec les huit émotions.

2.2.2.2 Filtrage

Prenons deux personnes (Laura et Émilie) discutant à côté d'une foule de manifestants. De leur point de vue, l'information qu'elles se donnent est plus importante que les sons émis par les autres manifestants. Leurs oreilles se chargent donc d'atténuer les sons émis par les manifestants afin que Laura et Émilie puissent s'entendre, se comprendre pour communiquer du mieux possible. Toute acquisition est accompagnée de bruit. En effet, dans toutes les acquisitions, des signaux parasites viennent se superposer aux signaux utiles. De même que les oreilles de Laura et Émilie sont importantes dans leur processus de communication à côté du groupe de manifestants, le filtrage l'est tout autant pour un système d'acquisition de signaux. Le filtrage consiste à enlever le bruit du signal utile et facilite l'extraction des caractéristiques.

Deux méthodes de filtrage sont principalement utilisées. La première méthode nécessite l'application de la transformée de Fourier (voir équation 2.1) sur le signal temporel et ensuite l'application d'un filtre passe bas qui permet d'éliminer les fréquences du bruit. La deuxième méthode qui est moins coûteuse en temps de calcul consiste à simplement effectuer une convolution du signal (voir équation 2.2) par un filtre.

$$F(k) = \sum_{j=1}^N s(j) \exp^{-2i\pi k \frac{j}{N}} \quad 1 \leq k \leq N \quad (2.1)$$

Où $F(k)$ est la transformée de Fourier du signal $s(k)$.

$$S(k) = \sum_{j=-\frac{l}{2}}^{\frac{l}{2}} s(k)h(k-j) \quad (2.2)$$

Avec $s(k)$ le signal physiologique acquis, $h(t)$ le filtre utilisé pour le dé-bruitage, $S(k)$ le signal filtré et l la taille de la fenêtre de filtrage.

Dr. Tayari, dans le cadre de sa thèse, a opté pour la méthode de convolution. Elle a étudié les filtres moyenneur, médian, gaussien, hanning et a finalement choisi le filtre

2.2 Approche monomodale

hanning dû au fait que ce dernier permet de détecter les pics tout en améliorant la forme spectrale. De plus, il permet une interprétation émotionnelle plus pertinente grâce à sa capacité à regrouper plusieurs variations du signal en un seul pic. Les coefficients de ce filtre sont obtenus en utilisant l'équation 2.3 où N représente la taille de la fenêtre, w les différents coefficients de ce filtre.

$$w(k) = 0.5 * (1 - \cos(\frac{2\pi k}{N})) \quad 0 \leq k \leq N \quad (2.3)$$

Nous verrons plus tard, l'importance du choix de la taille de la fenêtre car ce dernier influence considérablement la reconnaissance des émotions. L'étude des différentes fenêtres de filtrage a été faite par Dr. Tayari. Dans nos travaux, nous avons utilisé la fenêtre de filtrage hanning qu'elle a sélectionné à l'issue de ses études.

2.2.2.3 Extraction des vecteurs caractéristiques

Cette étape est cruciale dans notre méthodologie de détection des émotions. Après avoir isolé et filtré une portion du signal, nous procédons à la détection des pics. Considérons une personne qui marche tranquillement, elle présente un signal du rythme cardiaque fluctuant autour d'une certaine valeur, quand tout à coup, un lion surgit de nulle part devant elle. Le marcheur prend peur et instantanément son rythme cardiaque s'accélère (le signal croît) puis décélère (le signal décroît) au cours du temps. Cette croissance et décroissance ou vice-versa du signal s'assimile à un pic avec un Δt donc à la présence d'une activité émotionnelle.

Pour détecter les extremums (maximums et minimums) dans un signal continu, on calcule sa dérivée puis on calcule les antécédents du point zéro par rapport au signal. De la même façon, l'approche que nous utilisons consiste à calculer le gradient du signal et à détecter les passages par zéro du gradient. Dans les signaux discrets, il est très difficile de trouver des passages par zéro du gradient à cause notamment de l'échantillonnage. Ainsi, pour détecter les extremums, nous recherchons les changements de signe dans le gradient. Un point maximum est détecté lorsqu'on constate le passage d'un gradient positif à un gradient négatif et réciproquement, le passage d'un gradient négatif à un gradient positif donne un minimum. Pour détecter un pic, il faut nécessairement détecter dans cet ordre *un minimum suivi d'un maximum suivi d'un autre minimum*, la portion du signal se trouvant entre les deux minimums est alors isolée pour les traitements ultérieurs.

Après avoir extrait et isolé un pic, afin de trouver des caractéristiques types pour chaque émotion, nous calculons sept paramètres essentiels définis par les équations suivantes constituant ainsi un vecteur caractéristique pour chaque pic détecté. Les sept paramètres permettent de caractériser convenablement un pic donné. En supposant que chaque signal mesuré est généré par un processus Gaussien, avec des échantillons indépendants et identiquement repartis. Les deux fonctions physiologiques qui peuvent être utilisées pour caractériser un signal physiologique brut sont la moyenne et la variance. La moyenne du signal filtré et celle du gradient du signal filtré ainsi que la

variance du signal filtré sont calculées afin d'évaluer la tendance d'un signal. Enfin, la largeur et l'intensité d'un signal peuvent également fournir des informations pertinentes.

La moyenne du signal μ_S et du signal filtré $\widetilde{\mu}_S$:

$$\mu_S = \frac{1}{N} \sum_{n=1}^N S(t) \quad , \quad \widetilde{\mu}_S = \frac{1}{N} \sum_{n=1}^N \widetilde{S}(t) \quad (2.4)$$

La variance du signal σ_S^2 et du signal filtré $\widetilde{\sigma}_S^2$:

$$\sigma_S^2 = \frac{1}{T} \sum_{n=1}^T (S(t) - \mu_S)^2 \quad , \quad \widetilde{\sigma}_S^2 = \frac{1}{T} \sum_{n=1}^T (\widetilde{S}(t) - \widetilde{\mu}_S)^2 \quad (2.5)$$

La moyenne du gradient du signal filtré $\delta_{\widetilde{S}}$:

$$\delta_{\widetilde{S}} = \frac{1}{T-1} \sum_{t=1}^{T-1} |\widetilde{S}(t+1) - \widetilde{S}(t)| = \frac{1}{T-1} (s_T - s_1) \quad (2.6)$$

La largeur a du pic :

$$a = |t_{min2} - t_{min1}| \quad (2.7)$$

L'intensité b (l'amplitude) du pic :

$$b = |S(t_{max}) - \min(S(t_{min1}), S(t_{min2}))| \quad (2.8)$$

Les équations précédentes ont des caractéristiques très générales et peuvent être appliquées à un large type de signaux physiologiques (EMG, ECG, EDA, etc.). En utilisant ces caractéristiques, nous obtenons un vecteur caractéristique Y de 7 valeurs pour chaque échantillon. $Y = [\mu_S \quad \widetilde{\mu}_S \quad \sigma_S^2 \quad \widetilde{\sigma}_S^2 \quad \delta_{\widetilde{S}} \quad a \quad b]$

2.2.2.4 Construction d'une base d'apprentissage

Toutes les méthodes d'apprentissage mises en place dans la littérature pour la détection des émotions en utilisant les signaux physiologiques ont pour but d'établir des relations entre chacun des différents signaux (EMG, GSR, ECG, ...) et les états émotionnels considérés dans l'étude. En effet, ces algorithmes ont pour souci d'améliorer les performances (temps et efficacité) de la classification ce qui permet donc d'affiner le décodage des différents états émotionnels dans les signaux. En augmentant la complexité du ou des modèle(s), il est assez aisé d'avoir de très bonnes performances sur les données d'apprentissage (données connues). Les modèles encore appelés fonctions de prédiction seront désignés *f(.) dans la suite de ce manuscrit*. La difficulté réside au niveau de la généralisation qui donne les performances de *f(.)* sur des données inconnues. De ce fait, une validation des paramètres de *f(.)* est nécessaire afin d'optimiser au mieux son comportement sur les données inconnues.

2.2 Approche monomodale

2.2.2.5 Validation des paramètres

Dans la littérature, il existe deux principales méthodes pour la validation des paramètres de $f(\cdot)$. Ces méthodes généralisent l'apprentissage et la mesure des performances des algorithmes de classification :

- **La méthode optimiste** : Consiste à diviser la banque de signaux disponibles en base d'apprentissage ε et base de test τ . Dans cette méthode on apprend la fonction de prédiction $f(\cdot)$ sur ε , valide ses paramètres sur les données τ et teste/calcule les performances sur les mêmes données τ . Cette méthode n'est pas très rigoureuse car les paramètres sont validés et testés sur les mêmes données, on peut donc dire que c'est une méthode de classification biaisée [Jrad et al., 2011].
- **La méthode pessimiste** : Consiste à diviser la banque de signaux disponibles en trois parties : la base d'apprentissage ε , la base de validation ν et la base de test τ . Dans cette méthode, on apprend $f(\cdot)$ sur ε , valide ses paramètres sur les données ν et teste/estime les performances sur les données τ . Cette méthode élimine donc le biais sur la première méthode, car les données sur lesquelles les performances de la fonction de prédiction ont été estimées étaient inconnues pour la fonction $f(\cdot)$ au moment de sa mise en œuvre et la validation de ses paramètres. Cette méthode est donc une généralisation de la classification, car on peut s'attendre au même ordre de performances lorsque nous changeons les données de test [Flamary et al., 2015].

En raison des coûts d'acquisition des signaux physiologiques, les données générées ne sont pas très abondantes. Ceci amène les personnes travaillant sur ces signaux à faire un compromis entre réduire la base d'apprentissage au risque de diminuer la qualité de $f(\cdot)$ et réduire la base de test au risque de diminuer la fiabilité sur le taux de classification obtenu. Dans un souci de rigueur, nous avons implémenté l'approche de validation des paramètres par la méthode pessimiste. Cette méthode nous permet une généralisation de la fonction de prédiction $f(\cdot)$. C'est à dire qu'une fois que nous aurons défini $f(\cdot)$ avec les paramètres optimaux, notre classifieur pourra aussi bien fonctionner sur les données actuelles ou sur n'importe quelles autres données acquises des capteurs pour lesquels nous avons une relation entre signaux et états émotionnels. Le lecteur intéressé par les méthodes de validation des paramètres est invité à se reporter à ces différents travaux [Friedman et al., 2001], [Duda et al., 2012], [Jrad et al., 2011].

Dans la littérature, trois principales techniques sont utilisées pour découper les données en base d'apprentissage et de test comme l'illustre la figure 2.4. L'ensemble des données d'apprentissage, de validation et de test est représenté sur la figure 2.4a.

Hold out : Le Hold out (voir figure 2.4b) consiste à scinder aléatoirement la base de données en deux parties à savoir la base d'apprentissage ε et la base de test τ . Ensuite, on apprend une fonction de prédiction $f(\cdot)$ sur ε et les performances sont mesurées sur τ .

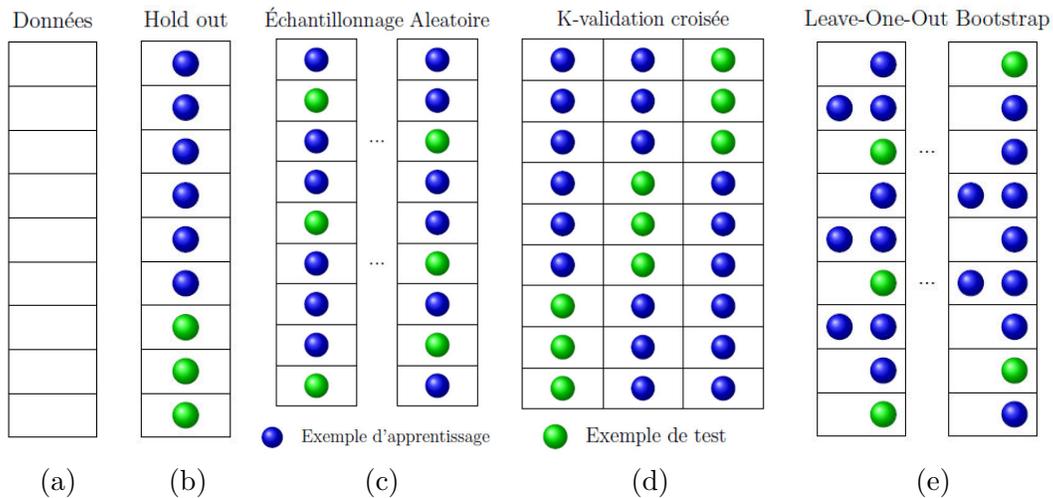


FIGURE 2.4 – Les techniques de séparation des données en base d'apprentissage et de test. Figure extraite de [Flamary et al., 2015]

Échantillonnage aléatoire : La technique *d'échantillonnage aléatoire* (voir figure 2.4c) consiste à répéter K fois celle de hold out afin de générer K bases d'apprentissage (ε) et K bases de teste (τ) indépendantes. La performance est obtenue en faisant la moyenne des différentes performances obtenues avec les bases ε et τ . Elle présente l'avantage de diminuer la variabilité des performances par rapport à la technique Hold out mais augmente considérablement le temps de calcul.

K-validation croisée et leave-one-out (LOO) Dans la validation croisée, les n exemples de la base de données sont scindés en K -sous ensembles de tailles égales (voir figure 2.4d). Sur l'un des K sous ensembles, on valide les paramètres et les sous ensembles restants constitueront la base d'apprentissage. On répète cette opération jusqu'à ce que chacun des K sous ensembles ait été pris comme ensemble sur lequel les paramètres sont validés et à chaque fois les performances sont estimées. La performance globale est calculée de la même façon que pour l'échantillonnage aléatoire. La validation croisée réduit la variance de l'estimation des performances du fait que chaque exemple ne figure qu'une et une seule fois dans l'ensemble de test [Flamary et al., 2015].

La technique Leave-one-out (LOO) n'est autre qu'un cas particulier de la technique k-validation croisée $k = n$. De ce fait, nous prenons $(n - 1)$ exemples comme ensemble d'apprentissage et un exemple comme ensemble de validation. Cette opération est donc répétée n fois afin que chaque exemple puisse être pris comme ensemble de validation des paramètres.

Bootstrap La technique bootstrap (voir figure 2.4e) consiste à générer à partir des n exemples dans la base de données la base d'apprentissage en effectuant n tirages uniformes avec remise. Du fait qu'elle utilise le tirage avec remise, certains exemples sont absents dans la base d'apprentissage tandis que d'autres s'y trouvent

2.2 Approche monomodale

à plusieurs reprises. Les performances sont calculées comme dans les deux autres cas. L'estimation des performances est moins optimiste par rapport au cas du leave-one-out; le biais est similaire à celui de la k-validation croisée avec $K = 2$ [Friedman et al., 2001],[Flamary et al., 2015].

Afin d'être rigoureux et de proposer des fonctions de prédiction générales (c'est à dire des fonctions sur lesquelles on peut avoir de bons taux de classification avec les données acquises après définition de ces fonctions), nous avons adopté l'approche pessimiste avec la k-validation croisée pour la création de la base d'apprentissage. La figure 2.5 met en évidence notre technique de découpage de la base de données disponible en base d'apprentissage, base de validation et base de test.

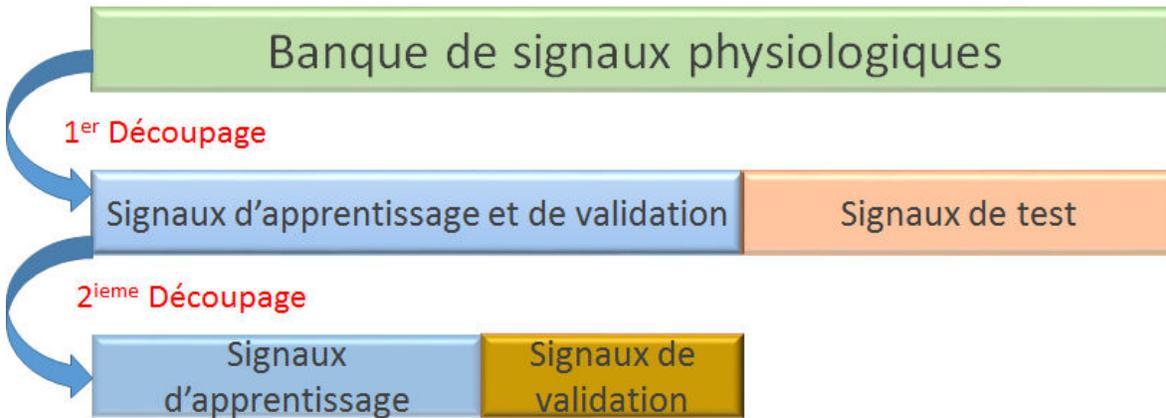


FIGURE 2.5 – Découpage de la banque de signaux disponible en base d'apprentissage, validation et test.

2.2.3 Détection

Dans cette phase, nous effectuons les même étapes de filtrage, d'extraction des vecteurs caractéristiques que dans la phase d'apprentissage sans passer par celle de découpage car ici on suppose ne pas avoir de connaissances a priori de l'émotion ressentie.

À fin de la phase d'apprentissage, nous aurons autant de bases d'apprentissage que de capteurs utilisés lors de la conception de cette dernière. Après avoir filtré le signal et détecté un pic, nous extrayons cette partie du signal et calculons un vecteur caractéristique. Ensuite, l'étape de classification prenant en entrée le vecteur caractéristique extrait dans le signal et la base d'apprentissage est appliquée afin de calculer un vecteur émotionnel. La figure 2.6 montre le synoptique des différentes étapes nécessaires à la phase de détection.

Dr. Tayari, dans le cadre de ses travaux de thèse, a proposé un modèle multidimensionnel de modélisation des états émotionnels. La dimension de la base d'émotion qu'elle a considérée est égale à huit et cette dimension est formellement définie par

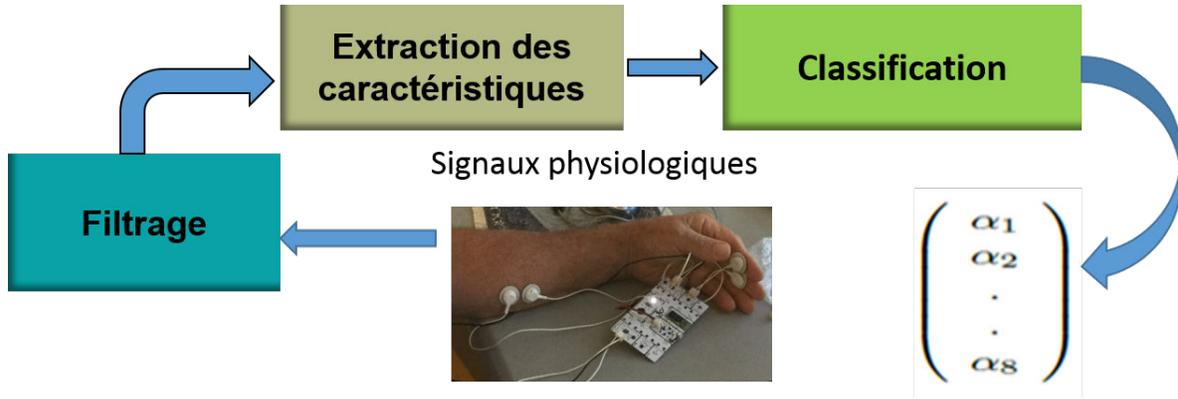


FIGURE 2.6 – Synoptique de la phase de détection

$(B) = (Joie, Tristesse, Acceptation, Dégoût, Peur, Colère, Surprise, Anticipation)$. Ainsi, toute émotion de façon formelle peut être représentée à partir d'une combinaison des autres émotions fondamentales qui définissent la Base (B) .

Constituant une base (B) avec les émotions de base de Plutchik (les émotions de la première dimension du modèle de Plutchik), elle propose de représenter chaque état émotionnel comme combinaison linéaire des émotions de la base (B) où chaque émotion correspond à un $u_i, i \in [1, \dots, 8]$.

$(u_i) = (Joie, Tristesse, Acceptation, Dégoût, Peur, Colère, Surprise, Anticipation)$
Chaque activité émotionnelle e est représentée par le vecteur d'émotions E d'écrit donc ainsi :

$$(e) = \sum_{i=1}^8 \langle E, u_i \rangle u_i \quad (2.9)$$

$$\text{où } E = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_8 \end{pmatrix}_B$$

L'équation 2.9, sous sa forme vectorielle devient :

$$(e) = \alpha_1 Joie + \alpha_2 Tristesse + \dots + \alpha_7 Surprise + \alpha_8 Anticipation \quad (2.10)$$

Avec $\alpha_i \in [0, 1]$. Plus α_i est proche de 1, plus l'émotion est ressentie et réciproquement une valeur de α_i proche de 0 montre que l'émotion n'est pas ressentie. Lorsque $\alpha_i = 1$, l'émotion ressentie est donc une émotion fondamentale. Les émotions composées sont un mélange de plusieurs émotions fondamentales. En se basant sur le

2.2 Approche monomodale

modèle de combinaison des émotions de Plutchik représenté sur la figure 2.7 et sur les intensités de chaque émotion, elle a proposé un dictionnaire incluant environ soixante dix émotions qui peut être étendu [Tayari et al., 2013]. Il est à noter que dans nos travaux, la base de données est sur les 8 émotions suivantes : neutre, colère, haine, chagrin, amour, amour romantique, joie et révérence. Chacune de ces 8 huit émotions peut être écrite comme combinaison linéaire des 8 huit émotions de base de plutchik constituant notre base (B). Par exemple, l'émotion amour peut être écrire combinaison

$$E_{joie} = \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_B \quad E_{acceptation} = \begin{pmatrix} 0 \\ 0 \\ \alpha_3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_B$$

$\alpha_1 \neq 0, \alpha_3 \neq 0. E_{amour} = E_{joie} \oplus E_{acceptation} = (\alpha_1 \ 0 \ \alpha_3 \ 0 \ 0 \ 0 \ 0 \ 0)_B$

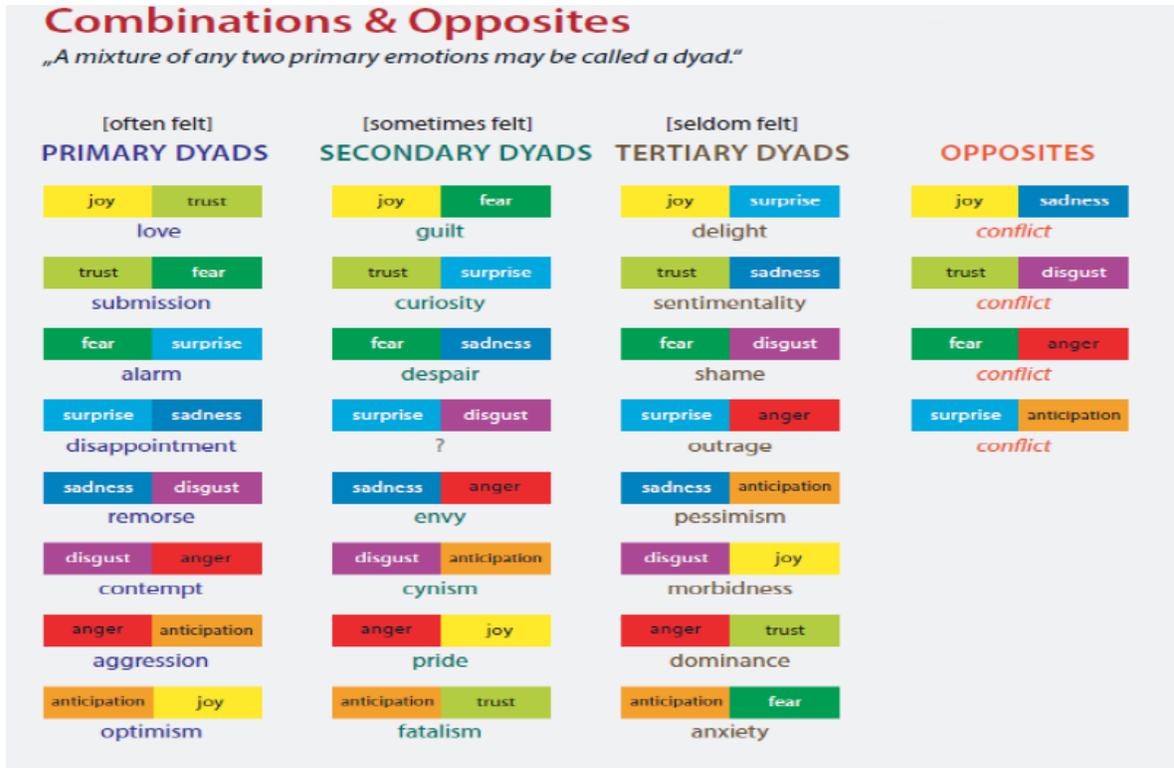


FIGURE 2.7 – Modèle de combinaison des émotions proposé par Plutchik [Plutchik, 1980].

La phase de détection a essentiellement pour but la détermination des différents coefficients α_i qui donnent les différentes probabilités qu'une émotion soit ressentie ou pas. Ainsi, l'approche de représentation des émotions sous une forme vectorielle permet

de remonter facilement aux émotions composées qui ne sont que des combinaisons des émotions de base. Dans toute notre méthodologie, nous garderons cette représentation des états émotionnels.

L'étape de classification a pour objectif de prédire une classe émotionnelle du vecteur caractéristique extrait en utilisant la base d'apprentissage. Ainsi dans ce qui suit, nous présentons quelques algorithmes de classification que nous avons étudié et implanté dans notre système en vue de déterminer les coefficients α_i .

2.2.4 Choix d'un classifieur

Dans le but de hiérarchiser et de catégoriser les objets, les animaux, les maladies, les documents d'une bibliothèque, nous avons besoin d'un système de classement organisé et efficace. Ce système de classification encore appelé *classifieur* apporte un éclairage sur la nature de la connaissance. Ainsi, Cournot dans ses travaux sur les fondements de nos connaissances définit la classification comme étant une opération de l'esprit. Cette opération pour la commodité des recherches ou de la nomenclature, pour le secours de la mémoire, pour les besoins de l'enseignement, ou dans tout autre but relatif à l'homme, groupe artificiellement des objets auxquels il trouve quelques caractères communs. Elle donne au groupe artificiel ainsi formé une étiquette ou un nom générique [Broad, 1913]. Classifier une connaissance revient donc à situer les connaissances les unes par rapport aux autres. Dans notre étude, les connaissances sont assimilables aux vecteurs caractéristiques que nous extrayons dans chaque activité émotionnelle pour mieux la caractériser et les étiquettes correspondent aux différentes émotions considérées.

Ainsi, dans les paragraphes qui suivent, nous introduisons deux algorithmes de classification que nous avons étudié et implanté dans notre système.

2.2.4.1 KNN [Larose, 2005]

La méthode KNN (k nearest neighbours en anglais signifiant k plus proches voisins) est une méthode d'apprentissage supervisé intuitive et simple à mettre en œuvre. D'une certaine façon, on peut dire qu'elle tient sa base de ce proverbe :

Dis-moi qui tu fréquentes, je te dirai qui tu es.

Ce proverbe signifie que L'homme s'identifie aux personnes qu'il côtoie. Ces personnes forment donc son réseau de fréquentation et ont un impact plus ou moins considérable sur son comportement. Par analogie à ceci, l'algorithme KNN consiste à calculer les distances entre le vecteur caractéristique d'une nouvelle observation x à classer et les vecteurs caractéristiques des signaux de référence qui forment la base d'apprentissage. Le signal à classer est alors assigné à la classe majoritaire parmi les k classes les plus proches en fonction du calcul de distance choisi. Il existe différents types de calcul de distance (distance Euclidienne, distance de Manhattan, distance de Minkowski et la distance de Tchebychev). Nous avons opté pour la méthode du calcul de la distance Euclidienne à cause notamment de sa simplicité de mise en œuvre et de son

2.2 Approche monomodale

efficacité. Cette méthode encore appelée *2-distance* permet de généraliser l'application du théorème de Pythagore à un espace de dimension n , ce qui permet d'accélérer le calcul de la distance. De plus, c'est la distance la plus intuitive.

Soient $X_p = (x_{p1}, x_{p2}, \dots, x_{pN})$ un vecteur caractéristique à classer, N le nombre de caractéristique dans ce vecteur et Y la base de référence.

$$d(X_p, Y) = \sqrt{\sum_{i=1}^N (X_{pi} - Y_i)^2} \quad (2.11)$$

Dans la méthode KNN, les coefficients α_i sont déterminés en utilisant la formule suivante :

$$\alpha_i = \frac{k_i}{k} \quad (2.12)$$

où k_i représente le nombre de i plus proches voisins trouvés. α_i représente la probabilité que l'émotion soit de rang i avec $\sum_{i=1}^M \alpha_i = 1$ où M représente le nombre de classes émotionnelles.

La difficulté dans l'algorithme des KNN est de trouver le k optimal. Plus k est élevé, plus grand est le risque de sur-apprentissage et les frontières entre les classes émotionnelles sont moins distinctes. L'utilisation des techniques de validation de paramètre dont nous avons parlé précédemment est donc nécessaire afin de trouver une valeur de k qui diminue le risque de sur-apprentissage et maximise le taux de reconnaissance des émotions.

L'algorithme de classification des k plus proches voisins est assez intuitif et simple à mettre en œuvre. Après l'implantation de ce dernier, en vue de l'amélioration des performances de notre système, nous avons développé et implanté un nouvel algorithme appelé séparateur à vaste marge (SVM). Dans la littérature sur les systèmes de classification, le SVM présente de meilleures performances par rapport au KNN [Zhang et al., 2006]. Nous allons donc comparer ces deux algorithmes en termes de performances sur nos données. Le paragraphe qui suit, introduit cet algorithme de classification.

2.2.4.2 SVM [Wang, 2005]

Le SVM (support vector machine en anglais signifiant séparateur à vaste marge) est une méthode d'apprentissage supervisé pour la classification binaire. Le principe consiste à définir un hyperplan séparant les 2 classes tout en maximisant la marge. Cette dernière est la distance entre la frontière de séparation et les vecteurs supports qui sont les plus proches échantillons. Dans le cas où les données sont linéairement séparables, on définit une fonction séparatrice optimale (maximisant la marge). Dans le cas contraire, les données sont projetées dans un espace de plus grande dimension où elles peuvent être linéairement séparables. Soient x_i, x_j les données de deux classes. Le SVM s'appuie sur le produit scalaire de ces vecteurs d'entrée $\langle x_i, x_j \rangle$. Dans le cas où les données ont été projetées dans un espace de plus grande dimension à l'aide de la transformation

$\phi : x \rightarrow \phi(x)$, le produit scalaire devient alors : $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ avec $K(x_i, x_j)$ la fonction noyau.

Différentes fonctions noyau existent comme :

— Noyau linéaire :

$$K(x_i, x_j) = x_i \cdot x_j$$

— Noyau polynomial de degré p :

$$K(x_i, x_j) = (1 + x_i \cdot x_j)^p$$

— Noyau Gaussien :

$$K(x_i, x_j) = \exp \frac{-\|x_i - x_j\|^2}{2 \cdot \sigma^2}$$

— Noyau sigmoïde où α et β sont des paramètres d'optimisation :

$$K(x_i, x_j) = \tanh(\alpha x_i \cdot x_j + \beta)$$

La fonction de décision est définie par :

$$\tilde{y} = \text{sign} \left[\sum_i \alpha_i y_i K(x, x_i) + b \right]$$

Où α et b sont des paramètres d'optimisation. Ils sont des termes de régularisation qui pondèrent l'attache aux données et la régularisation. L'attache aux données consiste en une mesure d'erreur sur une liste d'échantillons d'apprentissage contenant les exemples et la classe correspondante.

Dans notre étude, nous avons choisi d'implanter le noyau gaussien car les autres ne permettaient pas d'avoir de résultats probants. En effet, la gaussienne est une fonction capable de séparer des données ayant une grande similarité. L'application de l'algorithme classique du SVM ne permet pas d'aboutir directement à des probabilités à posteriori de classification donc de remonter aux coefficients α_i . Pour garder une approche statistique, la sortie du SVM doit être représentée sous forme de probabilité afin d'aboutir à une représentation vectorielle des émotions. Nous avons utilisé la toolbox LIBSVM [Chang and Lin, 2011] qui propose une méthode de calcul des probabilités à partir des paramètres obtenus par SVM. Les probabilités à posteriori sont ainsi déterminées à l'aide d'une fonction paramétrique.

Nos travaux s'inscrivent dans la problématique multi-classes car nous essayons de détecter plusieurs activités émotionnelles. Deux approches existent pour passer du SVM classique (classification binaire : 2 classes) au SVM multi-classes.

1. **Approche un contre tous** : Le principe de cette approche repose sur la transformation du problème à k classes en k classifieurs binaires. On apprend un modèle permettant de discriminer les groupes $y = k$ et $y \neq k$. Chaque modèle doit être capable de bien séparer une classe spécifique des autres. Les estimations des probabilités a posteriori trouvées pour chaque modèle donnent les coefficients α_i à partir desquels on forme le vecteur émotionnel.

2.2 Approche monomodale

- 2. Approche un contre un :** Dans cette approche, le principe consiste à considérer toutes les paires de classes possibles (C_k^2 possibilités de combinaisons). Prenons un exemple de 8 classes on va donc avoir $C_8^2 = 28$ donc 28 combinaisons possibles. L'approche consiste à constituer autant de modèles (fonctions prédictives) que de paires de classes. Les coefficients α_i sont calculés à partir du nombre de duels gagnés par chacune des classes émotionnelles.

La différence des performances de ces deux approches n'est pas si évidente que ça. Cependant, quelques études qui se contredisent ont été faites sur ces approches comme [Milgram et al., 2006] et [Allwein et al., 2000]. De ce fait, l'utilisation d'une approche ou de l'autre dépend de l'application visée et sa performance dépend de la finesse des modèles préétablis dans l'étape de classification binaire. C'est à dire chaque modèle devra être le plus efficace possible dans la classification d'une classe donnée afin d'augmenter la précision de l'algorithme multi-classes. De ce fait, nous avons choisi d'implanter l'approche un contre tous qui est l'approche la plus évidente et la plus intuitive. Cette approche nécessite moins de modèles (fonctions prédictives) à définir, ainsi, l'étape de validation nécessite moins de temps. La détection elle-même nécessite moins de temps de calcul que dans l'approche un contre un à cause du nombre de modèles à tester.

2.2.4.3 Résultats et comparaisons des performances des deux classifieurs

Après avoir regroupé tous ces signaux physiologiques, nous avons utilisé les techniques de validation de paramètres et de découpage des données en bases d'apprentissage, de validation et de test élucidées et choisies dans la section précédente afin de valider les hyper-paramètres de nos classifieurs. Ainsi, 84 vecteurs caractéristiques ont été extraits des signaux physiologiques pour chaque émotion. Étant donné que nous avons 8 émotions, 672 vecteurs caractéristiques ont donc été extraits pour chaque signal physiologique. Sur ces vecteurs caractéristiques, $34 * 8 = 272$ (environ 40%) ont été utilisés comme base de test pour tester nos algorithmes et les 60% restants ($50 * 8 = 400$ vecteurs caractéristiques) ont été utilisés comme base d'apprentissage et de validation. 60% des 400 vecteurs caractéristiques, soit $30 * 8 = 240$ vecteurs caractéristiques constituent donc la base d'apprentissage et enfin $20 * 8 = 160$ pour la base de validation. En somme, les bases d'apprentissage, de validation et de test sont respectivement composées chacune de 240, 160 et 272 vecteurs caractéristiques.

Après avoir subdivisé la base de signaux fournis en trois bases : *base d'apprentissage*, *base de validation* et *base de test*; nous avons appliqué l'algorithme des k plus proches voisins et celle des séparateurs à vaste marge sur ces données. Les résultats obtenus par notre algorithme du KNN avec l'approche uni-modale sont montrés sur la figure 2.8. Cette approche uni-modale permet d'avoir un taux de reconnaissance moyen de 57.24%.

Comme le montre la figure 2.8, certaines émotions sont mieux détectées avec certains capteurs de signaux physiologiques qu'avec d'autres. Nous pouvons bien remarquer que les émotions *neutre* et *Amour platonique* sont mieux détectées avec la modalité

BVP (Blood Volume Pressure). Les signaux de cette modalité se corrént bien avec le rythme cardiaque. Ainsi, un faible niveau de ce signal traduit la présence d’une émotion à valence neutre et un fort niveau traduit une émotion de grande intensité (joie intense, peur, colère). Grâce au fait que l’émotion amour est une combinaison de la joie et de l’acceptation, cette modalité peut alors bien détecter l’émotion *amour*. La modalité GSR (Galvanic Skin Response) permet de mieux détecter les émotions *haine* et *joie*. L’Activité ElectroDermale d’une personne croit linéairement avec son niveau d’excitation globale. Cette caractéristique fait de cette modalité un bel indicateur des émotions comme la haine, le stress l’attention et la joie. La modalité EMG (ElectroMyoGram activity) quant à elle permet plutôt de mieux reconnaître les émotions *colère* et *amour romantique*. Le signal EMG est un bel indicateur des émotions à valence négative.

L’analyse des différents résultats présentés permet de conclure qu’il existe une propriété de spécialisation des capteurs. En effet, chaque modalité est spécialisée dans la reconnaissance de certaines émotions spécifiques. Cette propriété est très importante car elle permettra de pondérer chacune des modalités, selon qu’elle peut mieux détecter une émotion ou non dans le but d’une détection plus efficace. Ainsi, nous utiliserons cette caractéristique dans notre approche multimodale afin d’améliorer considérablement le taux de reconnaissance.

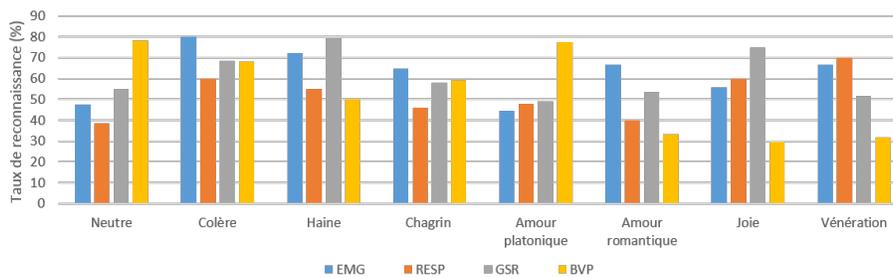


FIGURE 2.8 – Taux de reconnaissance mono-modal des émotions en utilisant l’algorithme du KNN.

Pour proposer un système de détection des émotions avec les meilleures performances, nous avons implanté un deuxième algorithme de classification à savoir le SVM. Cet algorithme selon l’état de l’art sur les algorithmes de classification, semble être plus rapide et plus efficace que le KNN. L’implantation des 2 algorithmes nous permet de comparer ces deux classifieurs sur les mêmes données et d’améliorer les performances de notre système. Les résultats obtenus par ce deuxième classifieur montrés sur la figure 2.9 donne un taux de reconnaissance moyen de **57.87%**.

Au regard de ces résultats, deux remarques peuvent être faites. La première montre bien que la caractéristique de spécialisation de chaque capteur de signaux physiologiques dans la reconnaissance de certaines émotions spécifiques se confirme. Le capteur EMG n’est plus le spécialiste pour l’émotion *amour romantique*. Deuxième remarque, nous voyons que le taux de reconnaissance obtenu par l’utilisation du classifieur SVM est très

2.2 Approche monomodale

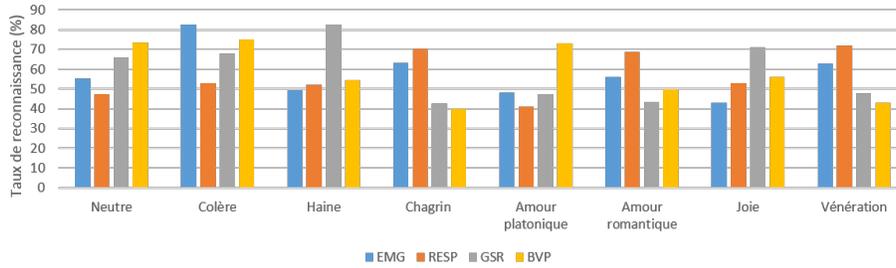


FIGURE 2.9 – Taux de reconnaissance mono-modal des émotions en utilisant l’algorithme du SVM.

légèrement supérieur à celui obtenu avec le classifieur KNN. Donc, en termes de taux de reconnaissance, l’utilisation de chacun des deux classifieurs sur ces données donne quasiment les mêmes performances dans l’approche monomodale. L’écart entre les taux de reconnaissance obtenus avec les 2 classifieurs est très faible : d’aucuns pourraient penser à une erreur de mesure mais non, il n’y en a point. Nous verrons dans l’approche multimodale si cet écart reste aussi faible ou pas afin de pouvoir conclure.

Le tableau 2.2 récapitule les différents résultats des méthodes de l’état de l’art et permet de comparer la performance de notre méthodologie et celles de l’état de l’art.

Méthodes	EMG	RESP	GSR	BVP	Taux de reconnaissance %
Baseline	39%	42%	42%	non	41
Fission basée sur les caractéristiques HHT	52%	60%	39%	non	50.33
Fusion basée sur les caractéristiques HHT	44%	59%	30%	non	44.33
Méthode proposée avec KNN	62.2%	52.2%	61.2%	53.38%	57.24
Méthode proposée avec SVM	57.48%	57.19%	58.67%	58.14%	57.84

TABLE 2.2 – Comparaisons des différentes méthodes de détection uni-modale instantanée des émotions

Les protocoles d’induction des émotions utilisés pour la génération des bases de données sont assez similaires dans toutes les méthodes présentées dans le tableau 2.2. De plus, chacune de ces méthodes ont été développées pour une détection instantanée et automatique des différents états émotionnels en utilisant les signaux physiologiques. Il existe plusieurs travaux sur la détection des émotions utilisant d’autres sources d’information que celle physiologiques que nous explorons dans cette thèse ou des travaux ne permettant pas des détections instantanées des émotions. Nous avons présenté ces différents travaux dans la section 1.1.2 et nous avons positionné nos travaux par rapport à ces derniers. La comparaison que nous faisons ne concerne pas ces travaux.

Nous remarquons selon les résultats présentés dans le tableau 2.2 que pour les modalités EMG et GSR, la méthode que nous proposons donne les taux de reconnaissance les plus élevés pour chacun de nos classifieurs. Pour ces modalités, les taux sont améliorés d'au moins 10,2% par rapport à la méthode de fusion basée sur les caractéristiques HHT et d'au moins 13,48% par rapport à la méthode de fusion basée sur les mêmes caractéristiques HHT [Kim et al., 2004]. Les méthodes de fusion et de fusion basées sur les caractéristiques HHT permettent d'avoir de meilleurs taux par rapport à notre méthode pour la modalité RESP. Pour chacune des modalités, nous obtenons de meilleurs résultats par rapport à la méthode Baseline [priya Muthusamy, 2012]. Nous voyons bien dans le tableau 2.2 que notre méthode est plus avantageuse en termes de taux de reconnaissance que les méthodes de l'état de l'art. La modalité BVP a été utilisée dans nos travaux contrairement aux autres travaux présentés dans le tableau 2.2.

Pour améliorer les taux de reconnaissance obtenus avec cette approche monomodale, nous proposons une méthode multimodale de reconnaissance des émotions basée sur la méthode de fusion tardive. Cette méthode de fusion est encore appelée fusion au niveau des décisions. Nous détaillons cette méthode multimodale de détection des émotions dans les sections suivantes.

2.3 Approche multimodale

Les émotions humaines par nature sont multimodales. En effet, une émotion ne se manifeste pas sur un seul canal mais sur plusieurs canaux simultanément. Par exemple, lorsqu'une personne est brusquement prise de peur, non seulement son rythme cardiaque s'accélère, elle se met aussi à trembler, à transpirer un peu plus, à rougir (dépend de l'origine) des fois. De plus sa voix, ses gestuelles sont affectées. Pour se rapprocher de la nature de l'émotion, il est nécessaire de proposer un système multimodal de reconnaissance des émotions. Dans le but d'interpréter au mieux les émotions des personnes qui nous entourent, nous utilisons toutes les sources d'information à notre disposition (l'intonation de leurs voix, leurs transpirations, sont-elles devenues pâles ? tremblent-elles ? ...). Ainsi, l'approche multimodale permet, en plus de disposer de meilleurs taux de reconnaissance de donner plus de robustesse lorsque l'une de ces modalités est acquise dans un environnement bruyant [Pantic and Rothkrantz, 2003]. Dans notre approche, nous avons choisi la multi-modalité utilisant plusieurs sources d'information au niveau physiologique. C'est à dire nous allons utiliser différents types de signaux physiologiques afin de proposer un système de reconnaissance multimodal des émotions. Pour proposer ce système, il est nécessaire dans un premier temps de fusionner les informations des différents signaux physiologiques.

2.3 Approche multimodale

2.3.1 les différentes techniques de fusion des signaux

Dans la littérature, il existe trois méthodes de fusion des signaux acquis de différents capteurs [Wagner et al., 2011] : fusion au niveau *signal* (fusion des signaux bruts), fusion au niveau *vecteurs caractéristiques* (fusion des vecteurs caractéristiques extraits dans chaque signal) et la fusion au niveau des *décisions monomodales* (fusion des vecteurs émotionnels déterminés avec l'approche monomodale).

2.3.1.1 Fusion au niveau signal

Cette fusion est directement effectuée sur les données brutes issues des capteurs de signaux physiologiques (voir figure 2.10). Elle a pour but de construire un signal physiologique à partir des différents signaux physiologiques. En raison des problèmes de synchronisation des capteurs, de résolution temporelle entre capteurs, de sensibilité au bruit des capteurs ... ; cette technique de fusion se trouve être très rarement utilisée [Hamdi, 2012].



FIGURE 2.10 – Fusion au niveau signal

2.3.1.2 Fusion au niveau vecteurs caractéristiques

Cette méthode de fusion (voir figure 2.11) réalisée sur les vecteurs caractéristiques est la plus utilisée. Elle a pour but de fusionner les vecteurs caractéristiques (un pour chaque signal physiologique) afin de construire un seul vecteur caractéristique multimodal. Cette méthode de fusion est moins sensible au bruit que la première. De plus, elle conserve des informations moins détaillées que dans la méthode de fusion des signaux car elle nécessite une sélection des paramètres importants [Sharma et al., 1998].

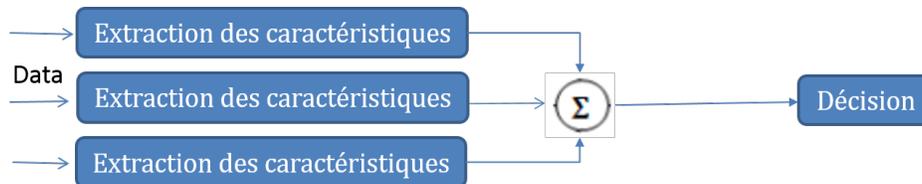


FIGURE 2.11 – Fusion au niveau des caractéristiques.

Plusieurs techniques de fusion au niveau des vecteurs caractéristiques ont été proposées dans la littérature [Zhi et al., 2001] notamment la concaténation des vecteurs caractéristiques uni-modaux [Huang et al., 1998], la méthode KNN [Tayari et al., 2013], l'analyse en composante principales [Deng et al., 2005], l'analyse

linéaire discriminante. La fission et la fusion basée sur la transformée de Hilbert-Huang (Hilbert-Huang Transform : HHT) dans l'extraction de la fonction appelée mode intrinsèque (Intrinsic Mode Function : IMF) en appliquant la décomposition en mode empirique [Zong and Chetouani, 2009] sont d'autres techniques utilisées à ce niveau.

- **Fission** : Sur chacune des fonctions mode intrinsèque, on extrait deux caractéristiques qui sont la fréquence instantanée et l'amplitude en utilisant la transformée de Hilbert-Huang.
- **Fusion** : On détermine l'utilisation de la fréquence moyenne pondérée et de la fréquence instantanée moyenne.

Cette technique de fusion sur les caractéristiques HHT a permis d'avoir un taux de reconnaissance de 62% tandis que la méthode hybride *Sequential Floating Forward Search-Fisher Projection (SFFS-FP)* du MIT a permis d'avoir un taux de 83% [Healey, 2000]. Cependant, la méthode SFFS-FP permet de déterminer les émotions durant un intervalle de temps fixé, elle ne permet donc pas une détection instantanée des émotions. Basée sur un scénario, la méthode SFFS-FP calcule les caractéristiques d'une émotion donnée, ensuite, réduit le nombre de données caractéristiques. Ainsi, elle ne permet non pas une détection instantanée mais une détection des états émotionnels sur un intervalle de temps relativement long. Contrairement à cette méthode, la nôtre permet une détection instantanée et automatique des états émotionnels.

2.3.1.3 Fusion au niveau des décisions monomodales

Cette technique de fusion encore appelée fusion tardive (voir figure 2.12) consiste à prendre une décision uni-modale pour chaque modalité et ensuite à fusionner ces différentes décisions afin d'avoir une vision plus globale. Contrairement aux deux autres méthodes, elle a l'avantage d'être indépendante de la nature des signaux et des caractéristiques utilisées pour la prise de décision. De plus, elle est très intuitive et s'implante facilement.

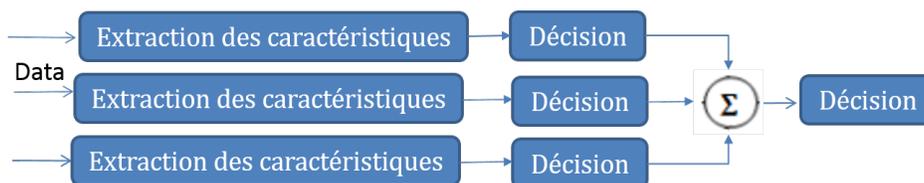


FIGURE 2.12 – Fusion au niveau des décisions.

Parmi toutes les techniques de fusion à ce niveau, la méthode de vote qui est aussi celle pour laquelle nous avons opté est la plus utilisée. À la sortie de chaque classifieur uni-modal, on construit un vecteur émotionnel et ensuite à ce niveau, le but est de fusionner ces différents vecteurs émotionnels. Prenons un jeune homme du nom de Daniel qui vient poser une question à trois médecins. Si la question de Daniel concerne le domaine de compétences de l'un des trois, bien qu'il écouterait toutes les réponses, il aura quand-même tendance à donner plus de crédit à la réponse donnée

2.3 Approche multimodale

par celui dont c'est le domaine de compétence. De la même façon, afin de prendre en compte la précision individuelle de chacun des capteurs physiologiques à détecter une émotion donnée, on a constitué une matrice de poids w_{ij} où sur les lignes on a les poids correspondants à un capteur et en colonne ceux des émotions. Ces différents coefficients w_{ij} ont été déterminés, d'une part, à l'aide d'une recherche dans l'état de l'art sur les caractéristiques de chaque capteur physiologique à reconnaître une émotion spécifique. D'autre part, par apprentissage statistique sur la pertinence de chaque capteur afin de déterminer un état émotionnel donné. Le mécanisme de vote implanté est donc un vote pondéré.

Nous pouvons dire que la méthodologie proposée est composée d'algorithmes assez intuitifs permettant de simplifier au maximum le traitement. Cette simplification de traitement est une contrainte que nous nous sommes imposés. Nous pourrions dans le futur envisager d'incorporer l'ensemble de nos algorithmes dans des appareils ne disposant pas d'une grande capacité de mémoire et de traitement pour stocker et traiter des algorithmes complexes. Notre simplification nous permettrait de garantir une bonne performance (taux de reconnaissance, temps d'exécution, autonomie) de notre système quelque soit l'unité de traitement dans lequel il est incorporée.

2.3.2 Résultats et discussions

A noter que l'ensemble de nos algorithmes ont été développés sur le logiciel Matlab. Ainsi, ce dernier a aussi été utilisé comme environnement de test. Le couplage de nos capteurs à Matlab pour l'acquisition des données est expliqué dans le chapitre 5.

Dans la suite de ce chapitre, nous avons considéré les signaux issus de quatre capteurs de signaux physiologiques *EMG*, *RESP*, *GSR*, *BVP*. En utilisant notre méthode DLF (Decision Level Fusion : fusion au niveau décisionnel), nous avons tout d'abord découpé la base de signaux fournis en trois bases comme dans l'étape mono-modale. Ensuite sur la base de validation, ayant opté pour le noyau gaussien, nous avons validé les paramètres c et γ permettant d'ajuster au maximum la gaussienne pour la classification de chacune des classes émotionnelles pour l'algorithme SVM. Le paramètre k pour l'algorithme de classification KNN a été validé sur les mêmes signaux fournis par le MIT que nous utilisons. Une fois les paramètres validés sur la base de validation, sur la base de test, nous avons vérifié la fiabilité de nos deux algorithmes de classification. Les résultats obtenus pour les algorithmes de classification KNN et SVM sont respectivement présentés sur les figures 2.13 et 2.14.

Aux vues de ces résultats, nous pouvons conclure que cette approche multimodale a permis d'améliorer considérablement le taux de reconnaissance des émotions par rapport à l'approche mono-modale. De plus, nos algorithmes permettent de détecter chaque activité émotionnelle et de distinguer cette dernière des autres avec de très bons taux de reconnaissance dont le minimum est d'environ 71%. Ce minimum est obtenu pour l'émotion *amour romantique* avec l'algorithme du KNN. En effet, avec l'algorithme du k -nearest neighbors, nous passons d'un taux de reconnaissance moyen de **57.24%** avec l'approche mono-modale à un taux de reconnaissance moyen de **81.69%** pour l'approche

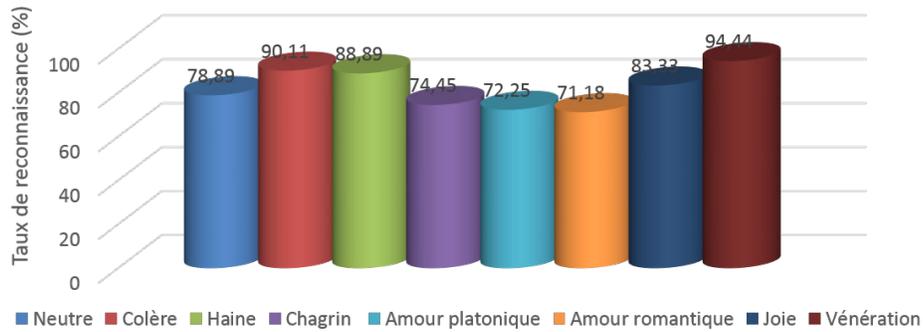


FIGURE 2.13 – Taux de reconnaissance multi-modal des émotions en utilisant l’algorithme du KNN.

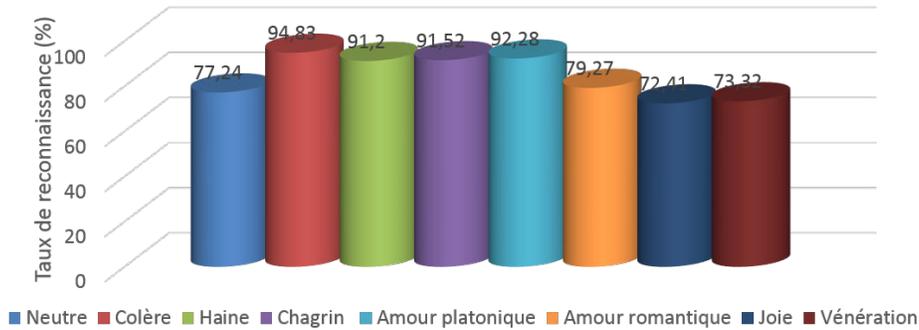


FIGURE 2.14 – Taux de reconnaissance multi-modal des émotions en utilisant l’algorithme du SVM.

multimodale. Avec l’algorithme support vector machine, nous avons obtenu un taux de reconnaissance moyen de **84%**. L’algorithme SVM a un meilleur taux de reconnaissance par rapport à l’algorithme KNN dans l’approche multimodale. Nous pouvons donc dire que sur ces données le SVM est plus efficace que le KNN car le taux de reconnaissance est amélioré à plus de **2%**.

En regroupant les émotions sur 3 classes suivant leurs valences.

- . Classe neutre : neutre, vénération
- . Classe négative : colère, haine, chagrin
- . Classe positive : amour platonique, amour romantique, joie

En prenant les taux d’erreurs de classification pour chaque émotion et en faisant la somme de ces différents taux sur chaque classe de valences, on trouve les taux d’erreurs regroupés dans le tableau 2.3.

Le tableau 2.3 permet d’avoir une idée sur la performance de notre méthodologie dans la distinction des différentes classes émotionnelles. Pour le KNN, sur les 37,67% d’erreur dans la classification d’une émotion de la classe neutre, 11% sont pour d’autres émotions de la même classe et 16,67% pour des émotions de la classe négative. Pour

2.3 Approche multimodale

Taux d'erreur avec l'algorithme KNN (%)			
Classe	neutre	négative	positive
neutre	11	16.67	0
négative	19.9896	24.34	2.2204
Positive	24.36	34.76	14.12
Taux d'erreur avec l'algorithme SVM (%)			
Classe	neutre	négative	positive
neutre	31.77	7.58	10.09
négative	1.28	7.52	13.65
positive	27.41	21.82	6.81

TABLE 2.3 – Taux d'erreur suivant les valences émotionnelles.

la classe négative, avec 24,34% le maximum d'erreur est obtenu pour des émotions de la même classe. Enfin quand notre algorithme fait une erreur dans la distinction d'une émotion de la classe positive avec le KNN, il détecte une émotion de la classe négative avec 34,76% contre 24,36% pour la classe neutre et 14,12% pour la positive. A défaut de proposer un système avec un taux de reconnaissance à 100%, l'objectif est de réduire les taux d'erreur de classification d'une classe à une autre. Pour l'algorithme du KNN, 33% du taux d'erreur sont pour une émotion de la même classe tandis que pour le SVM 36% du taux d'erreur sont pour une émotion de la même classe. Ces résultats montrent une fois de plus que l'algorithme du SVM est plus performante que celle du KNN sur ces données.

Le tableau 2.4 permet de faire une comparaison entre nos résultats et les différents résultats obtenus avec les méthodes de l'état de l'art qui permettent une détection instantanée des émotions.

Méthodes	Taux de reconnaissance %
Méthode de Kim [Kim et al., 2004]	61.2
Fusion basée sur les caractéristiques HHT [Zong and Chetouani, 2009]	62
Baseline [Tayari, 2013]	71
Méthode proposée avec KNN	81.69
Méthode proposée avec SVM	84

TABLE 2.4 – Comparaisons des différentes méthodes de détection multimodale instantanée

Les résultats présentés dans le tableau 2.4 montrent que la méthode que nous avons proposée avec chacun des deux algorithmes de classification permet d'avoir de meilleurs taux de reconnaissance que les méthodes de l'état de l'art. En effet, le taux de reconnaissance se trouve amélioré avec plus de 10%. Chacune des méthodes avec lesquelles nous comparons notre méthodologie permet une détection instantanée des

émotions et les protocoles d'induction des états émotionnels utilisés pour la conception des bases de données sont très proches.

2.4 Apport par rapport à l'existant

Nous avons travaillé sur les différents algorithmes définis au-cours des travaux de thèse d'Imen Tayari pour l'approche monomodale en vue de les optimiser. Cette optimisation visait à réduire les temps de traitement et à rendre les algorithmes plus performants en termes de taux de reconnaissance. Ensuite, nous avons ajouté une étape de validation des paramètres des classifieurs dans la phase d'apprentissage. La validation des paramètres des classifieurs vise à définir des fonctions de prédictions optimales. En effet, les fonctions de prédiction définies et validées aboutissent à une généralisation de la classification. Dans ce sens, les performances obtenues sur des données inconnues au moment de la définition de ces fonctions seront quasiment les mêmes que celles obtenues sur les données connues (données d'apprentissage et de validation). Elle permet une comparaison plus rigoureuse de différents algorithmes.

Après nos travaux sur la généralisation de notre méthode de détection des émotions, nous avons implanté un nouvel algorithme de classification supervisée et développé une approche multimodale pour la reconnaissance des émotions. Cette approche est basée sur la méthode de fusion des décisions monomodales (un vecteur émotionnel calculé pour chaque type de signal). Avec un mécanisme de vote pondéré, où les pondérations ont été déterminées, d'une part, à l'aide d'une recherche dans l'état de l'art sur les caractéristiques de chaque capteur physiologique à reconnaître une émotion spécifique, nous avons défini notre approche multimodale. Nos travaux ont permis d'améliorer considérablement les performances du système de détection des émotions. Les performances en temps de traitement de notre système de reconnaissance instantanée des états émotionnels d'une personne sont montrées dans le chapitre 5.

2.5 Conclusion

Dans ce chapitre, après avoir positionné nos travaux comme une continuité des travaux de Dr. Tayari; nous avons dans un premier temps proposé une méthode unimodale de détection instantanée des émotions en utilisation deux algorithmes de classification à savoir le KNN et le SVM. En effet, notre méthodologie de détection étant basée sur la détection d'un pic considéré comme la présence d'une activité émotionnelle a donc nécessité l'utilisation de quelques outils de traitements de signal afin de proposer un système efficace. Dans l'approche uni-modale, la méthode proposée a permis d'avoir un taux de reconnaissance des émotions raisonnables. Toutefois, face à notre insatisfaction vis à vis de ces taux de reconnaissance, nous avons donc étudié la multi-modalité. Cette étude nous a permis de proposer une méthode multimodale de détection automatique des émotions. Cette méthode est basée sur la fusion au niveau des décisions monomodales. En effet, après avoir filtré le signal de chaque modalité

2.5 Conclusion

et détecté une activité émotionnelle pour chacune des modalités, nous déterminons les différents vecteurs émotionnels (autant de vecteurs émotionnels que de modalités, si aucune activité émotionnelle n'est détectée dans les signaux d'un capteur alors aucun vecteur émotionnel n'est calculé pour ce dernier). Ensuite, ces différents vecteurs émotionnels sont fusionnés avec la méthode de vote pondéré. La méthode de vote pondéré permet de prendre en compte, la capacité de chacun des capteurs à détecter une émotion spécifique, et aussi, d'avoir une vision plus globale de l'émotion ressentie car les émotions sont de nature multimodale. Dans l'étape de classification, nous avons étudié et implanté deux algorithmes de classification à savoir l'algorithme des k plus proches voisins (KNN) et celui des séparateurs à vaste marge (SVM). Ces algorithmes ont été implantés de la façon la plus intuitive possible. Cette contrainte visait à optimiser au mieux notre méthodologie en vue de pouvoir proposer plus tard une version mobile de notre système avec les mêmes performances. Ainsi, avec l'algorithme des k plus proches voisins, nous avons un taux de 81.69% et avons obtenu un taux 84% pour l'algorithme des séparateurs à vaste marge.

Après avoir validé notre méthodologie de détection des émotions sur la base de données de signaux physiologiques fournie par le MIT, nous avons décidé de générer notre propre base de données liée à l'application d'e-santé que nous visons. La construction de cette base de données requiert l'utilisation d'un système d'acquisition dont l'architecture est à définir. Dans le chapitre suivant, nous présentons l'architecture globale d'un système d'acquisition pour la détection d'émotions et ainsi que notre modèle analytique d'estimation de la consommation énergétique de systèmes d'objets communicants.

Architecture du système d'acquisition et autonomie

Sommaire

3.1	Architecture de collecte de donnée pour la détection des émotions	74
3.2	Modèle de consommation	76
3.2.1	Quelques protocoles de communications	77
3.2.2	Modèle analytique d'estimation de la consommation et de l'autonomie	79
3.2.3	Caractérisation des paramètres	83
3.2.4	Test et résultats	85
3.3	Étude de l'influence des choix architecturaux sur la qualité de la détection	91
3.4	Conclusion	92

3.1 Architecture de collecte de donnée pour la détection des émotions

Dans le but de concevoir un système complet de détection des émotions, nous avons défini une chaîne d'acquisition et de traitement des données. Cette chaîne comporte, d'une part, un système d'acquisition adapté à l'application considérée et chargé de collecter et de transmettre les données via des communications sans fil. D'autre part, cette chaîne comporte un système de réception chargé du traitement et de l'interprétation de ces données. Le système d'acquisition a pour rôle de collecter des informations comme, par exemple, la pression, la température, le mouvement, ou comme dans le cadre cette thèse, elles peuvent représenter des signaux physiologiques... Ces informations servent à connaître l'état du système sur lequel les données sont collectées.

3.1 Architecture de collecte de donnée pour la détection des émotions

Ce système affecte une valeur à la grandeur caractérisant cet état. Il délivre alors ces informations sous la forme appropriée (un transducteur convertit la valeur en un signal électrique, le convertisseur analogique/numérique la convertit ensuite en binaire) pour une exploitation ultérieure.

Le système d'acquisition se compose de plateforme(s) de collecte de données dotée(s) de capteurs de signaux physiologiques, d'une unité de traitement, et d'un module de communication sans fil. Au sein de ces plateformes, l'unité de traitement assure d'une part la gestion : de l'ordonnancement des acquisitions, de la lecture et de la conversion des données acquises. D'autre part, elle sert à sauvegarder ces données dans la mémoire, et à préparer l'envoi des données qui sera réalisé par le module radio en charge des communications sans fil. Dans un premier temps, ces données seront envoyées à un ordinateur sur lequel les algorithmes pour la détection des émotions sont implantés. Il est aussi possible d'exécuter cette application sur un mobile Android. Dans ce cas, les données seront directement envoyées à ce dernier. On pourrait envisager de délocaliser une partie du traitement liée à la détection des émotions vers les plateformes d'acquisition. Cette délocalisation aurait pour conséquence d'augmenter la taille du code à exécuter et le temps d'activité, entraînant ainsi une réduction de l'autonomie de la plateforme d'acquisition. La figure 3.1 présente l'architecture du système d'acquisition et de traitement pour notre application de détection des émotions. Dans le cadre de

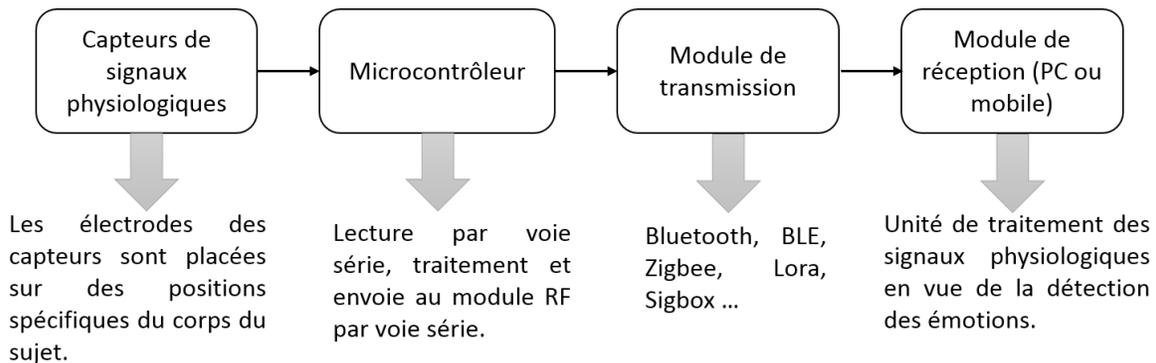


FIGURE 3.1 – Chaîne d'acquisition et de traitement pour la détection des émotions

notre application, la qualité de la détection comme l'autonomie du système d'acquisition constitue des contraintes primordiales. La qualité de la détection dépend des traitements qui ont été évalués dans le chapitre précédent mais aussi de différentes caractéristiques liées aux signaux collectés. Concernant l'autonomie du système d'acquisition, il est nécessaire de l'évaluer au plus haut niveau de la conception. Il est ainsi essentiel de disposer d'un modèle de haut niveau de la consommation d'énergie et de la puissance. Ces modèles permettront aussi de réaliser de l'exploration d'architecture, de configurations et d'évaluer la qualité des solutions qui permettrait de réduire la consommation d'énergie.

Notre objectif est de rechercher le meilleur compromis entre les performances (qualité de la détection) et l'autonomie par une exploration d'architectures et de

leurs configurations appropriées. Cette exploration permettra d'étudier l'influence des différents types d'unité de traitement, de capteurs mais aussi, des protocoles de communications sans fil. Dans ce qui suit, nous présentons brièvement les différents niveaux de modélisation de la consommation afin de positionner nos travaux. Ensuite, nous décrirons les protocoles de communications sans fil que nous avons utilisés avant de présenter le modèle d'estimation de la consommation de systèmes d'objets communicants pour l'e-santé que nous avons développé.

3.2 Modèle de consommation

Dans le domaine de la santé, le développement et la vente d'un objet comme un détecteur d'émotion ne peut être possible que si son encombrement et son autonomie apparaît comme acceptable vis à vis de l'utilisateur. Pour répondre au problème de l'autonomie, il est essentiel lors de la conception de pouvoir estimer la consommation d'énergie et de puissance, ce qui nécessite de définir des modèles adéquats. Cette estimation permettra de sélectionner le système le plus adapté en termes de consommation garantissant les contraintes d'autonomie.

Afin de proposer un modèle d'estimation de la consommation, il est d'abord nécessaire de discuter du niveau d'abstraction dans lequel il sera défini. Dans la section 1.3.4, nous avons présenté les différents niveaux d'abstraction ainsi que les outils et les techniques d'estimation de la consommation. Cependant, nos travaux ne ciblent pas la conception matérielle et électronique des capteurs, des unités de traitement et des modules radios, mais plutôt la construction d'architectures, regroupant ces différents composants et la définition de leurs configurations. De ce fait, les composants que nous envisageons d'utiliser sont déjà existants sur le marché. Il faut, de plus, noter que l'objectif est de concevoir des architectures capables de satisfaire les contraintes applicatives et utilisateurs. Ne disposant pas d'une description très précise, des composants disponibles sur le marché, aux deux premiers niveaux d'abstraction c'est à dire technologique et logique, nos études ont donc été menées aux niveaux *architectural* et *système*. Dans cette section, nous proposons un modèle analytique d'estimation de la consommation au niveau architectural. Ce modèle analytique est basé sur une décomposition en bloc fonctionnel du système d'acquisition. Ce modèle d'estimation de la consommation a été ainsi choisi afin de pouvoir réaliser une exploration d'architectures qui sera présentée au chapitre suivant.

Dans de nombreuses applications de type réseau de capteur sans fil, la consommation d'énergie du module radio est indiquée comme prépondérante. Aussi, afin d'évaluer l'impact du protocole de communication, dans le cadre de notre application, sur l'autonomie du système d'acquisition, nous commençons par présenter les caractéristiques principales de quelques protocoles de communication sans fil.

3.2 Modèle de consommation

3.2.1 Quelques protocoles de communications

Un protocole de communication peut être défini comme l'ensemble des règles, principes, méthodes régissant la communication d'entités de systèmes distribués par échanges de messages. Il est à la charge de la définition de formats pour les messages, il régle l'échange des données et définit un vocabulaire pour les messages valides (unités de données du protocole) et leur signification. Nous pouvons donc simplement dire que le protocole de communication est une infrastructure logicielle/matérielle garantissant la bonne transmission/réception d'une information donnée. De nombreuses approches ont été développées au sein des protocoles de communication pour répondre aux besoins des objets connectés de plus en plus divers et variés, des réseaux de capteurs sans fil. Certains de ces protocoles ont été standardisés et utilisés pour différents domaines applicatifs comme la surveillance industrielle, la santé, la domotique ou l'agriculture. Une application présente des besoins spécifiques en terme de communication : certaines nécessitent un grand débit alors que d'autres privilégieront une connexion hautement sécurisée et une longue autonomie. De ce fait, suivant l'application visée, les protocoles sont étudiés selon différents critères comme le débit, la sensibilité qui caractérisent la portée, la consommation qu'ils engendrent.

Dans le cadre de ces travaux, nous avons choisi d'étudier des protocoles présentant un débit adapté aux exigences de l'application tout en ayant une faible consommation. De plus, pour mener à bien notre étude sur la consommation, il était indispensable aussi de disposer de module radio sur lequel des mesures de consommation pouvaient être réalisées. Aussi nous avons sélectionné les protocoles suivant : Bluetooth, Zigbee et LORA (avec deux types de modulations). Le tableau 3.1 regroupent les caractéristiques principales des quatre protocoles de communication que nous avons choisi d'utiliser dans nos travaux. La figure 3.2 représente une classification des quatre protocoles explorés dans ces travaux de thèse.

Protocole	Bluetooth	Zigbee	Lora	
			modulation Lora	modulation GSKF
Standard IEEE	802.15.1	802.15.4	802.15.4	
Autonomie avec pile	jours	Années [info, 2017]	Années (10-20) [ADEUNIS, 2015]	
Nombre de nœuds	7	>65 000	>65 000	
Débits	1 Mbits/seconde	20- 250 Kbits/seconde	10,937 Kbits/seconde	50 Kbits/seconde
Portée	10-100 mètres	100 mètres	jusqu'à 20km	

TABLE 3.1 – Récapitulatif des protocoles de communication utilisés dans notre travail

Un protocole de communication s'appuie sur différentes couches protocolaires qui interagissent entre elles. Dans le cadre de notre application, le nombre de plateformes d'acquisition étant faible, la couche protocolaire réalisant le routage est relativement

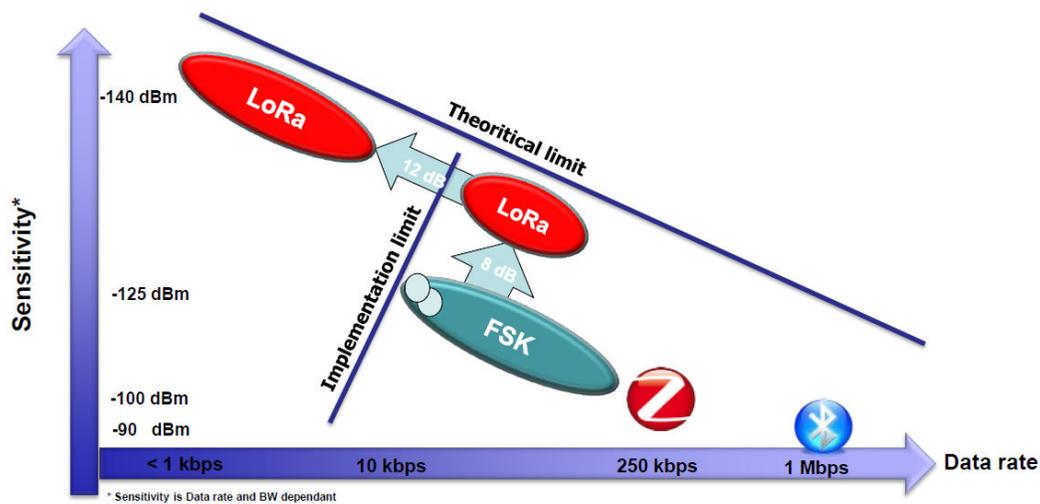


FIGURE 3.2 – Positionnement des 4 protocoles de communication suivant le débit et la sensibilité [Semtech, 2017]

simpliste. Concernant la couche MAC et la couche PHY, dans le modèle OSI, elles sont définies par le standard IEEE. Les spécifications de la couche PHY requièrent que la radio opère dans une des 3 bandes de fréquence de la bande ISM (Industrial, Scientific and Medical). Ces bandes sont 868-868.8 MHz (Europe, 1 canal), 902-928 MHz (Amérique du Nord, 30 canaux), et 2400-2483.5 MHz (mondialement, 16 canaux). Le débit lié à un protocole dépend entre autre de sa bande de fréquence. Par exemple, à 2,4 GHz pour le ZigBee, le débit peut être de 250 koctets par seconde. La couche PHY permet de configurer le matériel. A titre d'exemple, on peut citer le choix du canal, la puissance d'émission, la sensibilité à la réception, la modulation. De nombreux systèmes de modulation sont présents, mais les plus communs dérivent du DSSS (Direct Sequence Spread Spectrum). Pour réguler l'accès au canal, la couche MAC utilise souvent la technique CSMA-CA (Carrier Sense Multiple Access with Collision Avoidance). Un nœud avant d'envoyer une trame écoute d'abord le canal pour s'assurer qu'aucune activité est en-cours, et s'il n'y a pas d'activité, le nœud peut alors commencer à envoyer les données. Dans le cas où le canal est utilisé, c'est à dire lorsque le nœud détecte une activité sur le canal, alors, il attend pendant une période aléatoire puis vérifie de nouveau si le canal est occupé. Le nombre de tentatives d'accès au canal est limité pour que le nœud ne reste pas en état perpétuel d'attente pour envoyer une trame.

Les travaux présentés dans ce mémoire ne ciblent pas : la validation des protocoles de communication, la modélisation des couches protocolaires. L'objectif est plutôt de proposer une approche générique pour la modélisation de la consommation des modules radio pour différents protocoles de communication (dont ceux mentionnés dans ce paragraphe) et de coupler ce modèle avec celui des capteurs et de l'unité de traitement qui composent le système d'acquisition. Le but final est d'estimer l'autonomie de ce système pour une batterie donnée. Dans ce chapitre et le suivant, nous avons

3.2 *Modèle de consommation*

caractérisé et modélisé quatre protocoles de communication : Bluetooth, zigbee, lora avec la modulation Lora et la modulation GFSK (Gaussian Frequency Shift Keying modulation). Cette modélisation utilise leurs particularités comme leurs débits, leurs consommations, leurs réponses temporelles. Il faut aussi noter que les dimensions du système final c'est à dire le poids et la taille constituent une contrainte qui peut se décliner par l'encombrement du dispositif électronique mais aussi par son alimentation c'est à dire par la batterie qui est chargée de lui conférer une autonomie raisonnable.

3.2.2 **Modèle analytique d'estimation de la consommation et de l'autonomie**

De nombreux modèles analytiques permettant d'estimer la consommation des systèmes embarqués ont été développés et ont été décrits dans l'état de l'art. Ces modèles se concentrent généralement sur une fonction particulière du système afin de pouvoir par la suite jouer sur cette fonction ou module pour réduire la consommation du système complet. Ces modules élémentaires, dans le cadre de cette thèse, peuvent être des capteurs, des unités de traitement comme des microcontrôleurs, des systèmes d'exploitations, des mémoires, des modules radio. On peut à titre d'exemple citer les travaux suivant [Zhu et al., 2004], [Fornaciari et al., 1998], [Li and John, 2003], [Shelby et al., 2005], [Benini et al., 2001], [Tan et al., 2002]. Contrairement à ces travaux, nous nous intéressons à l'estimation de la consommation d'une plateforme d'acquisition complète. Ce modèle doit prendre en compte l'ensemble des paramètres qui conduisent à une configuration spécifique de la plateforme.

Dans le but de détecter les émotions, nous avons besoin de concevoir un système d'acquisition capable de collecter des signaux physiologiques tout en satisfaisant les contraintes applicatives, comme la qualité de la détection, ou celles liées à l'utilisateur, comme l'autonomie du système développé. D'autres contraintes sont aussi à prendre en compte comme le prix et les dimensions de la plateforme. La qualité de la détection dépend, entre autre de la qualité des signaux physiologiques et de l'algorithme de détection des émotions. Par exemple, dans l'avenir, nous prévoyons d'embarquer ce système. Pour cette raison, nous aurons besoin d'algorithmes peu gourmands en capacité de calcul et en temps de traitement, ce qui risque de diminuer le taux de reconnaissance des émotions c'est à dire la qualité de la détection. L'autonomie du système d'acquisition dépend de l'utilisation des ressources qui est liée au traitement local (du plus simpliste au plus compliqué) et à la qualité des données qui seront transmises et qui représentent les signaux physiologiques. Aussi, afin de tenir compte de ces deux aspects, notre but est de développer un modèle à haut niveau d'abstraction couplant l'estimation de l'autonomie et de la performance du système afin de rechercher l'architecture et la configuration qui présente le meilleur compromis ou capable de satisfaire les deux contraintes qui seraient fixées au préalable. La notion de performance dans la suite de ce document représente le taux de reconnaissance des émotions. Ce dernier est influencé, pour un algorithme de détection des émotions donné, par la qualité des signaux. Cette qualité dépend dans notre étude de la fréquence d'échantillonnage, de la qualité des

données qui varie en fonction de la qualité des capteurs et du nombre de capteurs. Afin de caractériser le système d'acquisition, nous proposons une approche basée sur l'identification de phases de fonctionnement permettant d'acquérir et de transmettre les données. Ces phases, au nombre trois, comportent deux phases pendant lesquelles le système est actif et différentes ressources sont activées et une phase de repos entre deux périodes d'acquisition/transmissions. La définition des ressources activées dans chaque phase ainsi que le temps passé dans ces phases vont permettre de calculer la consommation énergétique. Ces phases sont les suivantes :

- *Phase 1* : Association de la plateforme d'acquisition (composée des capteurs, du microcontrôleur et d'un module de transmission) au système de réception et de traitement (PC, mobile, ...). Cette association s'appuie sur un protocole de communication sans fil dans notre cas. Le microcontrôleur et le module radio sont activés. cette phase sera nommée phase de connexion.
- *Phase 2* : Acquisition et transmission(s) des données. Dans cette phase, les capteurs, le module de transmission et le microcontrôleur (MCU) sont alimentés.
- *Phase 3* : Dans cette dernière phase (phase de repos ou attente entre deux acquisitions/transmissions), tous les capteurs sont éteints, le module radio est désactivé, le MCU est actif ou peut être en mode repos.

La définition d'un modèle à haut niveau d'abstraction, pour estimer la consommation de plateforme mobile d'acquisition de signaux physiologiques doit tenir compte d'un ensemble de paramètres qui sont de deux types, soit architecturaux soit liés au scénario applicatif. Les paramètres architecturaux regroupent la consommation individuelle de courant des composants de la plateforme dans chaque phase, les temps nécessaire pour l'exécution de certaines fonctions, comme l'association, l'acquisition, la transmission, le traitement des données. A ces paramètres s'ajoutent ceux liés à la configuration de la plateforme comme la fréquence d'échantillonnage. Enfin, le scénario applicatif va permettre de spécifier les périodes d'acquisition et le temps nécessaire pour déterminer une tendance émotionnelle. Afin de déterminer l'autonomie du système d'acquisition, il faut spécifier la capacité de la batterie. Les différents paramètres utilisés dans notre modèle sont :

1. A : Autonomie de la plateforme exprimée en heures
2. Paramètres architecturaux
 - $T_{sensing}$: Temps d'acquisition
 - $T_{processing}$: MCU en mode actif, correspond au temps au-cours duquel le mcu acquiert puis transmet ces données
 - $T_{con_{radio}}$: Temps de connexion au module radio
 - $T_{trans_{screen}}$: Temps de transmission d'un échantillon
 - $I_{con_{radio}}$: Courant consommé durant la phase de connexion (phase 1)
 - $I_{trans_{radio}}$: Courant consomme durant la phase de transmission (phase 2)
 - $I_{sensors}$: Courant consommé par l'ensemble des capteurs
 - I_{static} : Courant statique du mcu

3.2 Modèle de consommation

- I_{mcu} : Courant additionnel du mcu dû au traitement. Ce courant peut être assimilé au courant dynamique du microcontrôleur.

$$I_{mcu_{total}} = I_{static} + I_{mcu}$$

3. Paramètre de configuration : F_s : Fréquence d'échantillonnage en Hz
4. Paramètres liés au scénario applicatif
 - T_{IE} : Temps d'activité de la plateforme, exprimé en heure
 - Nb_{IEJ} : Nombre d'interprétation par jour des émotions
5. Paramètres intermédiaires :
 - Nb_{trans} : Nombre de transmission nécessaire pour transmettre l'ensemble des signaux physiologiques acquis
 - Nb_{IE} : Nombre d'interprétation des émotions durant toute la durée de vie de la batterie. Ce nombre correspond au nombre de fois qu'on peut faire l'acquisition des signaux physiologiques sur la durée de vie de la batterie.

$$Nb_{IE} = \frac{A}{T_{IE}}$$

6. Paramètre lié à la batterie : $Q_{available}$: Quantité de courant utilisable et disponible sur la batterie

$$Q_{available} = \left(1 - \frac{1}{n}\right) * Q_{theoretic}$$

n dépend des caractéristiques de la batterie. Grâce à des campagnes de mesures sur la décharge des batteries qui étaient à notre disposition, nous avons trouvé une valeur de $n = 5$.

La pile (batterie) est vue comme une source finie de consommation. On fait donc l'hypothèse qu'elle contient une quantité linéaire de charge. La quantité de charge totale consommée par le système de détection des émotions peut-être représentée par la somme des quantités de charge consommées par les capteurs, l'unité de traitement et le module radio comme l'indique la formule suivante :

$$Q_{total} = Q_{radio} + Q_{processing} + Q_{sensing} \quad (3.1)$$

Où Q_{radio} , $Q_{processing}$, $Q_{sensing}$ représentent respectivement la quantité de courant consommée par les modules radio, l'unité de traitement et par les capteurs. Ces quantités de courant sont calculées comme suit :

$$Q_{sensing} = I_{sensors} * T_{sensing} * Nb_{IE} \quad (3.2)$$

$$Q_{processing} = I_{mcu} * T_{processing} * Nb_{IE} + I_{static} * A \quad (3.3)$$

$$Q_{radio} = Nb_{IE} * (Q_{con_{radio}} + Nb_{trans} * Q_{trans_{radio}}) \quad (3.4)$$

Où $Q_{con_{radio}}$, $Q_{trans_{radio}}$ représentent respectivement la quantité de courant consommé par le module radio durant la phase de connexion (Association) et la phase de transmission.

$$Q_{con_{radio}} = I_{con_{radio}} * \frac{T_{con_{radio}}}{3600} \quad (3.5)$$

$$Q_{trans_{radio}} = I_{trans_{radio}} * \frac{T_{trans_{screen}}}{3600} \quad (3.6)$$

Pour prendre en compte l'influence de la fréquence d'échantillonnage sur l'autonomie du système, une campagne de mesures préliminaires exploratoire a été réalisée. La plateforme d'acquisition était composée de différents capteurs de signaux physiologiques pour lesquels la fréquence de collecte des données était variable ainsi que la capacité de la batterie. À partir de cette exploration, nous avons remarqué que la consommation totale varie linéairement en fonction de la fréquence d'échantillonnage :

$$Q_{total} = Q_{available} - k_2 - k_1 * Fs \quad (3.7)$$

Où $k_1 = 1.49$ ($S * mAh$) and $k_2 = 18.5$ (mAh).

k_1 et k_2 sont des constantes calculées à partir d'une exploration des données par projection dans un repère.

À partir de ces différentes équations, on peut donc en déduire un modèle générique d'estimation de l'autonomie.

$$A = \frac{Q_{available} - (k_1 * Fs + k_2)}{I_{static} + I_{others}} \quad (3.8)$$

Avec

$$I_{others} = \frac{I_{mcu} * T_{processing} + I_{sensors} * T_{sensing} + \frac{Q_{radio}}{NbIE}}{T_{IE}} \quad (3.9)$$

Les valeurs des courants utilisés dans notre modèle ont été obtenues pour la plupart par des mesures à l'exception de quelques unes qui sont issues de documentation technique. Pour les modules radios, pour le protocole Bluetooth, nous avons fait une campagne de mesure à partir de la plateforme qui sera présentée dans la section suivante. Pour les autres protocoles (Zigbee et LORA), nous avons utilisé des résultats d'autres travaux réalisés au sein du laboratoire LEAT et/ou de documents techniques. Par ailleurs, les courants consommés par les capteurs, le microcontrôleur ont été caractérisés à partir de mesures réelles.

À Ce modèle qui permet une estimation de la consommation d'énergie et de l'autonomie d'objets communicants pour les applications d'e-santé est générique du fait que :

- . l'architecture considérée est modulaire et peut prendre en compte plusieurs configurations,
- . les paramètres du modèle ont été définis pour caractériser l'architecture type d'un système de détection des émotions. En effet, ce type de système d'acquisition est composé de façon générale de modules capteurs, d'un module de traitement et

3.2 Modèle de consommation

d'un module radio. Le modèle proposé tient compte des paramètres génériques de ces modules comme leur consommation individuelle, leur temps d'exécution, leurs fréquences de fonctionnement. Les paramètres liés au scénario applicatif ont été définis à un haut niveau d'abstraction afin d'être compatible avec d'autres applications.

Le tableau 3.2 présente un récapitulatif des paramètres du modèle dont les valeurs ont été mesurées sur la plateforme Bitalino.

Paramètres du modèle	Valeur mesurée
$T_{con_{radio}}$	6,932 secondes
$T_{trans_{screen}}$	1,099 seconde
$I_{con_{radio}}$	32,61 mA
$I_{trans_{radio}}$	25.25 mA
RX_{radio}	14,35 mA
$I_{sensors}$	10 mA (3 capteurs : ECG, EMG, EDA)
I_{static}	6,7 mA
I_{mcu}	11,05 mA à $Fs = 8Mhz$

TABLE 3.2 – Récapitulatif des paramètres mesurés sur Bitalino

3.2.3 Caractérisation des paramètres

Afin, d'une part, de caractériser les différents modules présents dans le système d'acquisition et de pouvoir valider le modèle d'estimation de l'autonomie, et d'autre part, de réaliser un système réel d'acquisition de signaux physiologiques pour la détection des émotions, nous nous sommes intéressés à différentes plateformes d'acquisition de signaux physiologiques. En plus des fonctionnalités présentées, nous avons analysé certaines caractéristiques comme le coût, la flexibilité, l'adaptabilité, l'autonomie Notre choix s'est finalement porté sur la plateforme de développement "BITalino" [da Silva et al., 2014]. Cette plateforme est un kit, équipé de bio-capteurs, qui vise à promouvoir le développement des dispositifs médicaux et notamment pour des applications de surveillance dans le domaine de la santé. Il possède tous les avantages d'un système d'acquisition de signaux physiologiques tout en restant bon marché. Il s'agit d'une plateforme de développement présentant un aspect modulaire, c'est à dire, permettant de détacher et d'ajouter des capteurs pour obtenir ainsi la combinaison de signaux physiologiques que nous voulons réaliser. Dans sa version initiale, cette plateforme d'acquisition possède :

- quatre capteurs (*ECG, EMG, GSR et accéléromètre*)
- un module radio avec le protocole de communication Bluetooth
- un microcontrôleur qui est du type ATmega328p
- une mémoire flash de 32 Kilo-octets
- une EEPROM de 1 Kilo-octets

- une SRAM 2 Kilo-octets (Static Random Access Memory)

Cette plateforme de développement offre la possibilité de régler la fréquence d'échantillonnage qui peut être calibrée à 1 Hz, 10 Hz, 100 Hz ou 1000 Hz. Cette configuration engendre un volume de données variable qui influe sur les périodes de transmission. Concernant la batterie, dans le kit initial, une autonomie de 4 heures est annoncée lorsqu'il est utilisé en continu en mode transmission. Il est possible d'adjoindre de nouvelles batteries qui présenterait une capacité de charge par exemple supérieure.

Nous avons ainsi procédé à l'extraction des paramètres du modèle défini pour l'estimation de l'autonomie, sur cette plateforme.

La première étape a été d'étudier la réponse temporelle de la plateforme Bitalino lors des différentes phases et par fonctionnalités. Cette étude a été menée en faisant varier la fréquence d'échantillonnage qui influe sur le nombre d'échantillons. Un des objectifs était d'évaluer l'influence de la fréquence d'échantillonnage sur le temps de transmission. Nous avons ainsi mesuré à l'aide de timers le temps lié à l'exécution des différentes phases (connexion, acquisition et transmission) et ceci pour différentes fréquences d'échantillonnage. Ainsi, on a noté que le temps de transmission était constant $T_{trans_screen} = T_{trans} = 1,099S$ car initialement, après chaque acquisition, une transmission est effectuée, la taille de l'échantillon ne varie donc pas, par contre le nombre de transmission est impacté. Dans un second temps, le temps de transmission a été mesuré pour des trames de longueur variable, longueur qui pourrait être calculée en fonction de la fréquence d'échantillonnage.

Dans la seconde étape, l'objectif était de mesurer la consommation de courant de la plateforme dans chacune des 3 phases. Cette étape nous a permis en plus de l'aspect extraction de paramètres pour le modèle de l'autonomie, de mettre en place une meilleure utilisation de ladite plateforme en termes de consommation. En se connectant aux bornes de la résistance de Shunt de la plateforme, nous avons mesuré les courants consommés par cette dernière dans les différentes phases et en fonction de l'activation des modules capteurs, radio. Ainsi le courant du microcontrôleur a été évaluée en premier, puis par différence, les courants consommés par les autres modules ont été extraits. Par exemple, les courants consommés par le module radio Bluetooth ont été évalués en faisant une différence du courant total consommé, lors de chaque phase, aux courants consommés par le microcontrôleur et les capteurs si ils étaient actifs. De ce fait, le courant consommé par le module Bluetooth est : $I_{trans_radio} = 25.25mA$, $RX_{radio} = 14.35mA$, $I_{static} = 6.7mA$.

Il existe trois modèles de Bitalino (voir figure 3.3). Ces modèles possèdent les mêmes fonctionnalités mais sont présentés sous différentes formats :

- **Board** : tous les capteurs sont pré-connectés au microcontrôleur.
- **Plugged** : Ce modèle est plus souple car les capteurs sont fournis en tant que composants individuels reliés par câble à la carte mère. Ce modèle donne de la flexibilité en terme de placement. Par exemple, la carte principale peut être placée à la taille du patient et les capteurs aux points de mesures adéquats.

3.2 Modèle de consommation

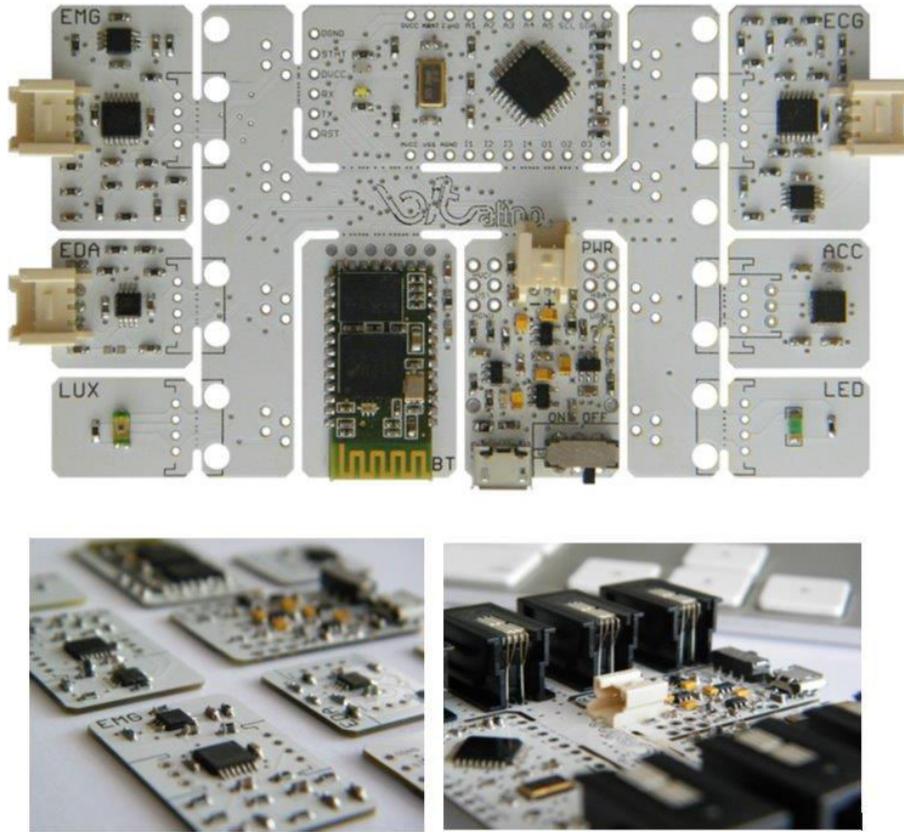


FIGURE 3.3 – Plate-forme Bitalino sous ses 3 formats à savoir Board, Freestyle et Plugged

- **Freestyle** : tous les blocs sont fournis individuellement, et on peut les disposer de façon à avoir un dispositif miniature mais il faut concevoir le support de communication.

3.2.4 Test et résultats

Afin d'évaluer l'autonomie du système d'acquisition, nous avons défini un scénario applicatif après discussion avec des médecins. Il a été proposé de faire l'interprétation des émotions deux fois par jour sur une durée de 20 minutes chacune. Selon ce scénario, Nb_{IEJ} qui est le nombre d'interprétation des émotions par jour sera pris égal à 2 ($Nb_{IEJ} = 2$ et T_{IE} qui représente le temps d'activité du système d'acquisition sera égal à 20 minutes :

$$T_{IE} = \frac{1}{3} \text{heure}$$

. Avec la plateforme BITALINO, nous disposons de trois capteurs de signaux physiologiques (ECG, EMG et GSR). Chaque capteur est relié à un ADC sur 10 bits et fournit une donnée codée sur 10 bits. Un échantillon possède ainsi une taille

de 30 bits pour les 3 capteurs. Afin de limiter le nombre de transmission qui sont consommatrices d'énergie, nous avons choisi contrairement à la configuration initiale de BITALINO, de n'opérer une transmission que lorsque la trame est remplie au maximum de sa capacité et de grouper le plus possible ces transmissions afin de réduire le temps d'activité de la radio. Ainsi lors des 20 minutes d'acquisition, les transmissions de données seront réalisées lorsque la mémoire de stockage des données est pleine et le nombre de transmissions adéquat sera effectué. Les équations, ci-après, permettent de calculer le temps maximal ($transmission_{period}$) qu'on peut attendre avant de transmettre les données acquises sans perte de données et le nombre de transmissions (Nb_{trans}) nécessaires pour transmettre l'ensemble des données acquises durant cet intervalle de temps.

$$nbSampleByte_{minute} = \frac{\lfloor ADC_{bitSize}/8 \rfloor}{60 * Fs} \quad (3.10)$$

$$transmission_{period} = \frac{protocole_{débit}}{nbSampleByte_{minute}} \quad (3.11)$$

$$Nb_{trans} = \lfloor \frac{60 * T_{IE}}{transmission_{period}} \rfloor \quad (3.12)$$

Où :

- $ADC_{bitSize}$ est le nombre de bit du convertisseur analogique numérique (ce nombre traduit la résolution des données),
- $protocole_{débit}$ est le débit du protocole de communication sans fil,
- $nbSampleByte_{minute}$ est le nombre d'octets de données acquis.

Les protocoles de communications Bluetooth, Zigbee, Lora avec la modulation Lora et Lora avec la modulation GFSK (Gaussian Frequency-Shift Keying) ont respectivement un débit de 1 Mégabits par seconde, 250 kilobits par seconde, 10938 bits par seconde et 50 kilobits par seconde en théorie. Par conséquent, selon ces débits, nous devons transmettre des données toutes les 17 minutes avec le protocole Bluetooth, toutes les 4 minutes avec le protocole Zigbee, toutes les 11,4 secondes pour le protocole Lora avec la modulation Lora et toutes les 52 secondes pour le protocole Lora avec la modulation GFSK lorsque la fréquence d'échantillonnage est $Fs = 10Hz$. Sur la base du scénario établi, pour chaque phase d'interprétation de 20 minutes, nous transmettrons 2 fois en utilisant le protocole Bluetooth, 5 fois en utilisant le protocole Zigbee, 106 fois en utilisant Lora avec la modulation Lora et 24 fois en utilisant Lora avec la modulation GFSK.

Le tableau ci-dessous, dans lequel Lora (lora) et Lora (GFSK) signifient respectivement le protocole de communication Lora avec la modulation lora et celui avec la modulation GFSK, montre les résultats obtenus à partir du modèle d'estimation de l'autonomie. Cette dernière dépend des quatre protocoles de communication avec ($Fs = 10 Hz$ et $I_{mcu} = 11.05 mA$). $A_{scenario}$, exprimé en jours, représente l'autonomie de la plateforme pour le scénario préalablement établi.

À la lumière des résultats présentés dans le tableau 3.3, il convient de noter que le protocole de communication n'influe pas de manière significative sur l'autonomie

3.2 Modèle de consommation

<i>Batterie (mAh)</i>	<i>A_{scenario} (jours)</i>			
	<i>Bluetooth</i>	<i>Zigbee</i>	<i>Lora (lora)</i>	<i>Lora (GFSK)</i>
500	19.4	19.3	18.4	19.2
800	32	32	30.5	31.7
1300	53.2	53.1	50.7	52.6

TABLE 3.3 – Autonomie selon différents protocoles et de chip RF : $F_s=10$ Hz, $I_{mcu}=11.05$ mA

du système d’acquisition pour cette fréquence d’échantillonnage. Cette étude est importante car elle permet de conclure que ces différents protocoles de communication pourraient être utilisés pour des systèmes d’acquisition de signaux physiologiques (système d’e-santé) sans affecter de manière notable l’autonomie.

Cette constatation s’explique par le fait que contrairement aux réseaux de capteurs actuels où le temps d’acquisition est relativement faible ce qui permet une mise en repos du noeud, notre système est en monitoring continu c’est à dire pendant tout l’intervalle d’interprétation, le système d’acquisition reste actif. Plus précisément, même si le courant consommé par le module radio est plus important que celui des modules de traitement et capteur ; en termes d’énergie, les modules de traitement (mcu) et d’acquisition (capteurs) consomment beaucoup plus que le module radio à cause de leur durée respective d’activité. Les modules mcu et capteurs sont actifs pendant tout l’intervalle d’interprétation des émotions alors que le module radio est actif pendant les périodes de transmission puis il est mis au repos. C’est la raison pour laquelle l’autonomie du système n’est pas très affectée par le changement de protocole de communication. Par conséquent, pour diminuer de manière significative la consommation d’énergie, il est nécessaire de réduire l’énergie consommée par les unités de traitement et d’acquisition.

Dans ce cadre, nous avons étudié en premier lieu le courant consommé par l’unité de traitement (MCU). La plateforme d’acquisition que nous avons choisi d’utiliser dispose, pour l’unité de traitement, d’un microcontrôleur du type atmega328p. Il fonctionne à 8 MHz et consomme 11,05 mA. Lorsque nous avons diminué la fréquence de fonctionnement à 2 MHz puis 1 MHz, les courants consommés étaient respectivement de $I_{mcu} = 7,21$ mA et $I_{mcu} = 6,77$ mA. Ces valeurs montrent que le courant varie linéairement par rapport à la fréquence de fonctionnement du microcontrôleur. Le tableau ci-dessous présente les résultats obtenus pour l’estimation de l’autonomie en fonction des quatre protocoles de communication, pour une fréquence de 1 MHz pour le mcu ($I_{mcu} = 6,77$ mA) et une fréquence d’échantillonnage identique aux résultats précédents soit de 10 Hz.

Les résultats présentés dans le tableau 3.4 montrent une amélioration de l’autonomie du système lorsque la fréquence de fonctionnement du microcontrôleur varie de 8 MHz à 1 MHz. En effet, avec une batterie de 500 mAh, l’autonomie passe de 19,4 jours à 25,5 jours pour le module de communication Bluetooth. En utilisant les protocoles Zigbee, Lora (lora), Lora (GFSK), les autonomies augmentent respectivement de 6,1 jours, 5.5

<i>Batterie (mAh)</i>	<i>A_{scenario}</i> (jours)			
	<i>Bluetooth</i>	<i>Zigbee</i>	<i>Lora (lora)</i>	<i>Lora (GFSK)</i>
500	25.5	25.4	23.9	25
800	42.2	42	39.6	41.6
1300	69.9	69.8	65.6	69

TABLE 3.4 – Autonomie en fonction du module radio : fréquence du mcu 1 MHz, $F_s=10\text{Hz}$, $I_{mcu}=6.77\text{ mA}$

jours et de 5,8 jours. Nous avons ensuite poursuivie cette étude en faisant varier la fréquence d'échantillonnage et nous avons considéré deux capacités de batterie.

<i>F_s(Hz)</i>	<i>A_{scenario}</i> (jours)			
	<i>Bluetooth</i>	<i>Zigbee</i>	<i>Lora (lora)</i>	<i>Lora (GFSK)</i>
1	20	20	20	20
5	19.8	19.8	19.3	19.7
10	19.4	19.3	18.5	19.2
20	18.6	18.5	16.9	18.2
50	16.2	16	12.9	15.4
80	13.8	13.6	9.9	12.7
100	12.2	11.9	8.2	11

TABLE 3.5 – Autonomie en fonction de la fréquence d'échantillonnage avec $I_{mcu}=11.05\text{ mA}$, et une batterie de 500 mAh

Les résultats présentés dans les tableaux 3.5 et 3.6 montrent l'influence de la fréquence d'échantillonnage sur l'autonomie du système. L'augmentation de la fréquence d'échantillonnage conduit à la diminution de l'autonomie du système d'acquisition car elle engendre une augmentation du nombre de données à collecter et à transmettre.

La figure 3.4 montre la variation de l'autonomie estimée du système en fonction du protocole de communication et de la fréquence d'échantillonnage. On peut noter, d'une part, que pour une fréquence d'échantillonnage comprise entre 5 Hz et 15 Hz, l'autonomie ne varie pas énormément selon le protocole de communication. Cette constatation est due au fait que, pour ces fréquences, le nombre de communication reste faible donc influe peu sur l'autonomie. D'autre part, les protocoles Zigbee et Bluetooth ont approximativement les mêmes valeurs d'autonomie en fonction de la variation de la fréquence d'échantillonnage. Par conséquent, nous pouvons conclure que Zigbee et Bluetooth sont tous les deux, des protocoles totalement adaptés à ce type de système.

Afin de valider le modèle proposé pour l'estimation de l'autonomie, nous avons mis en place une campagne de mesures adéquate visant à déterminer l'autonomie sur la plateforme BITALINO. Le tableau 3.7 montre l'écart entre les valeurs estimées et mesurées et la précision du modèle proposé. Les différents résultats présentés dans ce tableau ont été calculés ou mesurés pour $I_{mcu} = 11,05\text{ mA}$ et le protocole Bluetooth.

3.2 Modèle de consommation

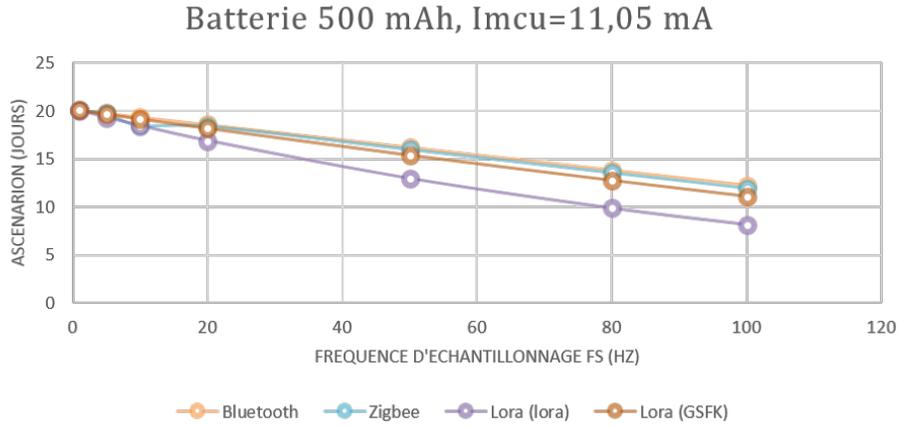


FIGURE 3.4 – Autonomie en fonction de la fréquence d’échantillonnage et du protocole de communication

$F_s(Hz)$	$A_{scenario}$ (jours)			
	<i>Bluetooth</i>	<i>Zigbee</i>	<i>Lora (lora)</i>	<i>Lora (GFSK)</i>
1	53.9	53.9	53.7	53.9
5	53.6	53.6	52.3	53.4
10	53.2	53.1	50.7	52.7
20	52.4	52.2	47.6	51.3
50	49.9	49.4	40	47.5
80	47.5	46.7	33.9	43.9
100	45.9	44.9	30.7	41.6

TABLE 3.6 – Autonomie en fonction de la fréquence d’échantillonnage avec $I_{mcu}=11.05$ mA et une batterie de 1300 mAh

$A_{measured}$ représente l’autonomie mesurée en utilisant la plateforme bitalino doté du protocole de communication Bluetooth, $A_{estimated}$ est l’autonomie estimée à partir de notre modèle proposé avec le protocole Bluetooth, l’erreur est calculée en utilisant la formule d’erreur de pourcentage suivante :

$$erreur = \frac{|A_{estimated} - A_{measured}|}{A_{measured}} * 100 \quad (3.13)$$

Pour mesurer l’autonomie, nous avons exécuté l’application de détection des émotions sur un PC/Matlab couplé et le système d’acquisition était réalisé par la plateforme BITALINO. Ces expérimentations ont été effectuées pour différentes fréquences d’échantillonnage. Ainsi, le niveau d’énergie de la batterie était mesuré après chaque transmission et par programmation des timers, nous avons extrait l’autonomie.

les résultats présentés dans le tableau 3.7 montrent que notre modèle présente une bonne précision et que les autonomies mesurées et estimées varient linéairement

Autonomie pour une batterie de 500 mAh			
Fs (Hz)	$A_{measured}$ (jours)	$A_{estimated}$ (jours)	erreur (%)
1	20,94	20	3.99
10	19,0751	19.4	1,63
100	13,12	12	6.74
Autonomie pour une batterie de 800 mAh			
Fs (Hz)	$A_{measured}$ (jours)	$A_{estimated}$ (jours)	erreur (%)
1	32,2001	32.8	1,87
10	30,5501	32	5
100	25	24.8	0,53
Autonomie pour une batterie de 1300 mAh			
Fs (Hz)	$A_{measured}$ (jours)	$A_{estimated}$ (jours)	erreur (%)
1	55,05	53.9	1.97
10	49,4	53.2	7,75
100	45,75	45.9	0,37

TABLE 3.7 – Comparaison entre autonomie mesurée et calculée

avec la capacité de la batterie. Ce fait confirme l'hypothèse de départ de cette étude qui considère la batterie comme une quantité linéaire de consommation. L'erreur maximale entre l'autonomie estimée et l'autonomie mesurée est de 7,75 %. L'erreur minimale de 0,37 % est obtenue avec une batterie de 1300 mAh et pour une fréquence d'échantillonnage de 100 Hz. Dans le modèle que nous avons proposé, lorsque la fréquence d'échantillonnage augmente, la fonction $(k_1 * Fs + k_2)$ augmente et la différence $Q_{available} - (k_1 * Fs + k_2)$ diminue. Les valeurs estimées sont très proches de celles mesurées. cette comparaison montrent que notre modèle tient compte des paramètres de haut niveau adéquats. On peut noter en particulier que la dépendance sur l'autonomie de la fréquence d'échantillonnage a été correctement modélisée. Par conséquent, le modèle proposé pourrait être utilisé pour estimer efficacement l'autonomie d'objets communicants dédiés à la problématique de la santé et ceci à un niveau d'abstraction élevé. Dans de la suite de ces travaux, afin d'évaluer l'impact de la configuration du système d'acquisition comme la fréquence d'échantillonnage qui agit sur la qualité des signaux, Il sera nécessaire de coupler cette étude avec le taux de reconnaissance des émotions. Il sera ainsi possible de réaliser une exploration d'architecture et de configuration qui tienne compte de contraintes liées, à la fois, à l'autonomie et à la qualité de la détection.

3.3 Étude de l'influence des choix architecturaux sur la qualité de la détection

Pour réaliser une exploration, d'architecture et de ses configurations, pertinente, il est nécessaire de pouvoir évaluer le taux de reconnaissance des émotions en fonction des choix architecturaux. Pour y parvenir, il apparaissait opportun de coupler la modélisation matérielle et les résultats donnés par les algorithmes de détection des émotions.

Ainsi pour différentes fréquences d'échantillonnage, le taux de reconnaissance des émotions a été déterminé à partir de l'algorithme de détection des émotions décrit sous Matlab. Nous avons fait varier la fréquence d'échantillonnage entre 20 Hz, et 1 Hz car la base de donnée que nous avons utilisé pour valider les algorithmes de détection des émotions ne possède que des signaux ayant une fréquence d'échantillonnage de 20hz, Il n'était donc pas possible d'étudier des fréquences supérieures. Les résultats que nous avons alors obtenus sont présentés dans le tableau 3.8. Nous avons constaté une diminution du taux de reconnaissance lorsque la fréquence d'échantillonnage décroît. En effet, toute division par **2** de la fréquence d'échantillonnage conduit approximativement à une diminution de **12%** du taux de détection et de classification d'une activité émotionnelle. En effet, nous avons trouvé respectivement avec les fréquences d'échantillonnage de 20 Hz, 10 Hz et 5 Hz, un taux de reconnaissance de 81.69%,69.01% et 55.83% comme le montre le tableau 3.8.

Fréquence d'échantillonnage (Hz)	Taux de reconnaissance (%)
20	81.69
10	69.01
5	55.83

TABLE 3.8 – Taux de reconnaissance des émotions en fonction de la fréquence d'échantillonnage

Une étude complémentaire pour des fréquences d'échantillonnage au delà de 20 Hz et jusqu'à 1000 Hz est nécessaire afin d'affiner le modèle qui lie le taux de reconnaissance et la fréquence d'échantillonnage. Pour y parvenir, il faut créer une nouvelle base de donnée qui fait l'objet du chapitre 5. Il devient ainsi possible de déterminer la fréquence d'échantillonnage optimale pour la détection des émotions utilisant des signaux physiologiques. Cette fréquence d'échantillonnage optimale est celle qui maximise à la fois le taux de reconnaissance et l'autonomie du système. Dans la section 4.3 du chapitre 5, nous étudierons plus en détail et, sur une nouvelle base de donnée, l'influence de la fréquence d'échantillonnage sur l'autonomie et la performance du système de détection.

3.4 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle approche pour estimer l'autonomie d'un système d'acquisition pour l'e-santé. Nous avons défini, à un haut niveau d'abstraction, un modèle générique et analytique d'estimation de la consommation et de l'autonomie. Le modèle développé est basé sur la décomposition fonctionnelle du système d'acquisition. Il a été nécessaire dans un premier temps de définir les paramètres adéquats de ce modèle. A la différence des modèles existants qui proposent une modélisation de la consommation d'énergie d'un module particulier (module radio, module de traitement, module capteurs), notre modèle considère la consommation d'énergie pour l'architecture complète d'une plateforme d'acquisition. Ainsi, il offre l'avantage aux concepteurs de ces systèmes d'estimer à haut niveau d'abstraction la consommation d'énergie et l'autonomie permettant alors de réaliser une exploration d'architecture et de ses configuration en fonction de contraintes sur l'autonomie. Sur la base de ce modèle, nous avons ainsi pu analyser l'impact du protocole de communication et de la fréquence d'échantillonnage sur l'autonomie. il faut aussi noter que contrairement à d'autres études, le module radio ne représente pas la part la plus importante en terme de consommation car les signaux physiologiques demandent une acquisition continue.

Le modèle développé a été évalué sur une plateforme réelle de développement appelée Bitalino et à partir d'un scénario applicatif défini avec des médecins. Les résultats montrent que le modèle proposé comporte une précision correcte car l'erreur minimale est de 0.37% et l'erreur maximale de 7.75%. En fin de ce chapitre, afin d'étudier l'influence de la fréquence d'échantillonnage sur le taux de reconnaissance des émotions, une étude complémentaire a été présentée et a permis de définir un modèle qui lie ces deux valeurs. Il est donc possible à l'aide de ces modèles de pouvoir réaliser une exploration d'architecture et de configuration et d'évaluer, pour chacune d'entre elles, l'autonomie et le taux de reconnaissance associé.

Chapitre 4

Vers un framework couplant autonomie et performance

Sommaire

4.1 Conception et implémentation d'un framework de simulation et d'exploration	94
4.1.1 Pourquoi un nouveau framework d'exploration et de simulation ?	95
4.1.2 Généralités sur la modélisation SystemC	97
4.1.3 Principes de la modélisation au niveau transactionnel (TLM)	99
4.2 Développement du framework de simulation/exploration	100
4.2.1 Architecture générale du framework	100
4.2.2 Modélisation du module de traitement	101
4.2.3 Modélisation du module radio	103
4.2.4 Modélisation du module sensors	104
4.2.5 Modélisation du module batterie	105
4.3 Résultats et discussions	106
4.3.1 Modélisation du système d'acquisition avec un noeud capteurs	106
4.3.2 Étude de l'influence du microcontrôleur sur la consommation	108
4.3.3 Étude de l'influence de la fréquence d'échantillonnage sur la consommation	111
4.3.4 Étude de l'influence du protocole de communication sur la consommation	113
4.3.5 Modélisation du système d'acquisition avec trois nœuds capteur : étude de l'autonomie	116
4.4 Conclusion	118

4.1 Conception et implémentation d'un framework de simulation et d'exploration

Dans le chapitre précédent, nous avons présenté un nouveau modèle pour l'estimation de la consommation énergétique et de l'autonomie d'un système d'acquisition. De plus, afin de tenir compte de l'impact de certains choix architecturaux sur la qualité de la détection des émotions, nous avons étudié l'influence de la fréquence d'échantillonnage sur le taux de reconnaissance des émotions. Notre objectif est maintenant de disposer d'un environnement de conception et de simulation au niveau système pour les objets communicants dédiés à des applications d'e-santé. Cet environnement de conception doit permettre de réaliser une exploration d'architecture, modulaire, et capable de tenir compte de ses configurations afin de ne retenir que les solutions architecturales capables de satisfaire des contraintes comme l'autonomie et la qualité de la détection. Aussi, cet environnement se compose :

- d'un module permettant de spécifier toutes les contraintes c'est à dire l'autonomie du système d'acquisition et la performance du système qui représente le taux de reconnaissance des émotions.
- d'un module permettant de composer l'architecture. Ce module s'appuie sur une librairie de modules de traitement, d'acquisition et de radio pour des protocoles de communication différents.
- d'un module relatif à l'application contenant d'une part la définition du scénario applicatif et permettant de voir le fonctionnement global de notre système.

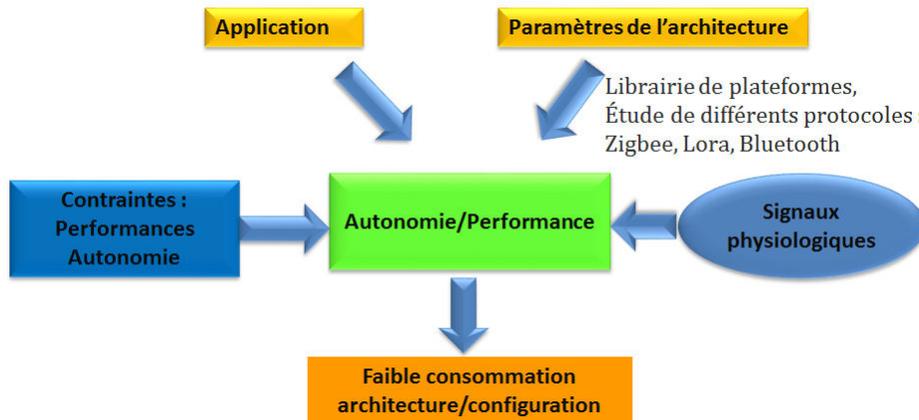


FIGURE 4.1 – Environnement pour l'exploration d'architecture

L'originalité de cet environnement d'exploration/simulation est de pouvoir simuler le comportement du système d'acquisition en lien avec l'application et d'estimer à la fois l'autonomie d'une solution architecturale et la qualité de la détection. Pour y parvenir, nous avons choisi de modéliser cet environnement composé de nœuds de capteurs en SystemC-TLM. Nous avons choisi cet environnement de simulation car il est capable de modéliser à haut niveau d'abstraction les communications à la fois interne

4.1 Conception et implémentation d'un framework de simulation et d'exploration

aux capteurs/système d'acquisition mais aussi externes. Le framework ainsi conçu offre aux concepteurs de systèmes pour la santé un cadre de conception et de simulation en leur donnant la possibilité d'évaluer les performances de leurs systèmes avant la conception matérielle réelle. Il permet ainsi d'étudier l'impact de certains paramètres, de configuration et architecturaux pour ces types de systèmes d'objets communicants. Ainsi, ce framework propose une description matérielle et logicielle de ces types de systèmes. Dans ce chapitre, nous étudierons deux architectures différentes pour la collecte des signaux physiologiques. La première est composée d'un seul noeud doté de trois capteurs alors que la deuxième se compose de trois noeuds doté d'un seul capteur. L'objectif visé par cette étude est de développer une architecture non-invasive où chaque noeud de capteur peut être encapsulé dans un boîtier afin de faciliter l'acquisition des données.

Ce chapitre est organisé comme suit. Tout d'abord, dans la section 4.1.1, nous justifions l'intérêt de concevoir un nouvel environnement d'exploration/simulation d'architecture et nous introduisons brièvement SystemC et la modélisation au niveau transactionnel (TLM) de plateformes d'e-santé. Ensuite, dans la section 4.2, la modélisation des différents modules du système d'acquisition et leurs interactions sont présentées. Enfin la section 4.3 montre les différents résultats de simulation obtenus.

4.1.1 Pourquoi un nouveau framework d'exploration et de simulation ?

La modélisation d'un système consiste entre autre à l'association de modèles décrivant ou représentant approximativement les comportements ou les caractéristiques de ce système. Elle sert, d'une part, à estimer le fonctionnement global de ce système avant sa conception réelle et d'autre part, elle permet une exploration d'architecture à haut niveau d'abstraction. Le fonctionnement d'un système complet est influencé aussi bien par son architecture matérielle, que par son environnement extérieur c'est à dire par le contexte applicatif. Il est donc nécessaire de modéliser le système en tenant compte de ces deux types de paramètres. Le choix d'un simulateur pour réseaux de capteurs sans fil est donc délicat du fait de cette influence de l'environnement extérieur donc de l'application visée. Sur les nombreux simulateurs pour réseaux de capteurs sans fil, les exigences principales sont entre autres :

- La fidélité, pour rendre le modèle simulé le plus proche possible du comportement réel du système.
- La capacité de simuler un grand nombre de noeuds de capteurs sans augmenter considérablement les temps de simulation et la quantité de mémoire occupée.
- La modularité du simulateur afin de donner la possibilité d'ajouter ou d'enlever certains modules.
- La modélisation de la consommation d'énergie, qui permet d'estimer l'autonomie des noeuds et d'évaluer des stratégies visant à rendre ces noeuds plus économes en énergie.

4.1 Conception et implémentation d'un framework de simulation et d'exploration

Les résultats obtenus avec les simulateurs classiques de réseaux de capteurs sont très approximatifs du fait d'une part, qu'ils ne tiennent pas compte de manière précise de l'architecture interne des nœuds de capteurs. D'autres part, ni l'application cible, ni le système réel ne sont simulés mais seulement un modèle approximatif de leur comportement est pris en considération. En conséquence, ces outils ne fournissent aucune indication quant à l'implémentation du système ou au développement de code. Ainsi, l'utilisation de ces simulateurs n'est pas préconisée lorsqu'il s'agit de déboguer ou d'optimiser une application ou de réaliser des choix quant à son implantation. Leur intérêt réside à un niveau plus élevé, durant le processus de conception de l'application [CHELIUS and FRABOULET, 2007].

Dans l'état de l'art, les simulateurs de réseaux de capteurs sans fil sont divisés en 3 catégories : les simulateurs de réseaux, les émulateurs et les simulateurs de nœuds. Les simulateurs de réseaux s'appuient en général sur des techniques qui calculent l'interaction entre les différents blocs qui composent le réseau, en captant réellement des échanges d'information ou virtuellement à partir d'un réseau de production réel et en rejouant les observations expérimentales [Pan and Jain, 2008]. Les émulateurs de nœuds, y compris les émulateurs de systèmes d'exploitations et les simulateurs de jeux d'instructions (ISS) de l'unité de traitement (microcontrôleur, DSP, CPU, etc) modélisent l'exécution du logiciel embarqué, tandis que les simulateurs de nœuds sont généralement développés dans *System Level Design Language (SLDL)* avec des modèles de description comportementale des nœuds de capteurs [Du et al., 2011]. Le tableau 4.1 liste quelques simulateurs de réseau de capteurs sans fil existants.

<i>Catégories des simulateurs</i>	<i>simulateur</i>	<i>Langages</i>
Simulateurs de réseaux	NS-3 [Henderson et al., 2008]	C++, OTcl
	OMNeT++ [Du et al., 2011]	C++
	NeTopo [Shu et al., 2011]	Java
	PowWow [Berder and Sentieys, 2010]	nesC
Émulateurs de nœuds	TOSSIM [Levis et al., 2003]	C
	ATEMU [Polley et al., 2004]	
	Avrora [Titzer et al., 2005]	Java
	Asymphony [Riliskis and Osipov, 2015]	C++
Simulateurs de nœuds	WISENES [Kuorilehto et al., 2008]	langage SDL
	SNOPS [Damm et al., 2010]	SystemC-TLM
	ATLeS-SN [Hiner et al., 2010]	
	SCNSL [Fummi et al., 2008]	SystemC, C++
	IDEA1 [Du et al., 2011]	

TABLE 4.1 – Quelques simulateurs de réseaux de capteurs sans fils

Contrairement aux simulateurs présentés dans le tableau 4.1, notre environnement d'exploration/simulation vise à fournir, en premier lieu, un simulateur, au niveau système, dédié en priorité au dispositif d'acquisition pour la détection des émotions

4.1 Conception et implémentation d'un framework de simulation et d'exploration

(applications de soins de santé). Ce framework couple Matlab et SystemC pour fournir un cadre de conception et d'exploration d'architecture en tenant compte de toutes les contraintes applicatives et utilisateurs. Ces applications ne peuvent pas être simulées avec les simulateurs actuels de réseau de capteurs en raison des différences suivantes :

- Contrairement au système de détection des émotions, dans les réseaux de capteurs traditionnels, le temps d'acquisition est faible et le nombre de sessions d'acquisition par jour est élevé.
- Les réseaux de capteurs actuels ne prennent pas en compte les contraintes utilisateurs liées aux dimensions des capteurs, leur poids et à la qualité des signaux physiologiques. Le plus souvent, ils ne couplent pas les contraintes applicatives, d'architecture et utilisateur.
- Le nombre de nœuds du système d'acquisition pour la détection des émotions est faible alors que ces simulateurs de réseaux de capteurs ont été conçus pour simuler un nombre important de nœuds et d'évaluer la performance en fonction, entre autre, de la latence des communications.
- La plupart des simulateurs actuels ne possèdent pas de modèles d'architectures cibles contrairement au nôtre qui, contiendra les différents types d'architectures de collecte que nous utilisons pour l'application visée. Les systèmes de détection des émotions que nous considérons utilisent des modules fonctionnels existants sur le marché. Notre framework s'appuiera sur un modèle fin de l'architecture du système d'acquisition (unités de traitement, radio, et capteurs) afin de garantir la précision de notre simulateur. Cette garantie nous permettra d'évaluer de manière pertinente à la fois l'autonomie et le taux de reconnaissance des émotions.

Pour toutes ces raisons, nous avons choisi de concevoir notre propre environnement de simulation/exploration. Cet environnement qui a été développé en SystemC-TLM s'adresse à des architectures de collectes de données de signaux biomédicaux composés de nœuds de capteurs en faible nombre. Le comportement des différents modules est représenté par des machines d'états et la communication entre certains de ces modules est modélisée par des transactions TLM. La synchronisation des différents modules est assurée par des événements SystemC (*sc_event*). Dans ce qui suit, nous introduisons les deux outils SystemC et TLM qui nous ont permis de modéliser ce framework.

4.1.2 Généralités sur la modélisation SystemC

SystemC est une librairie de classes C++ pour la conception et la simulation de systèmes. Il offre l'avantage de permettre une modélisation matérielle et logicielle, ce qui permet de modéliser un système complexe dans son ensemble à un très haut niveau d'abstraction. Ainsi il offre la possibilité de réaliser une co-conception matérielle/logicielle qui permet de gagner du temps lors la conception et de diminuer les coûts de conception. En plus de cette co-conception matérielle/logicielle, il est possible de modéliser un système à plusieurs niveaux d'abstraction. Une architecture peut ainsi être modélisée à différents niveaux d'abstraction en fonction de la précision attendue et

4.1 Conception et implémentation d'un framework de simulation et d'exploration

de l'étude de phénomènes observée. Comme SystemC utilise les classes C++, il garantit la possibilité de réutiliser des modèles avec les notions de surcharges, d'héritages ... ce qui augmente la vitesse et diminue la complexité de conception de l'architecture complète du système. En terme d'efficacité, le modèle SystemC peut atteindre une vitesse dix fois plus rapide qu'un modèle VHDL au même niveau d'abstraction [Aboulhamid, 2001] et une version multi-threadée est en phase d'être disponible ce qui accélèra encore les temps de simulation grâce à la répartition de la charge sur plusieurs cœurs du processeur.

La modélisation en SystemC regroupe différents types de constructeurs (les modules, les processus, les ports et les signaux) permettant de modéliser n'importe quel système matériel/logiciel, synchrone/asynchrone ...

- Les modules (*SC_MODULE*) servent à modéliser les sous-systèmes d'une architecture. Une architecture comporte généralement plusieurs composants qui interagissent entre eux. Aussi en SystemC, ces composants matériels ou logiciels peuvent être modélisé comme une classe *SCModule*.
- Ils existent en SystemC 3 types de processus qui sont appelés lorsqu'un événement auquel ils sont sensibles se produit. Chaque processus peut être rattaché à un module. Les processus de types *methods* sont ceux dont l'exécution ne peut être suspendue. Une fois lancée, l'ensemble du processus est exécuté. A la fin de leurs exécutions, ce type de processus rend le contrôle au noyau SystemC. Pour les processus de types *threads*, ces derniers peuvent être suspendus en appelant la fonction *wait()*. L'exécution pourra reprendre lorsque le temps d'attente est dépassé ou lorsqu'un événement se produit. Les processus de types *Cthread* quant à eux sont des cas particuliers de *threads*. Ils sont uniquement sensibles au signal d'horloge. Ainsi, les actions des processus *methods* sont utilisés pour des actions devant être traitées jusqu'à la fin avant de passer la main contrairement au *threads*. Enfin, les *cthread*s sont utiles pour la synthèse du matériel.
- Les entités *ports* et les *signaux* correspondent aux interfaces des modules, leur permettant de communiquer entre eux. Les ports sont les entrées/sorties des modules. Ils sont interconnectés à l'aide de signaux qui peuvent être, comme les ports, de n'importe quel type de données SystemC.

L'ordonnanceur de la librairie SystemC permet de gérer tous les processus lancés. Un processus peut être dans l'un des trois états suivants :

1. Élu : le processus s'exécute et utilise le processeur à l'instant courant.
2. Éligible : le processus est prêt à être élu et attend que l'ordonnanceur le choisisse.
3. Bloqué : le processus attend une donnée (entrées-sorties, événements, ...)

Pour sélectionner les processus devant être exécutés, l'ordonnanceur SystemC passe par quatre principales étapes. La première étape correspond à *l'initialisation*, étape au cours de laquelle tous les processus sont exécutés. Ensuite, durant l'étape *d'évaluation*, les processus s'exécutent et se suspendent eux-mêmes. L'étape *update* est appelée quand il n'y a plus de processus éligibles. Une mise à jour des différents signaux est faite durant

4.1 Conception et implémentation d'un framework de simulation et d'exploration

cette étape. Enfin, dans la quatrième étape *time elapse*, le temps simulé est mis à jour c'est à dire avancé.

La simulation SystemC est paramétrée et contrôlée dans la fonction principale *sc_main* comme le *main* dans un programme C++. Les modules sont tout d'abord initialisés et les ports des modules sont interconnectés par des signaux. La simulation débute par l'appel de la fonction *sc_start*. En cours de simulation, le noyau SystemC met à jour les valeurs des signaux et à chaque cycle de simulation, les processus actifs sont invoqués et le cycle de simulation avance. A noter que dès qu'un événement de la liste de sensibilité d'un processus se produit, ce dernier devient actif. Comme tous les processus actifs sont invoqués au même instant logique de simulation, cela crée une illusion de concurrence. Enfin, lorsque le temps du simulateur atteint le point où l'application se déclenche (*sc_stop*), la simulation s'arrête et tous les objets sont supprimés. Le lecteur intéressé par le noyau de simulation SystemC peut trouver plus de détails dans les travaux de Mueller [Mueller et al., 2001].

4.1.3 Principes de la modélisation au niveau transactionnel (TLM)

TLM (Transaction Level Modeling) qui signifie modélisation au niveau transactionnel est un protocole largement utilisé par les concepteurs de systèmes embarqués afin de simplifier et d'accélérer la simulation d'architecture. En effet, il offre la possibilité aux concepteurs de se concentrer uniquement sur les fonctionnalités les plus intéressantes pour l'application visée. Cette modélisation intégrant à la fois le développement matériel et logiciel se concentre sur les communications entre les différents blocs matériels/logiciels tout en garantissant une validation du système à haut niveau d'abstraction. La modélisation TLM s'appuie sur un composant de calcul et un autre de communication.

La modélisation en TLM se base donc essentiellement sur la communication entre des modules. C'est justement cet aspect de la modélisation qui nous intéresse, par exemple, pour certains modules comme le Timer. Cette communication est assurée par des modules *initiateurs* et *targets* (cibles) interconnectés par des sockets. Une transaction pouvant être définie comme une structure de données passée par appel des fonctions d'interfaces (read, write ...) entre un module qui initie de nouvelles transactions et celui qui y répond. Un module peut à la fois être un initiateur et une cible "target" comme les modules d'interconnexions qui ne font que transmettre des transactions sans modifier leurs contenus. La modélisation TLM offre l'interopérabilité entre les modèles de niveau transaction et ceci est assurée en grande partie grâce à l'utilisation des transactions de types *tlm_generic_payload*.

Le module qui doit réaliser une transaction utilise un de ses sockets initiateurs pour démarrer la transaction en appelant la méthode *b_transport* de son interface de transport. La transaction est reçue par le socket cible qui est connecté au socket initiateur. A la réception de la transaction, une fonction du module contenant le socket cible est appelée. Cette fonction définit le ou les actions qui doivent être réalisées suite

à la transaction entrante. Un objet de type `payload` est associé aux transactions. Il possède un ensemble standard d'attributs liés à un bus, comme principalement, une adresse, un pointeur de données, une longueur de données, une commande TLM et un état de réponse. Les deux commandes supportées sont `read` et `write`. L'attribut `address` indique l'adresse la plus basse à laquelle les données doivent être lues ou écrites. Le pointeur de données pointe vers une mémoire tampon dans le socket initiateur. Si la commande TLM est `read`, les données lues seront écrites dans cette mémoire tampon. Si la commande est `write`, les données à écrire seront lues dans cette même mémoire tampon.

4.2 Développement du framework de simulation/exploration

Afin de réaliser à un haut niveau la simulation et l'exploration d'architecture d'un système d'acquisition, ce framework a été développé en SystemC-TLM et C++. Les nœuds de capteurs de signaux physiologiques sont modélisés dans SystemC et les interconnexions entre les nœuds sont représentés par des transactions TLM et/ou en C++. Ce framework s'appuie sur une librairie composée d'unités de traitement et de modules radio qui sont associés à des protocoles de communication différents. Dans le chapitre précédent, les résultats ont montré qu'une part importante de la consommation était due à l'unité de traitement. Aussi plusieurs types de microcontrôleurs ont été modélisés, mais aussi différents types de protocoles de communication et de capteurs de signaux physiologiques. Parmi les microcontrôleurs, nous avons choisi, entre autre, de modéliser ceux compatibles avec le monde de l'Arduino car ils sont les plus utilisés actuellement par les concepteurs de plateforme de développement. Comme nous l'avons dit dans les sections précédentes, les concepteurs de systèmes d'objets communicants pour la santé sont à la fin de la chaîne de conception : ils utilisent assez souvent des modules Arduino ou équivalent afin d'explorer très rapidement différentes architectures pour optimiser au mieux la performance du système final. Ces modules sont très simples d'utilisation et garantissent des consommations d'énergies raisonnables.

4.2.1 Architecture générale du framework

L'architecture d'un système d'acquisition est composé d'une unité de traitement (microcontrôleur), de capteurs, d'un module émetteur-récepteur Radio Fréquence (RF), de timers et d'une batterie. Dans notre framework, comme le montre la figure 4.2, chaque composant est modélisé indépendamment des autres comme étant un module SystemC. Il communique avec les autres modules par le biais de canaux spécifiques et/ou via des transactions TLM. La synchronisation des communications est garantie par l'utilisation des événements SystemC (`SC_event`).

Le noyau de la librairie SystemC se charge de l'ordonnancement des processus, à chaque cycle de simulation, les processus actifs sont appelés par le noyau et ces processus

4.2 Développement du framework de simulation/exploration

se synchronisent entre eux grâce aux événements SystemC. Ceci permet de garantir une synchronisation des différents composants du système afin de garantir la cohérence de la simulation.

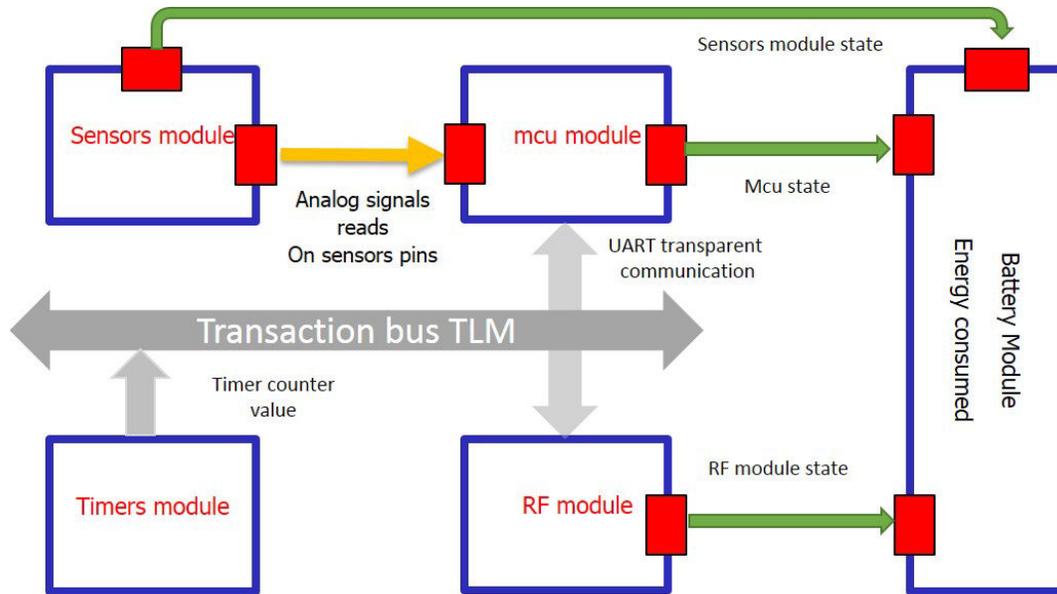


FIGURE 4.2 – Architecture d'un noeud capteur

4.2.2 Modélisation du module de traitement

Nous avons élaboré un modèle générique pour le fonctionnement de l'unité de traitement (le microcontrôleur noté mcu), basé sur une représentation par machine d'états finis (FSM) comme l'illustre la figure 4.3. Cette FSM du mcu est contrôlée par les interruptions provenant du module radio et de l'application (conditions sur le temps d'acquisition, la longueur des trames ...). L'état initial du microcontrôleur est l'état IDLE. Il quitte l'état IDLE après réception d'une interruption de l'application qui l'amène à l'état SENSING afin d'acquérir les différents signaux physiologiques. Dans l'état SENSING, le mcu exécute la génération de signaux modélisée dans le module capteurs qui sera décrite dans la section 4.2.4. Le mcu lit ensuite les données générées, les convertit avec le convertisseur ADC (Analog Digital Converter) et prépare les paquets de transmission avant de les sauvegarder dans un buffer. Après avoir sauvegardé les données, il peut retourner à l'état IDLE ou aller à l'état TX selon les conditions exprimées sur la taille du paquet à transmettre et si le mcu ne reçoit aucune autre interruption. Dans l'état TX, le microcontrôleur transmet les données acquises au module radio via le bus UART (Universal Asynchronous Receiver Transmitter). Le module radio sera en charge d'envoyer ces données en tenant compte du protocole de communication, c'est à dire suivant les spécifications sur le débit, les trames à transmettre, etc. Le module de traitement (mcu) peut alors quitter l'état TX quand

4.2 Développement du framework de simulation/exploration

les paquets sont bien transmis. Ceci implique la réception d'une interruption du module radio pour informer le mcu que le paquet transmis a bien été reçu. Après une transmission, le mcu reste à l'état TX et attend le *time on air* (temps entre deux transmissions consécutives) s'il est nécessaire de transmettre un autre paquet. Dans le cas contraire, il retourne à l'état IDLE ou SENSING suivant l'état de la file de donnée (vide ou pas), ou l'état du sensing (terminé ou pas). Dans l'état IDLE, quand il reçoit une interruption du module radio, c'est à dire dans le cas d'une réception de donnée, le mcu passe dans l'état RX pour lire les données sauvegardées dans le buffer. Enfin, dans l'état RX, il peut soit aller dans les états IDLE, SENSING ou TX selon les différentes conditions. Lorsqu'il y a apparition d'une interruption du mcu, l'automate quitte l'état dans lequel il se trouve.

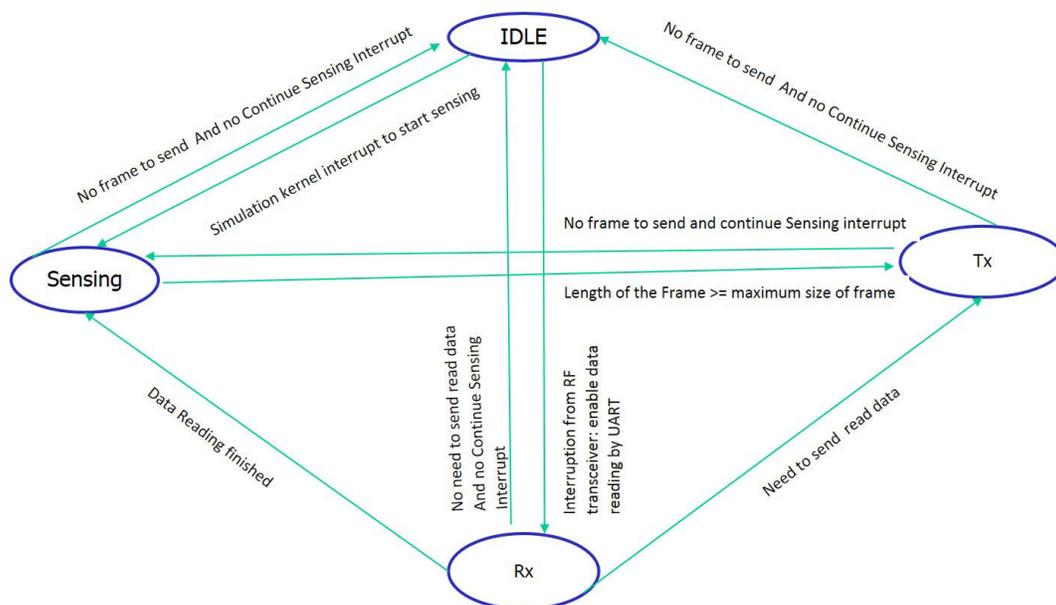


FIGURE 4.3 – Machine d'états finis associée au microcontrôleur

Il faut aussi noter que ce module de traitement communique avec les modules timer et radio. La communication entre le mcu et le module radio est établie et assurée par des transactions TLM. La communication UART a aussi été modélisée par des transactions TLM dans le but d'accélérer le temps de simulation.

Dans le chapitre précédent, les résultats ont montré que l'unité de traitement est celle qui consomme le plus en termes d'énergie, nous avons donc étudié différents microcontrôleurs. La bibliothèque de ce module contient les modèles de sept microcontrôleurs à savoir :

- quatre microcontrôleurs de la famille Atmel : ATMEGA328P, ATMEGA128, ATTINY85, ATTINY841 ;
- deux microcontrôleurs PICS : le PIC16F873A et le PIC18F252 ;
- et enfin le mcu STM32L476.

4.2 Développement du framework de simulation/exploration

Le tableau E.1 donne les courants consommés par les différents microcontrôleurs considérés dans notre étude. Ces courants extraits des datasheets des mcus sont donnés suivant l'état dans lequel se trouve les microcontrôleurs.

MCUs	Attiny85	Attiny841	Atmega128	Atmega328p	STM32L476	PIC16F873A	PIC18F45Q10
Mémoire (Octets)	512	512	4096	2048	98304	192	192
Sleep mode current (mA)	0,01	0,004	0,025	0,84	0,0014	0,02	0,0002
PinInput current (mA)	0,001	0,001	0,00005	0,1	0,005	0,001	0,001
Active mode current (mA)	3	2,4	9,5	11,05	0,312	1,6	11,05

TABLE 4.2 – Taille de la mémoire et courants consommés par les MCUs à 8 MHz

4.2.3 Modélisation du module radio

Comme pour l'unité de traitement, ce module est modélisé par une machine à états. Lorsque le microcontrôleur est à l'état TX, un module radio est requis pour transmettre les données acquises. Nous avons modélisé en premier lieu le module bluetooth HC-05. Ce module radio est celui utilisé sur la plateforme de développement Bitlino que nous avons utilisée dans le chapitre 3 pour valider le modèle d'estimation de la consommation que nous avons développé. Ce chip RF Bluetooth se trouve être aussi l'un des chips radio les plus utilisés dans les applications utilisant les modules arduinos, mbeds ... en raison de sa simplicité de déploiement, sa taille, sa consommation ...

Nous proposons un diagramme (voir figure 4.4) composé de machines d'états permettant de décrire le fonctionnement de ce dernier. Le diagramme ainsi proposé permet de modéliser d'autres types de modules radio en raison de sa généralité puisqu'il considère l'ensemble des étapes requises dans de tels modules. A gauche, la figure montre le protocole de lecture et de transmission des données. A droite, le module radio (en premier lieu le HC-05 est représenté en SystemC selon ses différentes consommations en courant et le temps de chaque état.

Le module radio a été modélisé en SystemC avec deux modes de fonctionnement (le mode *COMMAND* et le mode *TRANSMISSION*). Dans le mode *command*, le microcontrôleur configure la transmission radio donc le module radio (lecture et mise à jour du nom du transceiver, choix du data rate, ...) . Dans le mode *transmission*, le microcontrôleur exécute l'acquisition et le paquetage des données ainsi que l'envoi des données au module RF. Les paquets construits sont transmis au module RF, par la voie série (UART) en fonction du choix du microcontrôleur et de la configuration du

4.2 Développement du framework de simulation/exploration

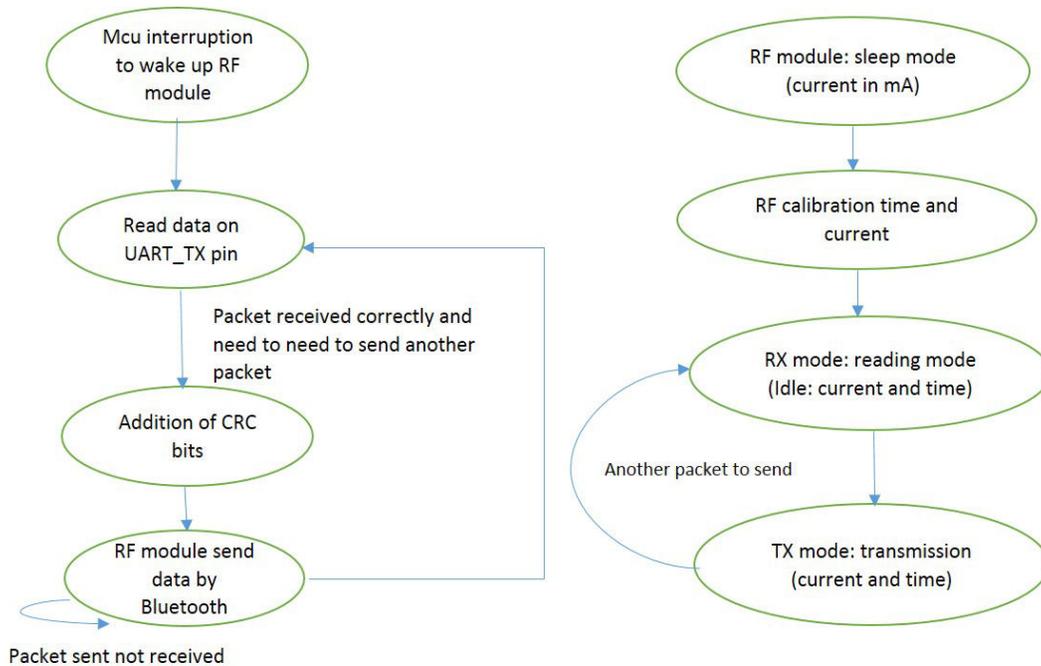


FIGURE 4.4 – Machine d'états du module radio

module RF. A leur réception, le module RF transmet les paquets selon le protocole de communication sans fil choisi et configuré.

Dans un souci de donner plus de flexibilité pour l'utilisateur et d'étudier la consommation, deux modes de transmissions des données ont été configurés dans notre framework. Le premier consiste à transmettre de manière continue après chaque acquisition ne permettant pas une mise en repos de la radio. Le second mode que nous appelons *transmission à la fin de l'acquisition* consiste à sauvegarder les données dans la RAM jusqu'à ce que le temps d'acquisition soit fini ou que la mémoire de stockage des données (mémoire RAM du mcu) soit pleine. Ensuite, les données ainsi groupées seront transmises. Ce deuxième mode est modélisé de telle sorte à garantir la non perte des données grâce à la synchronisation de l'acquisition et des traitements de données. Ces 2 modes de transmission améliorent l'exploration d'architecture en permettant d'étudier l'influence du module de transmission sur la consommation dans chacun de ces deux cas de figure.

4.2.4 Modélisation du module sensors

Le module sensors effectue l'opération d'acquisition de données. Dans le contexte de la détection des émotions ou d'application d'e-santé, il est nécessaire d'utiliser des capteurs physiologiques. De ce fait, nous avons modélisé trois types de capteurs physiologiques, à savoir *le Rythme cardiaque (ECG)*, *l'activité électromyographique (EMG)*, *la réponse électrodermale (GSR)*. Afin de valider en simulation le

4.2 Développement du framework de simulation/exploration

fonctionnement réel de l'application, les modules capteurs sont dotés de générateur de signaux physiologiques. Deux modes de génération de signaux ont été développés dans ce module `sensors`. Pour le premier mode, le module `sensors` répertorie des données synthétiques générées et pré-enregistrées à partir de capteurs réels. En utilisant les couches physiques et de transport réelles, cette fonctionnalité permet aux développeurs de tester la réception des paquets et de modifier les valeurs des données en continu. Pour le deuxième mode, ce module lit en temps réel les données acquises sur un sujet sur lequel les capteurs sont attachés. Ce second modèle a nécessité de connecter SystemC et Octave afin de favoriser cette lecture de données en temps réel. Nous avons utilisé Octave à la place de Matlab dans cette phase d'interconnexion car d'une part nos algorithmes de détection des émotions fonctionnent à la fois sur Matlab et sur Octave. D'autre part, ayant développé notre Framework sur une machine linux, nous avons préféré utilisé Octave qui est un logiciel de calcul numérique comparable à Matlab.

Le module `sensors` pour une plateforme d'acquisition doit être flexible et adaptable car il est possible d'associer un nombre variable de capteurs. Il en est de même pour la taille des données recueillies. Le nombre et la taille des données représentant les signaux physiologiques à acquérir peut varier et l'acquisition peut se faire avec différentes fréquences d'échantillonnage.

Contrairement aux réseaux de capteurs actuels où le temps d'acquisition dans une session est assez faible, un système de détection des émotions est en monitoring continu pendant un intervalle de temps relativement élevé (environ 20 minutes par session d'acquisition). De plus, le nombre de sessions d'acquisition par jour est faible (environ 1-2 fois par jour) dans un système de détection des émotions comparativement aux réseaux de capteurs sans fil actuels où les sessions d'acquisition quotidiennes se font à intervalle de temps régulier et relativement court (environ 15 minutes entre deux sessions : monitoring de la température, de l'humidité ...). Deux sessions d'acquisition seraient suffisantes [Tayari et al., 2012] pour des applications comme la lutte contre la dépression où l'objectif serait de voir si les sujets ont des émotions négatives consécutives pendant un certain nombre de jours. Ainsi, un scénario d'acquisition doit être proposé en fonction des contraintes d'application pour une meilleure estimation de la consommation d'énergie et pour faire une exploration d'architecture basée sur la consommation d'énergie et la qualité du service.

4.2.5 Modélisation du module batterie

Afin de modéliser la charge de la batterie et de déterminer son autonomie, il est nécessaire de disposer des courants unitaires consommés par les différents modules mais aussi des temps d'activation de ces modules. Ces temps d'activation sont extraits à partir de timers. Les timers sont modélisés dans SystemC-TLM et communiquent avec les autres modules par des transactions TLM. Un timer permet de calculer la durée de chaque état de transition de chaque composant et d'enregistrer le temps d'activation de chaque composant matériel et selon son mode de fonctionnement. En fonction des caractéristiques techniques des modules RF, des microcontrôleurs et

des capteurs modélisés dans le Framework, des courants consommés par ces derniers ont été associés aux différents états de fonctionnement du système. Les états de transition sont automatiquement mis à jour et envoyés au module de batterie pour calculer la consommation d'énergie. L'énergie consommée à chaque état est modélisée comme l'illustre l'équation 4.1 où $E_i(mAs)$, $I_i(mA)$ sont respectivement l'énergie et le courant consommé à l'état i ; $T_i(S)$ est le temps passé dans l'état i . V est la tension d'alimentation et n est le nombre de transition.

$$E_i = \sum_{i=1}^n V * I_i * t_i \quad (4.1)$$

A partir de l'équation 4.1, on peut facilement proposer la formule permettant de calculer la consommation énergétique totale qui est donnée par l'équation 4.2. Cette énergie consommée est exprimée en mAs. Dans cette formule, m représente le nombre total d'états et E_i l'énergie totale consommée dans un état de transition.

$$E_{total} = \sum_{i=1}^m E_i \quad (4.2)$$

Par ailleurs, le modèle analytique d'estimation de la consommation énergétique à un haut niveau d'abstraction proposé dans la section 3.2.2 du chapitre précédent a aussi été implémenté dans le framework afin de faire une étude comparative entre les deux modèles et aussi de les coupler en vue de l'amélioration de l'estimation énergétique.

A partir de la quantité de courant consommée (calculé par le framework) et du modèle défini dans le chapitre précédent pour estimer l'autonomie, on peut déterminer l'autonomie du système modélisé.

4.3 Résultats et discussions

Afin de faire de l'exploration d'architectures, différentes unités de traitement, de capteurs et de modules radio ont été modélisés dans la librairie du framework. Deux types de système d'acquisition ont été modélisés, le premier comprend un seul noeud d'acquisition comportant les trois capteurs physiologiques. Dans le second, trois noeuds capteurs seront utilisés permettant un meilleur placement sur le corps.

4.3.1 Modélisation du système d'acquisition avec un noeud capteurs

Pour simuler une architecture d'acquisition, il est nécessaire de fixer entre autre la fréquence d'échantillonnage. Dans le chapitre précédent, nous avons étudié l'impact de la fréquence d'échantillonnage sur le taux de reconnaissance. Dans le cadre de cette nouvelle étude, nous avons fait des acquisitions de signaux à différentes fréquences afin de voir lesquelles nous permettraient d'avoir des formes de signaux assez proches de la théorie. Cette étude s'est simplement basée sur l'analyse des courbes obtenues,

4.3 Résultats et discussions

en les comparant avec des signaux théoriques. Par exemple, pour l'ECG, le signal attendu doit contenir une onde QRS (voir figure 4.5). Nous avons donc noté, après l'analyse des signaux sur la plage de fréquences de $[1, 1000 \text{ Hz}]$, que la fréquence de 100 Hz permettait de garantir une bonne acquisition et une autonomie acceptable. Les fréquences supérieures quant à elle permettent de se rapprocher des critères de nyquist et de mieux acquérir les données mais l'avantage apporté par la montée en fréquence est surtout importante lorsqu'une analyse dans le domaine fréquentiel est requise, ce qui n'est pas notre cas. De plus, dans notre étude, nous éviterons de monter très haut en fréquence afin de réduire la consommation. En effet, comme nous l'avons montré dans le chapitre précédent, plus la fréquence d'échantillonnage est élevée, plus le système consomme et plus faible sera l'autonomie. La fréquence de 100 Hz permet donc de faire un compromis entre la pertinence des signaux acquis et l'autonomie du système. Comme perspective sur ce travail, nous prévoyons de concevoir des bases de données pour plusieurs fréquences d'échantillonnage afin de mieux étudier l'influence de cette dernière sur le taux de reconnaissance. Les différents résultats que nous pouvons extraire du framework permettent de faire une exploration au niveau de la fréquence d'échantillonnage, et ceci en fonction de la composition de l'architecture, la taille de la RAM, le type de microcontrôleur, le nombre de capteurs, le type de module RF utilisé. Dans la suite de ce chapitre, nous présentons quelques résultats extraits de cette exploration. Le choix d'une architecture se fera sur le plan technique suivant son autonomie, le taux de reconnaissance, sa dimension.

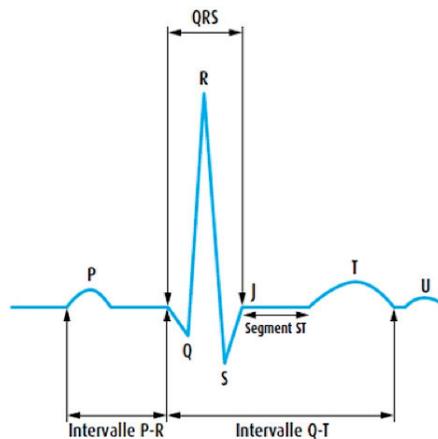


FIGURE 4.5 – Représentation d'un signal ECG

Dans le Framework, deux modes de transmissions sont possibles à savoir le mode CONTINUOUS et END_ACQ. Le premier consiste à acquérir et transmettre en continu. Dans le second mode, les données acquises sont transmises dès que la mémoire est pleine ou si le temps d'acquisition est fini. On note que les simulations des acquisitions qui suivent ont toutes été faites pour une durée d'acquisition de 20 minutes, temps nécessaire pour déterminer une tendance émotionnelle comme expliqué dans le chapitre 5. Sur les graphiques présentés dans la suite de ce document, il est à noter

que ceux montrant la consommation représentent la quantité de courant consommée par l'architecture complète de collecte des signaux physiologiques (capteurs, module RF, microcontrôleur). Ces différentes consommations sont exprimées en mAs. Tous les microcontrôleurs fonctionnent à la même fréquence (8 MHz). Ces simulations ont donc été réalisées pour un intervalle de temps de 20 minutes d'acquisition et ceci pour différents types de microcontrôleur, de protocole de communication et un nombre variable de capteur. L'objectif étant par la suite de sélectionner l'architecture du système d'acquisition qui présente le meilleur compromis.

4.3.2 Étude de l'influence du microcontrôleur sur la consommation

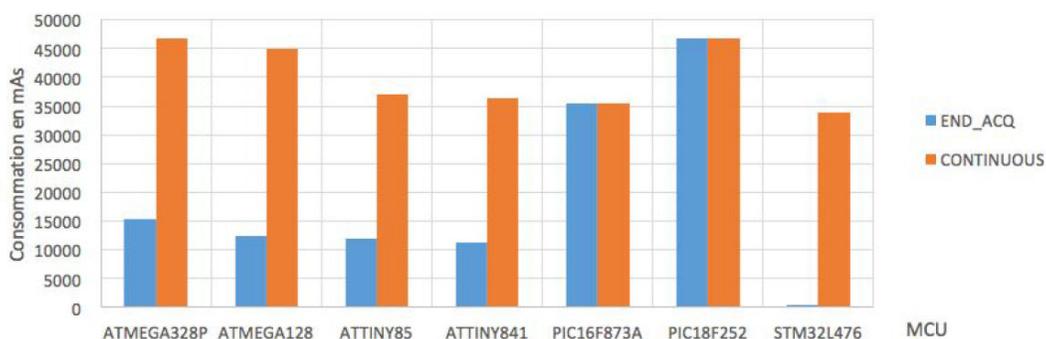
L'architecture la plus économe que nous pouvons concevoir est un noeud de capteur composé d'un seul capteur et des composants comme l'unité de traitement et la radio, et fonctionnant, pour une qualité juste acceptable, avec une fréquence d'échantillonnage de 10 Hz. Avec cette architecture nous avons bien remarqué une différence entre les deux modes de transmission. On constate que le mode de transmission *END_ACQ* est le plus intéressant au niveau de la consommation par rapport au mode de transmission continue (*CONTINUOUS*). Cette différence de consommation montre, comme on pouvait s'y attendre, que plus le nombre de transmission augmente, plus grande est la consommation du noeud. A l'issue de cette constatation, l'idéal serait de diminuer considérablement le nombre de transmission mais cette diminution requiert une capacité mémoire suffisante pour stocker l'ensemble des données acquises. Pour cette raison, il est nécessaire d'étudier l'influence de la taille de la mémoire RAM du microcontrôleur.

Dans le cas idéal, l'objectif serait de ne réaliser qu'une seule transmission à la fin de l'intervalle d'acquisition. Pour ce faire, la capacité de la RAM doit permettre de mémoriser l'ensemble des données. Aussi les microcontrôleurs présentant une capacité mémoire suffisamment grande devrait permettre de réduire la consommation car ils limitent le nombre de communications. L'objectif est d'occuper au maximum la mémoire RAM. A titre d'exemple, le STM32 possède une capacité mémoire suffisamment grande mais aussi une faible consommation de courant. Pour résumer ce premier résultat, pour favoriser le respect des contraintes utilisateurs (autonomie) et applicatives (taux de reconnaissance), il est préférable de disposer d'unité de traitement capable de mémoriser un nombre suffisant de données.

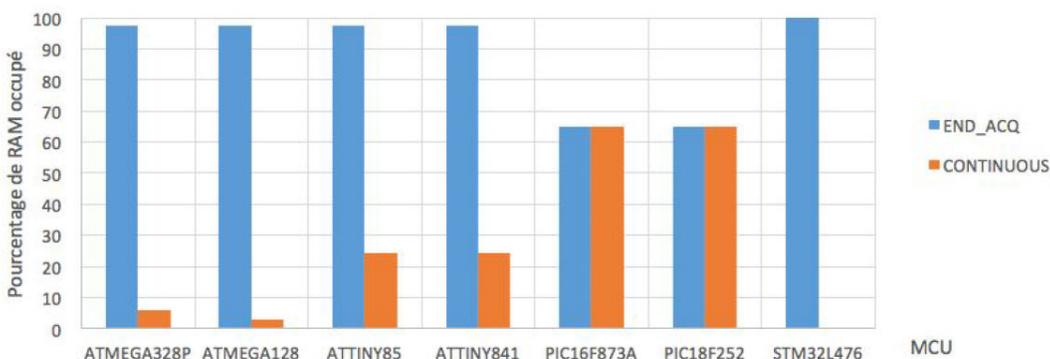
La simulation d'une architecture composée d'un module radio Bluetooth, d'un capteur avec une fréquence d'échantillonnage de 100 Hz et d'une unité de traitement STM32 consomme environ 450 mAs (voir figure 4.6a) et utilise presque 100% de sa mémoire RAM (voir figure 4.6b). L'ATMEGA328P consomme environ 15 As et l'ATMEGA128 et les ATTINY consomment environ 12 As. Ces 4 derniers types de mcu utilisent environ 98% de leur mémoire. Les PICs présentent la pire performance avec des consommations de 35 et 45 As, et une utilisation non optimale de la mémoire RAM (environ 65%). Avec cette configuration, les ATTINY possèdent les meilleures performances en terme de consommation et gèrent aussi bien l'occupation de la mémoire

4.3 Résultats et discussions

que les ATMEGA. Mais en résumé, c'est le microcontrôleur STM32 qui présente une consommation de courant très faible par rapport aux autres microcontrôleurs mentionnés en raison d'une faible consommation intrinsèque mais aussi d'une capacité de mémoire RAM adaptée. A noter que, le choix de la meilleure architecture dépend de sa consommation d'énergie mais aussi de l'utilisation de la mémoire RAM qui impacte le nombre de transmission.



(a) Quantité de courant consommée

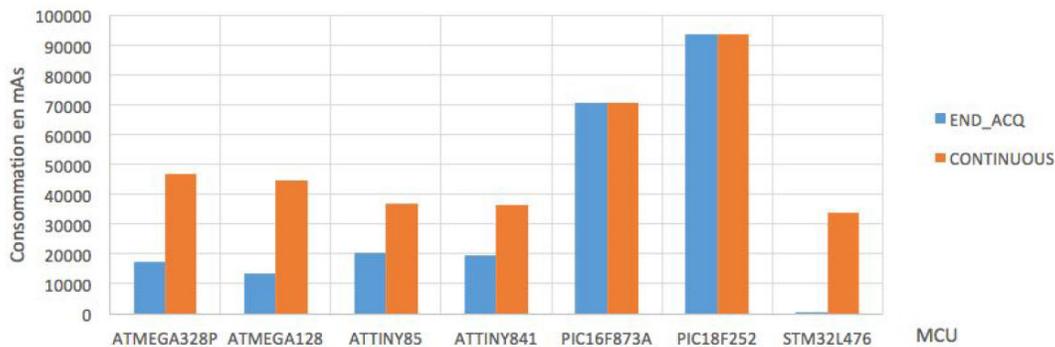


(b) L'occupation de la mémoire

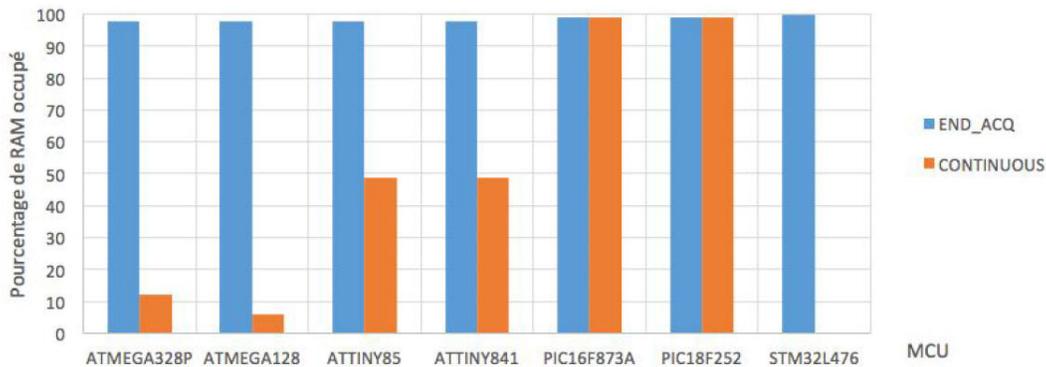
FIGURE 4.6 – Résultat de l'exploration d'architecture pour un noeud composé d'un capteur et fonctionnant à 100 Hz. A gauche, la consommation en courant et à droite, l'occupation de la mémoire

Considérons maintenant la même architecture dotée cette fois de deux capteurs et fonctionnant à une fréquence d'échantillonnage de 100 Hz. Les microcontrôleurs atteignent entre 97% et 99.95% du taux d'occupation de la mémoire RAM avec une transmission en fin d'acquisition comme on peut remarquer sur la figure 4.7. Dans le cas d'une transmission continue, après le STM32, les ATTINY ont la meilleure performance en terme de consommation (autour de 36 As : voir figure 4.7a), même s'ils utilisent moins de 50% de leur mémoire RAM. On peut noter que les PIC présentent une utilisation plus optimale de la mémoire avec presque 100% d'utilisation. Pour cette architecture, les mcus de type ATMEGA en mode de transmission à la fin de l'acquisition garantissent les meilleures performances. Dans le cas de trois capteurs

(voir figure 4.8), les microcontrôleurs s’approchent des limites de leurs mémoires, et les deux modes de transmission deviennent presque équivalents, sauf pour les ATMEGA et le STM32, qui ont des espaces mémoires supérieurs ou égaux à 2 Ko. Ainsi, grâce à leur mémoire RAM, pour le mode de transmission à la fin de l’acquisition, c’est le STM32 puis l’ATMEGA128 qui consomment le moins. D’autre part, pour le mode de transmission continue, c’est les ATTINY et le STM32 qui possèdent une consommation minimale (environ 38 As : voir figure 4.8a). Pour conclure cette étude, il serait intéressant en termes d’autonomie et d’utilisation de la mémoire d’utiliser un ATMEGA couplé au mode de transmission à la fin de l’acquisition. Si on analyse l’effet du nombre de capteurs de signaux physiologiques sur la consommation du système, nous pouvons dire que l’architecture avec un microcontrôleur de type ATMEGA (ATMEGA128 ou ATMEGA328P) permettrait de garantir une utilisation optimale de la mémoire RAM et présente une consommation de courant relativement faible.



(a) Quantité de courant consommé

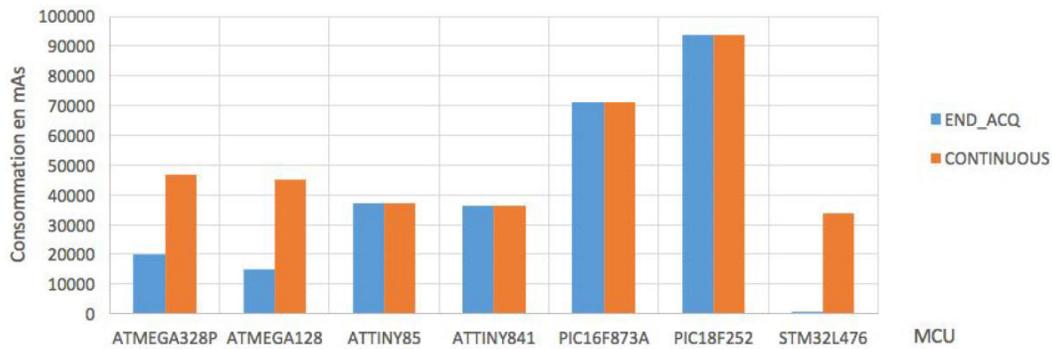


(b) L’occupation de la mémoire

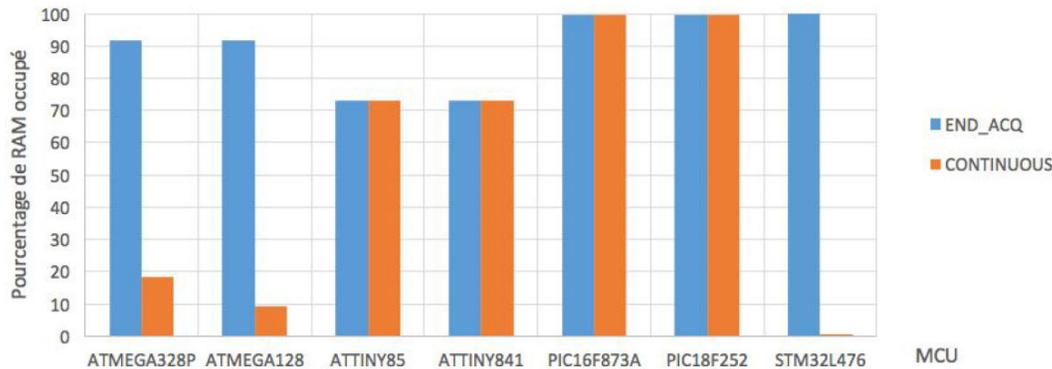
FIGURE 4.7 – Résultat de simulation pour 2 capteurs à 100 Hz

Dans la suite de cette section, nous proposons d’étudier l’influence, sur la consommation, de la fréquence d’échantillonnage des capteurs de signaux physiologiques et l’utilisation de la mémoire RAM pour une architecture composée de trois capteurs. Nous rappelons que tous les microcontrôleurs fonctionnent à 8 MHz.

4.3 Résultats et discussions



(a) consommation en courant



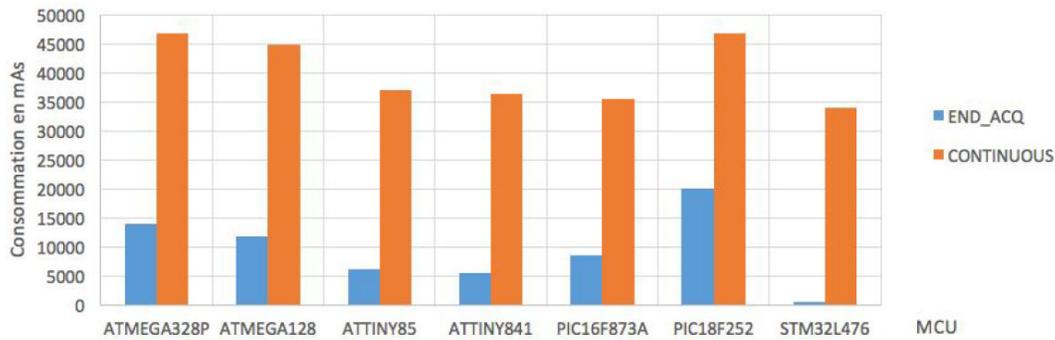
(b) L'occupation de la mémoire

FIGURE 4.8 – Résultat de simulation pour 3 capteurs à 100 Hz

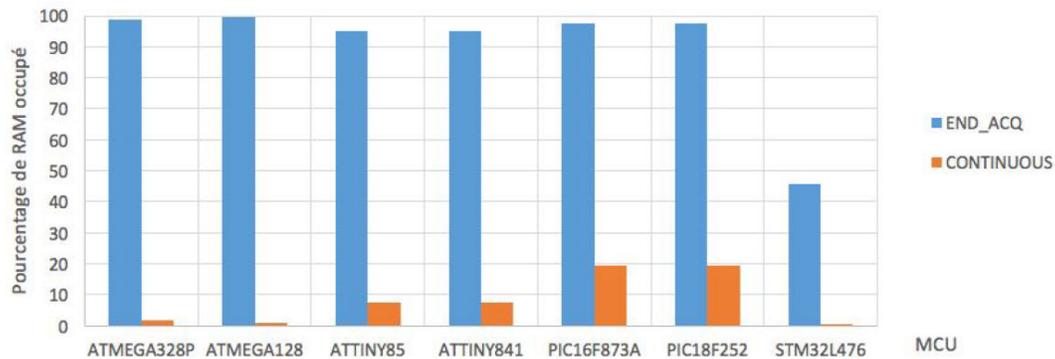
4.3.3 Étude de l'influence de la fréquence d'échantillonnage sur la consommation

Pour une fréquence d'échantillonnage de 10 Hz, la consommation est minimale pour une transmission à la fin de l'acquisition. Si on compare les microcontrôleurs, le PIC18F252 (20 As) est le plus consommateur suivi par les ATTINY qui consomment autour de 5 As. Le STM32 consomme très peu, environ 419 mAs. En mode de transmission continue, les ATMEGAs consomment autant que le PIC18 (autour des 45 As) alors que les autres MCU consomment autour de 35 As (voir figure 4.9a). En ce qui concerne l'utilisation de la RAM, le STM32 est un cas particulier vu la taille de sa mémoire. D'autre part, vu le nombre de capteurs, même pour une fréquence de 10 Hz, la mémoire est occupée à peu près à 100% pour une transmission *END_ACQ* (voir figure 4.9b). Finalement, on peut noter dans ce cas que la petite taille de la RAM des ATTINY font d'eux les microcontrôleurs les mieux adaptés en raison de leur utilisation optimale de la RAM et de leur basse consommation (voir figure 4.9).

Pour une fréquence d'échantillonnage de 100 Hz (voir figure 4.8a), la consommation est la même pour les deux modes de transmission pour les MCU ayant de petites mémoires RAM. C'est le cas des ATTINY et des PIC. D'autre part, pour les MCU qui



(a) Quantité de courant consommée



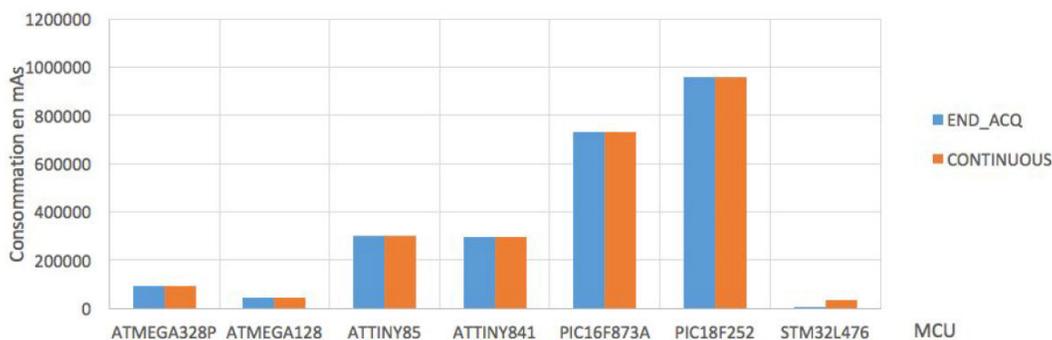
(b) L'occupation de la mémoire

FIGURE 4.9 – Résultats de simulation pour 3 capteurs à 10 Hz

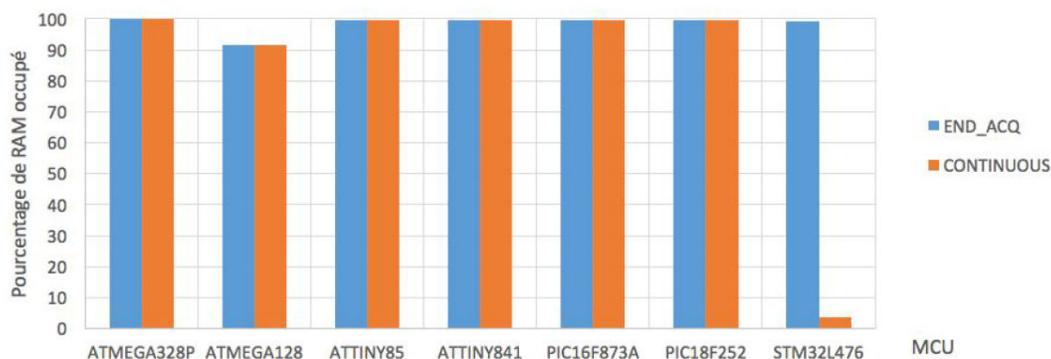
ont une taille de RAM plus grande, comme les ATMEGA (2 et 4 Ko), une transmission à la fin de l'acquisition reste largement plus avantageuse au niveau énergétique avec une consommation maximale en *END_ACQ* (pour l'ATMEGA328P) de 20 As contre 45 As en transmission continue. Encore une fois, le STM32 est un cas particulier avec une consommation énergétique très basse pour un mode de transmission à la fin de l'acquisition grâce à la taille de sa RAM. Le graphique de l'utilisation de l'espace RAM (voir figure 4.8b) confirme nos remarques précédentes. En effet, les ATTINY et les PIC atteignent les limites de leur RAM. La taille des données (3 capteurs * 100 échantillons * 10 bits d'ADC = 3000 bits/seconde) explique le fait que seul 73% environ de la RAM des ATTINY (4096 bits) est utilisée. Ce fait explique également pourquoi les deux modes de transmission présentés deviennent équivalents. D'autre part, nous constatons aussi que les MCU avec des RAM plus grandes, qui n'atteignent pas encore leurs limites dans ce cas, présentent une utilisation plus optimale de leur mémoire dans le cas d'une transmission à la fin de l'acquisition. Pour conclure, dans ce cas, il serait intéressant d'utiliser l'ATMEGA128 par exemple, couplé au mode de transmission à la fin de l'acquisition. Si on augmente encore la fréquence d'échantillonnage c'est à dire à 1000 Hz, la figure 4.10 montre les résultats obtenus pour l'architecture doté de trois capteurs. Les deux modes de transmission deviennent équivalents pour presque tous les

4.3 Résultats et discussions

microcontrôleurs car même avec un mode de transmission à la fin de l'acquisition, la RAM se sature très vite ce qui amène l'ordonnanceur à envoyer les données aussitôt acquises afin d'éviter la perte de données. Il n'est pas envisageable d'avoir des pertes de donnée pour notre application car ceci pourrait diminuer le taux de reconnaissance des émotions ou conduire à des erreurs de détection. Ceux sont les microcontrôleurs avec le plus grand espace mémoire qui présentent une consommation minimale.



(a) Quantité de courant consommée



(b) L'occupation de la mémoire

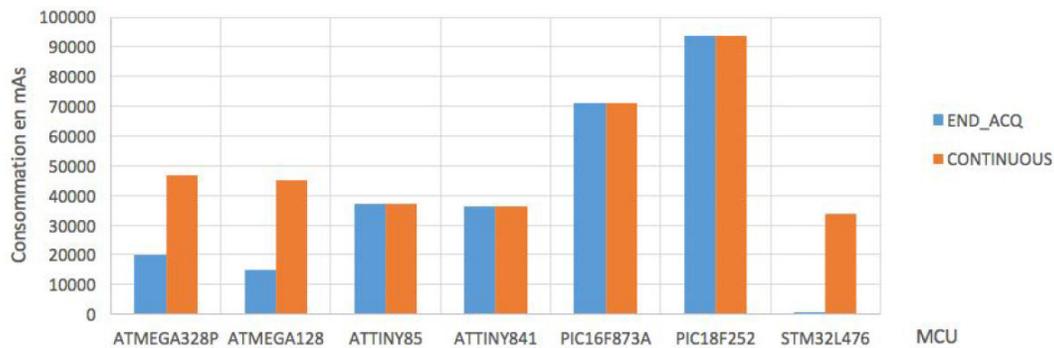
FIGURE 4.10 – Résultats de simulation pour 3 capteurs à 1000Hz

4.3.4 Étude de l'influence du protocole de communication sur la consommation

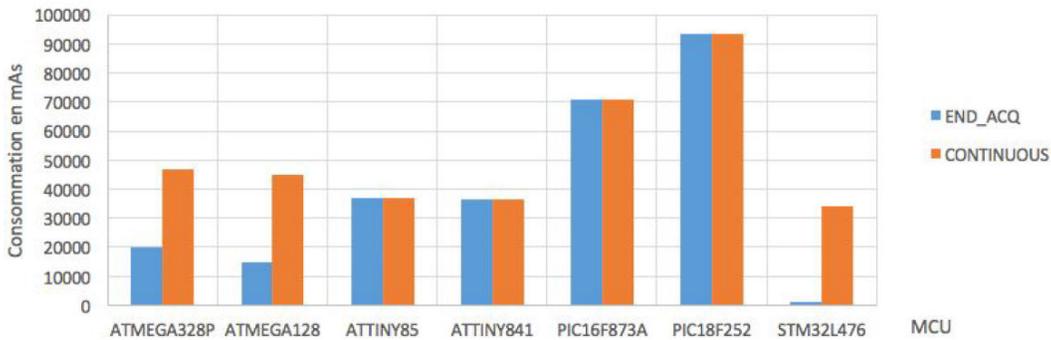
Dans le chapitre précédent, nous avons étudié l'impact du choix du protocole de communication sur la consommation. Dans cette partie du document, cette étude va être complétée par l'influence du choix de l'unité de traitement.

Les quatre protocoles de communication que nous avons sélectionnés présentent des débits différents. Par exemple, le protocole de communication Zigbee possède un débit de 250 Kbits/seconde alors que celui du Bluetooth est de 1 Mbits/seconde. Pour ces deux protocoles, la figure 4.11 présente les différents résultats de simulation obtenus pour les différents microcontrôleurs modélisés dans notre framework. Les résultats obtenus en

terme de consommation pour ces deux protocoles sont à peu près identiques. Ce fait est du que, tant que le nombre de bits stockés dans la mémoire ne dépasse pas le débit du protocole, les données sont transmises le même nombre de fois. Dans le cas contraire, le nombre de transmissions augmente et entraîne une augmentation de la consommation de courant. On constate ainsi une augmentation de la consommation pour le mcu STM32 (pour le mode de transmission à la fin de l'acquisition) car dans le cas de notre application, il est nécessaire et possible de stocker dans la mémoire RAM 786 Kbits, capacité supérieure au débit du Zigbee (>250 Kbits/s). En effet, la consommation, dans ce cas, passe de 524 mAs pour le Bluetooth à 1,328 As pour le Zigbee. On note que l'occupation de la mémoire n'est pas modifiée quand le protocole de communication passe du Bluetooth au Zigbee.



(a) protocole Bluetooth



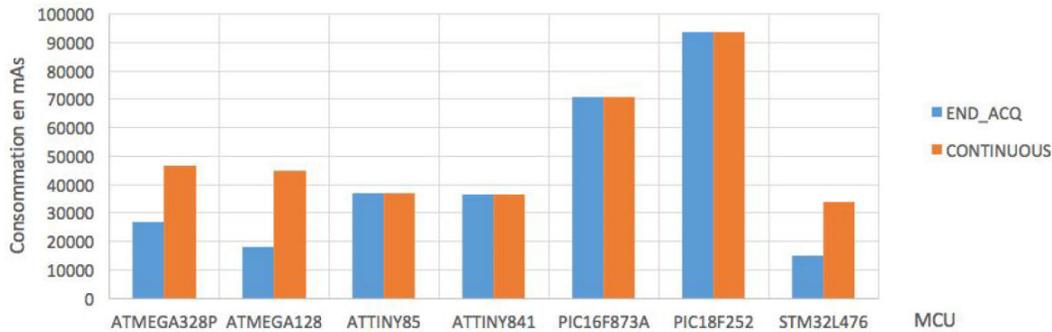
(b) protocole Zigbee

FIGURE 4.11 – Résultat de simulation de consommation en courant pour les protocoles Bluetooth et Zigbee

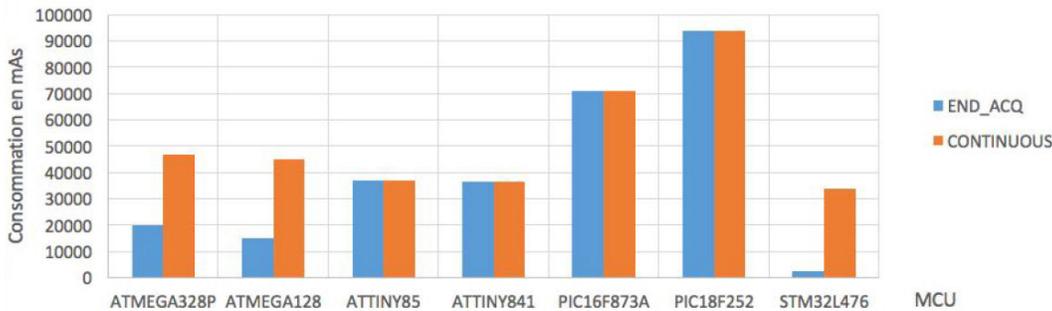
Dans le cas du protocole de communication LoRa avec la modulation LoRa, le débit est de 10938 bits/s. Cette forte diminution du débit par rapport aux protocoles de communication Zigbee et Bluetooth devrait engendrer un nombre supérieur de transmission donc une augmentation de la consommation (voir figure 4.12). En fait, dans le cas des ATTINY et des PIC, il n'y a pas de changement réel de consommation, vu que la capacité mémoire de ces mcu est inférieure au débit actuel, le nombre de transmission est relativement stable. Pour les MCU avec des mémoires RAM plus grandes, comme

4.3 Résultats et discussions

les ATMEGA et le STM32, on observe une augmentation de la consommation qui est liée à l'accroissement du nombre de transmission. Par exemple, la consommation de l'ATMEGA128 passe de 15 As environ à 18 As. Malgré cette augmentation, le STM32 et l'ATMEGA128 restent les deux microcontrôleurs les plus économes. Pour le protocole de communication LoRa avec la modulation GFSK, le débit est de 50 Kbits/s. Ce débit, qui est supérieur à celui de la modulation LoRa, donne des résultats presque similaires à ceux du Zigbee. Pour le cas du mcu STM32, la consommation obtenue est plus grande que dans le cas du Bluetooth avec 2,5 As au lieu de 570 mAs environ.



(a) protocole Lora avec la modulation Lora



(b) protocole Lora avec la modulation GFSK

FIGURE 4.12 – Résultat de simulation de consommation en courant pour le protocole Lora avec la modulation Lora à gauche et la modulation GFSK à droite

L'exploitation de ces différents résultats permet de noter que les microcontrôleurs avec le plus grand espace mémoire RAM tendent à diminuer la consommation d'énergie, à condition d'utiliser le mode de transmission à la fin de l'acquisition. Ceci s'explique par le fait que, quand la taille de la RAM augmente, le nombre de transmission (la taille des données augmente mais cette augmentation n'impacte pas de façon considérable la consommation) tend à diminuer. Cependant une limite existe, elle est liée au débit de chaque protocole de communication sélectionné. Quand ce débit diminue, le nombre de transmission augmente, ce qui par voie de conséquence accroît la consommation. D'autre part, l'augmentation du nombre de capteurs ou de la fréquence d'échantillonnage augmentent la quantité de données à transmettre, et donc le nombre de transmissions effectuées, tend à accroître la consommation. Concernant le choix de l'unité de

traitement, on peut conclure que le STM32L476 et les ATMEGA (ATMEGA128 et ATMEGA328P) paraissent intéressants dans le cas de notre application car ils présentent une consommation énergétique et un taux d’occupation de la mémoire RAM acceptables. En effet, les différents résultats présentés dans les tableaux ci-dessus montrent que l’architecture optimale est composée d’un type de ces 2 mcus pour une fréquence d’échantillonnage de 100 Hz. De plus, l’association de ces mcu avec le protocole de communication Bluetooth avec le mode de transmission END_ACQ permettent de rendre l’architecture encore plus autonome et performante.

Batterie	Modèle d’estimation	$A_{scenario}$ (jours)
500 mAh	Modèle analytique	16.2625
	Modèle du framework	19.5175
800 mAh	Modèle analytique	33.0496
	Modèle du framework	31.2280
1300 mAh	Modèle analytique	61.0282
	Modèle du framework	50.7456

TABLE 4.3 – Comparaison des deux modèles d’estimation développés

4.3.5 Modélisation du système d’acquisition avec trois nœuds capteur : étude de l’autonomie

Afin de positionner les capteurs physiologiques aux points essentiels sur le corps et de façon à occasionner le moins de gêne à l’utilisateur, nous avons proposé de concevoir un réseau de capteurs permettant d’acquérir les signaux physiologiques. L’architecture de ce réseau est composée de trois nœuds dotés individuellement d’un capteur unique de signaux physiologiques. ces nœuds intègrent aussi un module radio de type Bluetooth, un module d’alimentation et une unité de traitement, l’ATMEGA328P. (Nous avons choisi ces modules (mcu, radio) en fonction des plateformes disponibles comme BITALINO qui nous permettront par la suite de faire une acquisition réelle des signaux. La figure 4.13 montre l’architecture matérielle du noeud de capteur que nous proposons. Ces nœuds capteur ont été modélisés au sein de notre framework. Le modèle permettant de calculer la quantité de courant consommée a été couplé avec le modèle défini dans le chapitre précédent pour estimer l’autonomie du système pour une batterie donnée. Ainsi, le tableau 4.3 présente les résultats obtenus pour trois capacités de batterie.

Dans ce tableau (4.3), nous présentons aussi les résultats sur l’autonomie obtenus selon deux principes de calcul pour la quantité de courant consommée. Le modèle analytique, présenté dans le chapitre précédent, s’appuie entre autre sur les caractéristiques du scénario applicatif préalablement établi. Rappelons que ce scénario consiste à acquérir les données deux fois par jour et chaque session d’acquisition a une durée de 20 minutes. Le deuxième modèle est celui proposé par notre framework

4.3 Résultats et discussions

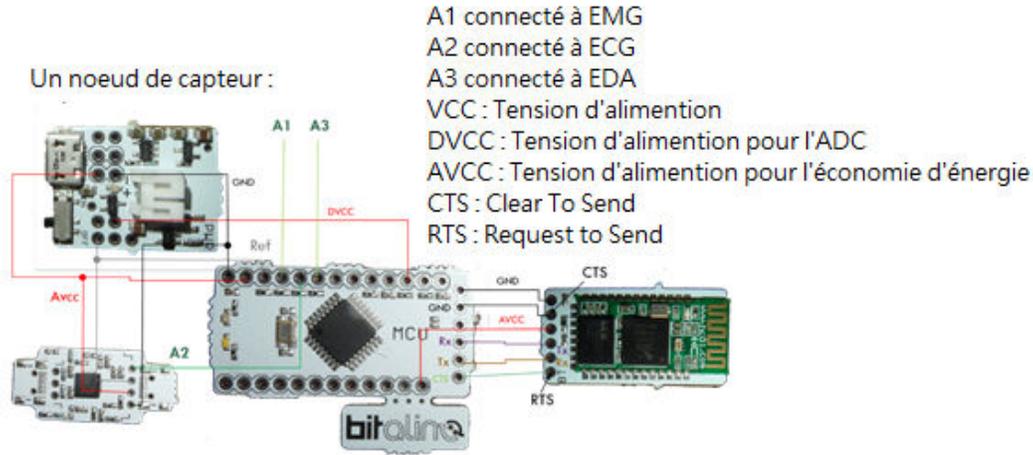


FIGURE 4.13 – Architecture matérielle proposée pour un noeud de capteur

qui s'appuie sur une simulation des communications entre les modules du système d'acquisition.

Les deux modèles produisent des valeurs sur l'autonomie qui peuvent être éloignées dans certains cas. C'est le cas en particulier pour la batterie de plus haute capacité c'est à dire de 1300 mAh. Le framework développé a l'avantage d'étudier, en plus de la consommation d'énergie, l'occupation de la mémoire, la synchronisation des données, la comparaison de différentes combinaisons architecturales et la prise en compte de contraintes. D'autre part, les résultats présentés dans le chapitre 3 montre que le modèle analytique d'estimation de la consommation est beaucoup plus proche des valeurs mesurées sur un noeud de capteur réel. En effet, ce modèle analytique d'estimation prend en compte non seulement les contraintes d'architecture mais aussi les contraintes applicatives, ce qui rend ce modèle plus performant que celui du Framework. De ce fait, nous avons remplacé le modèle énergétique du Framework par le modèle analytique présenté dans le chapitre 3. Cette modification améliore l'estimation de la consommation dans le Framework, les performances ainsi estimées seront plus proches des performances réelles du système. Ainsi le modèle modifié de consommation de courant dans le framework 3 est donné par l'équation 4.3 :

$$I_{consommé} = I_{static} + \frac{I_{mcu} * T_{processing} + I_{sensors} * T_{sensing} + Q_{conradio} + Nb_{trans} * Q_{transradio}}{T_{IE}} \quad (4.3)$$

$I_{consommé}$ est la consommation totale de courant du noeud de capteur pendant le temps d'acquisition (" T_{IE} ") imposé par le scénario applicatif. Ainsi, dans le framework, nous calculons la consommation de chaque module (processing, sensors, radio) à chaque fois que le noeud rentre dans une fonction de ces modules. A la fin de l'intervalle temporel d'acquisition, la somme de ces quantités de courant est divisée par T_{IE} . Le courant statique est ensuite ajouté comme illustré par l'équation, du courant total

consommé, 4.3. Les autres paramètres de ce modèle d'estimation ont été introduits et expliqués dans le chapitre précédent. Cette modification dans le modèle d'estimation de la consommation énergétique permet d'une part de palier au problème de la simplicité du premier modèle proposé. En effet, l'absence de la prise en compte des contraintes dans ce modèle simplifie ce dernier et diminue les performances de l'estimation de la consommation d'énergie. D'autre part, cette modification permet de coupler ces deux travaux sur l'estimation de la consommation et ainsi de rendre notre modèle plus performant et plus rigoureux. Ainsi, les estimations des performances d'autonomie de l'architecture du nœud de capteur que nous proposons sont les mêmes autonomies obtenues avec le modèle analytique. Le modèle de consommation donne une erreur maximale de 7.75% par rapport aux mesures sur une plateforme réelle (voir chapitre 3).

Comme montré sur la figure 4.13, la dernière architecture considérée s'appuie sur un réseau distribué de collecte de données. Chaque nœud est composé d'un seul capteur, d'une unité de traitement, d'un module RF et d'une batterie. Cette architecture en plus d'être non-invasive, c'est à dire qu'elle n'est pas contraignante en terme de poids, de taille et de complexité, est modulaire dans le sens où on peut facilement ajouter ou enlever des nœuds sans impacter le fonctionnement global du système d'acquisition. Cette nouvelle architecture offre de la mobilité en rendant possible les acquisitions lorsque la personne est en mouvement. Cette nouvelle architecture garantit une meilleure autonomie par rapport à la première. En effet, pour une fréquence d'échantillonnage de 100 Hz, nous obtenons une autonomie de 38 jours avec cette nouvelle architecture contre 12 jours avec l'architecture composée d'un nœud doté de tous les capteurs.

4.4 Conclusion

Au début de ce chapitre, nous avons présenté un bref état de l'art sur les simulateurs des réseaux de capteurs sans fil. Cet état de l'art, ainsi que l'analyse des spécificités de l'application, nous ont permis de montrer le besoin de développer notre propre Framework pour la simulation et l'exploration d'architecture de système d'acquisition pour la détection des émotions ou plus généralement pour des applications d'e-santé. Ce framework étant développé en SystemC, nous avons, tout d'abord, introduit les concepts sous-adjacents à cet environnement de simulation et ceci au niveau TLM. Le framework développé permet de simuler à un haut niveau des objets communicants portables pour l'e-santé. Il vise à fournir une aide à la conception en permettant d'estimer de manière rapide et concise la consommation d'énergie d'un nœud/réseau et de faire une exploration architecturale en tenant compte des contraintes de l'application, de l'architecture et de l'utilisateur. L'application visée étant la détection des émotions, il est nécessaire de noter que ce type d'application diffère des applications classiques de réseau de capteurs sans fil, du fait d'une part que le temps d'acquisition est élevé, le nombre de séances d'acquisition par jour est faible et le temps entre deux sessions

4.4 Conclusion

d'acquisition est relativement élevé. D'autre part, le réseau est constitué d'un faible nombre de nœuds capteur, au nombre de cinq au maximum pour un système de détection des émotions et selon une topologie en étoile. Le système d'acquisition ne doit pas être invasif et il faut noter qu'à partir de trois nœuds, on obtient des taux de reconnaissance satisfaisants. Les contraintes de ces systèmes sont différentes de celles des réseaux de capteurs sans fil classiques qui priorisent souvent la latence sans vraiment tenir compte de la qualité des signaux. Les différents modules qui composent notre système d'acquisition ont été modélisés par des machines d'états. Cette modélisation a permis de proposer des modèles génériques et de réaliser rapidement la simulation d'un système d'acquisition. Cette simulation permet, d'une part, d'estimer l'autonomie, d'évaluer la performance et de faire de l'exploration d'architecture à l'aide des bibliothèques de modules de traitement, radio, capteurs. La synchronisation entre les modules est réalisée en utilisant des événements SystemC et s'appuie sur le module Timer qui communique avec les transactions TLM.

Dans le but de faire de l'exploration d'architecture pour un système d'acquisition, de nombreux résultats ont ainsi été produits par notre framework. Nous avons, entre autre, étudié l'influence des spécificités de l'unité de traitement, de la fréquence d'échantillonnage, du protocole de communication et du nombre de capteurs sur la consommation. L'analyse des nombreux résultats fournis par le Framework nous a permis de proposer une architecture matérielle pour le système Emotica moins invasive, mobile (offrant la possibilité d'acquérir les signaux physiologiques lorsque le sujet est en mouvement), et présentant une autonomie acceptable. L'architecture ainsi obtenue possède une faible consommation d'énergie (grande autonomie) et un taux de reconnaissance des émotions élevé. Dans le chapitre suivant, cette architecture matérielle sera utilisée pour collecter les données et créer une nouvelle base d'apprentissage. De même, l'architectures logicielle de traitement y sera décrite.

Le système Emotica

Sommaire

5.1	Architecture de collecte de données d'Emotica	122
5.2	Architecture de traitement des données dans le système Emotica	124
5.2.1	Application Matlab	124
5.2.2	Application mobile	127
5.3	Application visée et conception de notre propre base d'apprentissage de signaux physiologiques	129
5.4	Résultats et discussions	134
5.5	Conclusion	136

Le projet Emotica, qui cible la détection des émotions, permet le développement d'applications qui peuvent s'inscrire dans les défis de l'e-santé mais aussi dans certains défis socio-économiques (robotique, maisons intelligentes, villes intelligentes ...). En effet, l'idée de détecter, d'analyser les émotions est aujourd'hui fortement convoitée dans de nombreux secteurs d'activités professionnelles. A titre d'exemple, ce type de système est déjà utilisé, de nos jours, par certains constructeurs automobiles pour optimiser le plaisir de conduite ou encore dans les jeux vidéo pour détecter une éventuelle addiction des joueurs. Elle est aussi appréciée au niveau du marketing. Les grandes entreprises l'ont bien compris, appréhender les émotions que procurent leurs produits aux clients en temps réel permettrait de s'adapter aux besoins des consommateurs et d'augmenter le chiffre d'affaire en conséquence. Dans le cadre de notre étude, nous avons souhaité nous concentrer sur des applications orientées vers le médical.

Emotica (capture d'émotion) est un système de détection des émotions basé sur l'acquisition de signaux physiologiques à l'aide, de capteurs de signaux du système nerveux autonome. Comme le montre la figure 5.1, le système est composé de deux phases. La première consiste à former un base d'apprentissage c'est à dire à apprendre des méthodes permettant de détecter une activité émotionnelle et de la distinguer d'une autre. Dans la deuxième phase, Emotica se sert des connaissances apprises lors de la

première phase pour détecter et quantifier automatiquement une activité émotionnelle. Le système Emotica est donc un système complet de détection des émotions. En effet, il est composé d'une architecture matérielle et logicielle permettant l'acquisition et le traitement des données acquises.

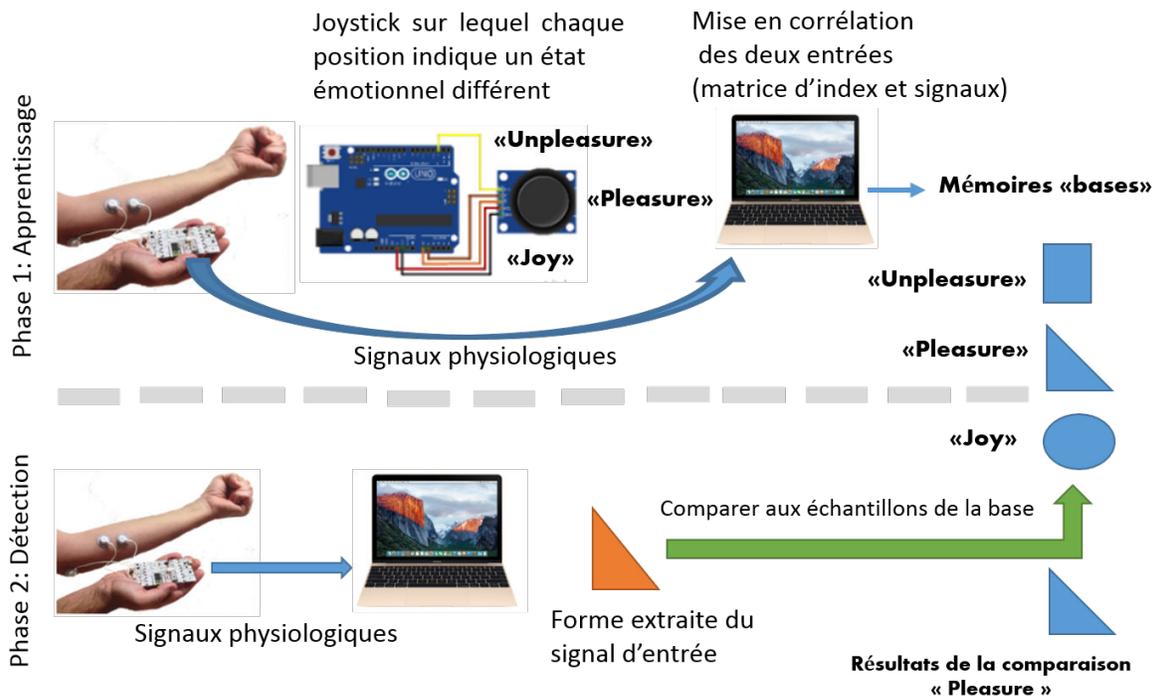


FIGURE 5.1 – Fonctionnement global du système Emotica

La figure 5.1 permet de mieux comprendre le fonctionnement d'un système de détection des émotions en assimilant cette détection à une reconnaissance de forme. Dans la phase d'apprentissage, Emotica caractérise chaque émotion et lui attribue une forme particulière. Ainsi, les émotions *unpleasure*, *pleasure*, *joy* sont respectivement assimilées aux formes carrée, triangulaire et circulaire. Ces différentes formes géométriques forment la base d'apprentissage. Dans la phase de détection, lorsqu'une émotion est détectée, Emotica la caractérise en déterminant sa forme puis regarde dans sa base d'apprentissage la forme géométrique la plus proche. L'émotion reliée à la forme la plus proche devient l'émotion détectée.

Ce chapitre est exclusivement consacré à la mise en place d'une application basée, sur le système de détection des émotions Emotica, en mettant un accent particulier sur la conception de la base de données. Afin d'utiliser le système EMOTICA pour une application concrète médicale, nous avons choisi avec l'aide de la fondation Noizier à Sophia Antipolis, de l'appliquer à la "Mnémothérapie" musicale pour les patients atteints de la maladie d'Alzheimer. Dans un premier temps, l'architecture matérielle du système complet de collecte est décrite, suivie par celle de l'architecture logicielle sous-jacente. Ces deux architectures serviront, entre autre, à la conception d'une nouvelle

base de données de signaux physiologiques dont le but est de caractériser différents états émotionnels. Dans l'objectif de faciliter les mesures, une version sur mobile a été développée. Il a ensuite été conçu une base de données pour la "Mnémothérapie" qui s'appuie sur une approche de stimulation musicale. Les différentes expérimentations seront alors présentées ainsi que la constitution de la base de données.

5.1 Architecture de collecte de données d'Emotica

Emotica signifie capture d'émotions. Pour capturer des émotions, le système Emotica a besoin de signaux à analyser. Ces signaux sont acquis par des capteurs. Pour concevoir notre système Emotica, nous avons dans un premier temps défini et réalisé l'architecture matérielle de nos nœuds de capteurs. Cette architecture comme le montre la figure 5.2 est composée de trois nœuds de capteurs. Le premier nœud *ECG node* permet l'acquisition des signaux du rythme cardiaque, le deuxième *EDA node* monitoré la variation des signaux électro-dermales et le troisième *EMG node* mesure l'activité électrique des muscles.

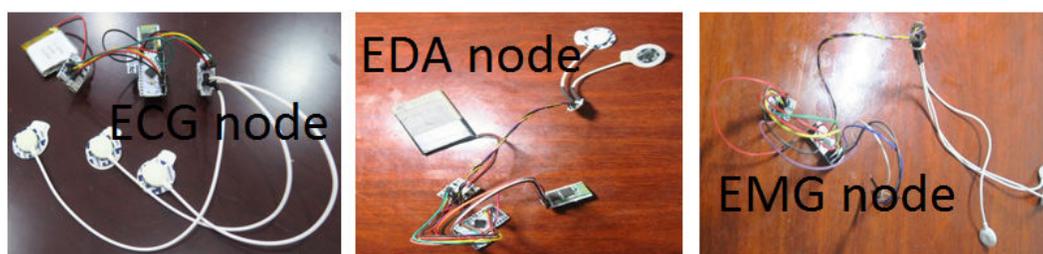


FIGURE 5.2 – Les trois nœuds d'Emotica

Cette architecture en plus d'être non-invasive, c'est à dire qu'elle n'est pas contraignante en terme de poids, de taille et de complexité, est modulaire dans le sens où on peut facilement faire varier le nombre de nœuds sans impacter le fonctionnement global du système d'acquisition. Les quatre points suivants montrent en quoi cette architecture à trois nœuds est moins invasive qu'une architecture qui comporterait les trois capteurs dans un seul nœud.

1. Dans l'architecture composée d'un seul nœud, doté de trois capteurs, les fils des électrodes sont nombreux et il est difficile de les positionner correctement. De plus, cet aspect complique la mobilité du système et entraîne beaucoup de sessions inutiles car les électrodes se détachent souvent.
2. Dans l'architecture qui comporte trois nœuds, chaque nœud capteur sera encapsulé dans un petit boîtier placé au plus proche du point d'acquisition. Par exemple, le nœud ECG pourra ainsi être placé dans la poche d'une chemise et le nœud EMG dans un petit bracelet à placer sur le bras.

5.1 Architecture de collecte de données d'Emotica

3. Les fils actuels d'interconnexions des différents modules (module RF, mcu, capteur) qui composent un nœud seront remplacés par un PCB dans le but de rendre les nœuds de capteurs encore moins invasifs.
4. Enfin, certaines études dans l'état de l'art montrent qu'il est possible de remplacer les électrodes par des patchs collés directement sur le corps. Ceci nous permettrait de rendre le système à trois nœuds encore moins invasif et plus portatif.

Les nœuds capteur du système EMOTICA transmettent leurs données à un unique système de réception chargé de traiter et d'interpréter. La synchronisation des acquisitions est gérée grâce à un ordonnancement réalisé au sein de l'architecture logicielle d'acquisition. L'algorithme d'ordonnancement utilisé est de type *Round Robin* (voir l'annexe D). Nous avons choisi ce type d'ordonnanceur en raison de sa simplicité, et de plus au vu du faible nombre de nœud qui compose l'architecture. Nous aurions pu concevoir un nœud intermédiaire dédié à la collecte des données qui aurait servi principalement d'ordonnanceur mais ceci, aurait pour effet de rajouter une consommation d'énergie supplémentaire. De plus, la latence de communication serait augmentée et pourrait engendrer des pertes des données. Ce fait diminuerait la fiabilité du réseau.

Dans le but de réaliser une base d'apprentissage pour EMOTICA, il a fallu définir une méthode d'expérimentations. Il faut rappeler que cette étape est primordiale car elle conditionne la qualité des résultats de détection futurs. Aussi, en plus des nœuds de capteurs constitués pour la collecte des signaux physiologiques, nous avons conçu une manette composée d'un joystick, d'un arduino, d'un afficheur LCD. L'objectif est de provoquer des émotions sur un sujet équipé des capteurs cités précédemment. Le sujet dispose aussi de cette manette pour renseigner son état émotionnel. Grâce à ce dispositif, dans la phase d'apprentissage, le système Emotica fait l'acquisition pour sa base de données de deux entrées couplées : signaux physiologiques et information émotionnelle associée. La manette a été conçue sur la base des émotions qui seront utilisées pour l'application de "Mnémothérapie". Aussi, pour la conception de la base de données associée, nous ne considérerons que 5 états émotionnels. Le choix de ces états émotionnels sera décrit dans la section 5.3. La manette ainsi conçue peut prendre cinq positions différentes dont chacune symbolise une émotion donnée. Ainsi, au cours de chaque session de test, le sujet manipule cette manette suivant son état émotionnel et une matrice associée, comportant entre autre l'indice de début et fin de l'activité émotionnelle, est mise à jour. La première colonne de la matrice indique le nombre d'échantillon (indice dans le signal acquis) et la deuxième colonne indique l'émotion exprimée par le sujet. De ce fait, à la fin de la session d'acquisition, nous extrayons les différentes portions (étape de découpage dans la phase d'apprentissage) de signaux correspondants à une émotion donnée. L'importance de cette manette est double dans notre système Emotica. Tout d'abord, elle permet de recenser facilement les périodes durant lesquelles chaque émotion a été ressentie par le sujet grâce à l'annotation automatique des signaux. Cette annotation permet d'accélérer le processus

5.2 Architecture de traitement des données dans le système Emotica

d'apprentissage et sa validation. Il est important de noter que la manette garantit une condition très importante dans le système de détection des émotions à savoir l'annotation sans biais donc l'acquisition de vrais signaux émotionnelles. En effet, nous savons qu'il n'est pas possible de vivre une émotion pleinement et profondément et d'en parler en même temps, donc, demander au sujet en pleine émotion de donner un feedback sur son état émotionnel le ferait automatiquement sortir de son émotion. Par ailleurs, le sujet peut correctement indiquer son état émotionnel en manipulant un dispositif sans sortir de son émotion si ce dispositif est ludique et simple d'utilisation, sa manipulation doit donc être simple et naturelle. Pour toutes ces raisons, la plateforme comportant le joystick, conçue et présentée sur la figure 5.3 rend notre processus d'apprentissage pertinent.



FIGURE 5.3 – Dispositif servant à l'indexage des signaux physiologiques

5.2 Architecture de traitement des données dans le système Emotica

Après avoir conçu l'architecture matérielle composée des différents nœuds capteur et de la manette servant à annoter les signaux, il est nécessaire de définir l'architecture logicielle. Cette architecture a pour rôle, d'une part, de contenir les algorithmes de détection des émotions et, d'autre part, elle doit disposer d'une interface graphique afin de configurer le système d'acquisition et de visualiser à la fois les signaux reçus ainsi que le résultat de l'analyse émotionnelle. Deux applications ont donc été développées à cet effet : l'une sous Matlab où les premiers algorithmes de détection ont été réalisés, l'autre sous android pour faciliter la mobilité.

5.2.1 Application Matlab

L'application matlab est la première à avoir été développée pour valider les algorithmes de détection développés. Elle est principalement composée de deux parties : l'interface graphique et les algorithmes de détection des émotions. Nos algorithmes ayant déjà été présentés dans le chapitre 2, nous ne présentons dans cette partie que l'interface

5.2 Architecture de traitement des données dans le système Emotica

graphique Matlab incluant l'ensemble des fonctionnalités de configurations du système Emotica. En effet, cette interface graphique permet l'acquisition des données en temps réel, la sauvegarde des données acquises, l'annotation des signaux acquis afin de pouvoir faciliter l'apprentissage et par la même occasion faciliter le test des performances du système. De plus, elle permet l'analyse des signaux acquis et la détection des émotions.

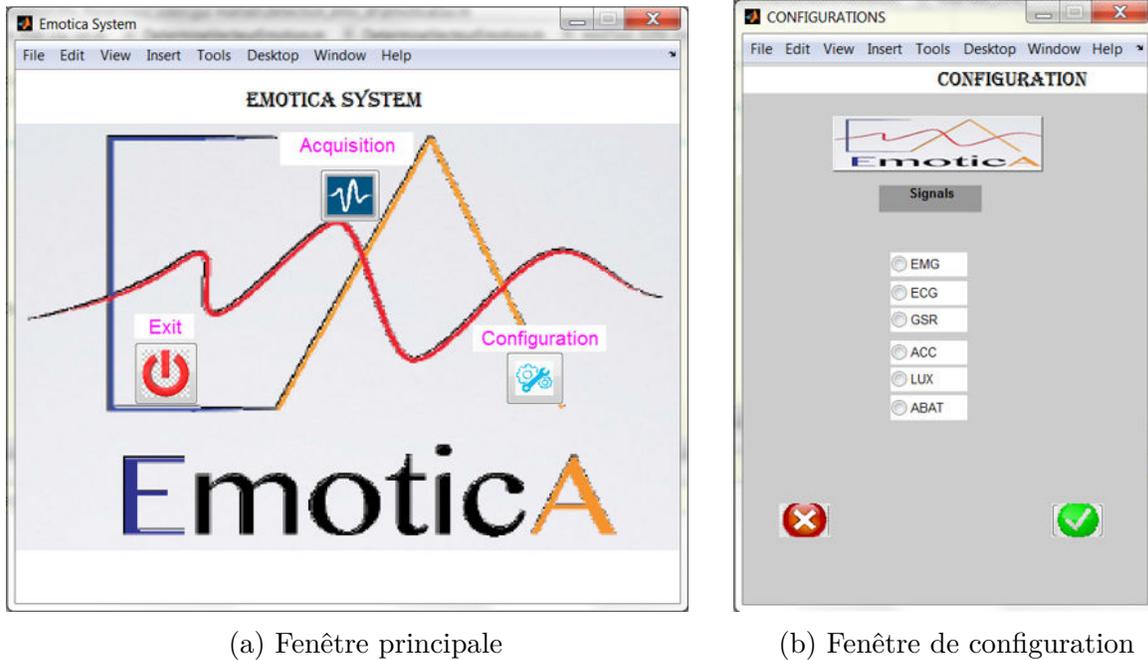


FIGURE 5.4 – Fenêtre principale et celle de configuration du GUI Matlab d'Emotica

La figure 5.4a et la figure 5.4b représentent respectivement la fenêtre principale et la fenêtre de configuration de l'interface graphique d'Emotica sur Matlab. Sur la fenêtre principale, en appuyant sur le bouton **Exit**, l'application quitte. En appuyant sur le bouton **Configuration**, la fenêtre de configuration représentée sur la figure à droite s'ouvre. Cette dernière permet de configurer l'acquisition des données physiologiques. A l'aide des boutons radio, l'utilisateur sélectionne les signaux physiologiques qui seront pris en compte. Chaque signal sélectionné permet d'activer le noeud sur lequel le capteur associé est connecté. Une fois que cette configuration a été réalisée, il est nécessaire de la valider en appuyant sur les boutons poussoirs disposés à cet effet. Si la configuration n'est pas validée, aucun noeud ne sera actif et l'acquisition ne peut se faire, la fenêtre de configuration se ferme alors automatiquement. Par ailleurs, après avoir validé les configurations, seule la fenêtre principale reste affichée. Nous pouvons donc appuyer sur le bouton **Acquisition** (voir figure 5.5) pour commencer une session d'acquisition et visualiser en temps réel les signaux sélectionnés dans l'étape de **Configuration**.

Lorsque nous cliquons sur le bouton **Acquisition** le système va se connecter au(x) noeud(s) sélectionné(s) à l'aide du protocole de communication disponible dans le module RF du ou des noeud(s). Une fois connectée, la fenêtre d'affichage s'ouvre et

5.2 Architecture de traitement des données dans le système Emotica

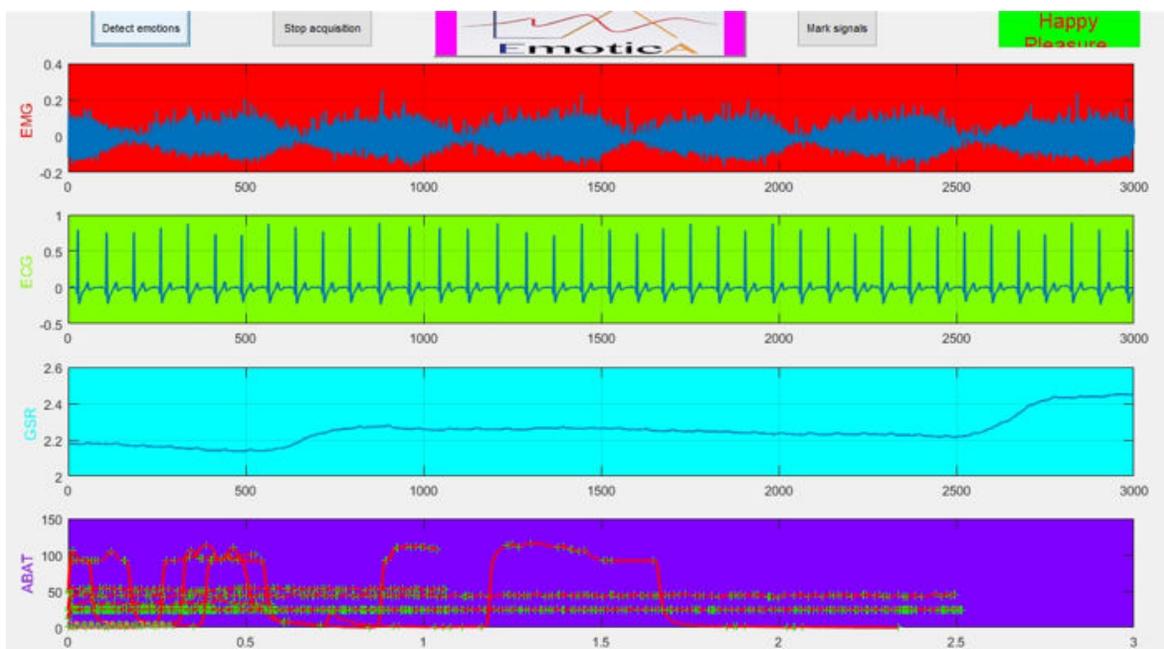


FIGURE 5.5 – Fenêtre d’acquisition du GUI Matlab d’Emotica

l’affichage des signaux en temps réel démarre. Sur la figure 5.5, nous pouvons observer en haut trois boutons et un champs d’édition de texte.

- . En appuyant sur le bouton poussoir **Mark signals**, l’indice du signal est mis dans une matrice pour informer du début et de la fin d’une activité émotionnelle et des lignes verticales sur l’indice du signal sont tracées en discontinue sur les affichages des signaux.
- . L’appui sur le bouton poussoir **Stop acquisition**, stoppe l’acquisition des signaux en désactivant les nœuds actifs et sauvegarde les données brutes acquises dans une première matrice. En parallèle, une seconde matrice regroupe les indices liés aux appuis sur le bouton **Mark signals** et qui correspondent aux débuts et fin d’activité émotionnelle.
- . Emotica permet de faire une détection automatique des émotions. Dans cette application, au lieu d’afficher toutes les activités émotionnelles détectées, un bouton nommé **Detect emotions** permet à l’utilisateur de n’afficher l’état émotionnel que lorsqu’il le souhaite. Suite à l’appui sur ce bouton, les signaux acquis sont filtrés, puis analysés pour extraire des caractéristiques. Si une activité est détectée dans un signal, un vecteur émotionnel associé est déterminé. Les vecteurs émotionnels extraits sont alors fusionnés avec la méthode de vote pondérée que nous avons proposée afin de déterminer une seule émotion traduisant une vue globale de l’état émotionnel d’une personne. Cette émotion, est alors affichée dans la case d’édition de texte colorisée en vert en haut à droite sur la figure 5.5. Pour plus de détails sur ces traitements, le lecteur est invité à relire la section 2.3 décrivant notre approche multimodale de détection

5.2 Architecture de traitement des données dans le système *Emotica*

des émotions. Après chaque appui sur le bouton `Detect emotions`, la matrice contenant les signaux acquis est transformée en une matrice vide. Ceci permet de ne prendre en compte que les signaux acquis entre deux appuis sur ce bouton.

5.2.2 Application mobile

La constitution d'une base de données élargie et la validation sur de nombreux sujets ne sera possible que si on dispose d'un système de détection réduit c'est à dire par exemple inclus dans un téléphone. Aussi, l'état d'avancement du projet *Emotica* a requis le développement d'une première version mobile. L'idée était de proposer un système complet de détection des émotions utilisable par tout le monde et garantissant des performances optimales. Suite à une première version non portable sous Matlab, la volonté d'adapter *Emotica* à un téléphone sous Android est née afin de rendre le système indépendant d'un logiciel précis. Cette volonté est liée entre autres :

- aux contraintes liées à Matlab. Matlab est un logiciel dont la taille est conséquente (3-4 Go) et le prix également. C'est un logiciel présentant une faible adaptabilité. Les applications développées sous Matlab sont dépendantes de ce logiciel. Ainsi, il ne permet pas le développement d'applications grand public.
- aux besoins de fournir un système mobile simple d'utilisation
- au côté pratique pour les médecins exerçant dans des lieux retirés ou lors de leur visite à domicile. Il est dans ce cas plus aisé d'utiliser un téléphone portable.

L'ensemble des algorithmes de détection des émotions ont donc été écrits en Java puis adaptés à Android pour la conception de la version portable. De plus, nous avons conçu et réalisé entièrement une base de données qui regroupe les connaissances nécessaires au système. Cette base de données est également utilisable par les deux applications (Matlab et mobile) et un ensemble de fonctions médicales ont été implémentées (voir annexe B).

Cette application mobile permet d'acquérir et d'afficher en temps réel, sur des graphiques, les activités enregistrées par les nœuds capteurs. Les codes convertis d'*Emotica* ayant été ajoutés, la détection des émotions peut s'effectuer. Cette application mobile accentue les contraintes sur l'autonomie, les performances du système en termes de temps de traitement et de gestion de la mémoire. L'exploration d'architecture réalisée dans les chapitres précédents a permis de proposer une architecture matérielle et logicielle optimale.

Sur la figure 5.6a, nous visualisons la fenêtre principale de l'application. L'interface graphique d'*EMOTICA* mobile permet à l'utilisateur de configurer facilement les nœuds de capteur et d'accéder à différents paramètres. Le menu latéral (à gauche) permet de démarrer une séance d'acquisition, d'accéder aux paramètres de configuration, de consulter la liste des émotions testées, d'aller dans la partie historique, de visionner l'aide et enfin de se renseigner à propos de l'application.

Le menu `settings` permet de configurer l'état des nœuds en activant ou désactivant un ou plusieurs nœuds ainsi qu'en définissant leur fréquence d'échantillonnage.

5.2 Architecture de traitement des données dans le système Emotica

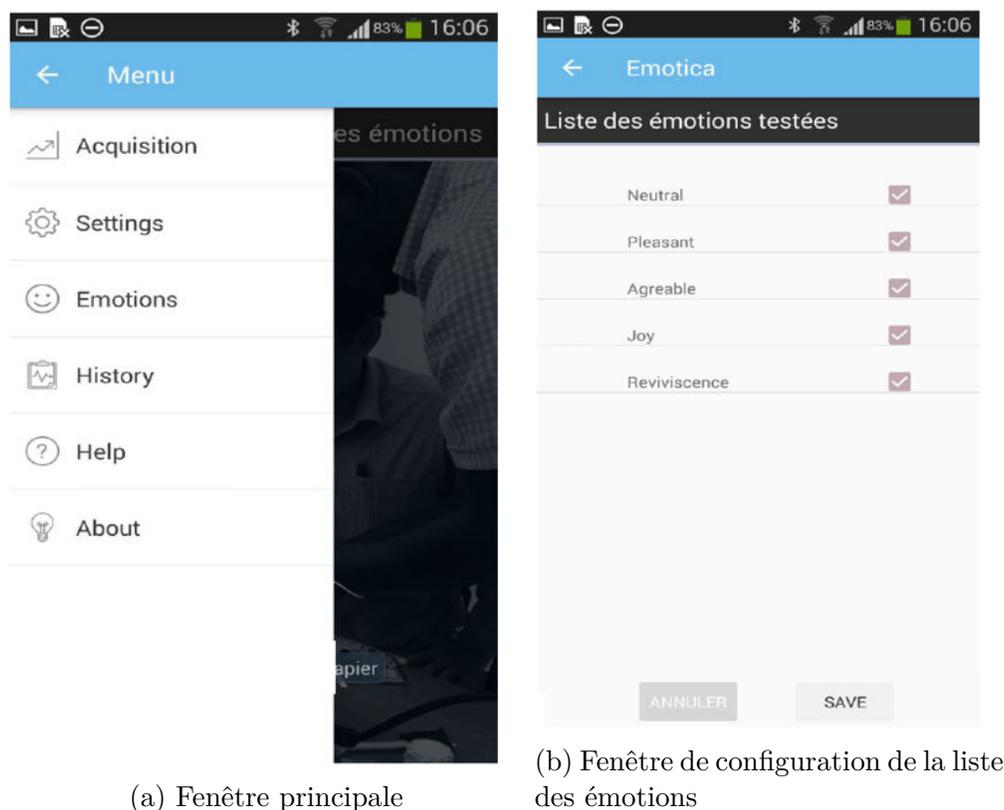


FIGURE 5.6 – Fenêtres principale et de configuration de la liste des émotions sur Android

Sur la figure 5.6b, le menu `Emotion` a été mis en place pour montrer à l'utilisateur les émotions actuellement détectables c'est à dire les émotions prises en compte lors de la création de la base d'apprentissage et qui sont donc disponibles. L'utilisateur a aussi la possibilité de sélectionner celles qu'il l'intéresse. Son choix pourra être sauvegardé.

Dans la conception de cette application, nous avons essayé de prendre en compte tous les paramètres qu'un utilisateur de ce système de détection aura besoin. L'une des fonctionnalités nécessité en premier lieu la création d'un nouveau profil utilisateur qui requiert de renseigner trois champs obligatoires : le genre, le nom de famille et le prénom. Cinq champs optionnels ont été ajoutés aux champs obligatoires. Ces différents champs optionnels renseignent l'âge, le métier, l'origine, les goûts musicaux (puisque nous avons choisi une stimulation de type musical qui sera expliqué dans la section suivante), les paramètres importants à renseigner.

Emotica étant plutôt destiné à une utilisation médicale, les fonctionnalités de celle-ci doivent être optimisées et proches des attentes de la profession. Les fonctionnalités déjà prises en compte dans la présente application sont donc :

- la recherche filtrée
- La recherche d'utilisateurs
- La sauvegarde des données d'une séance d'acquisition
- La création d'un nouveau profil utilisateur

5.3 Application visée et conception de notre propre base d'apprentissage de signaux physiologiques

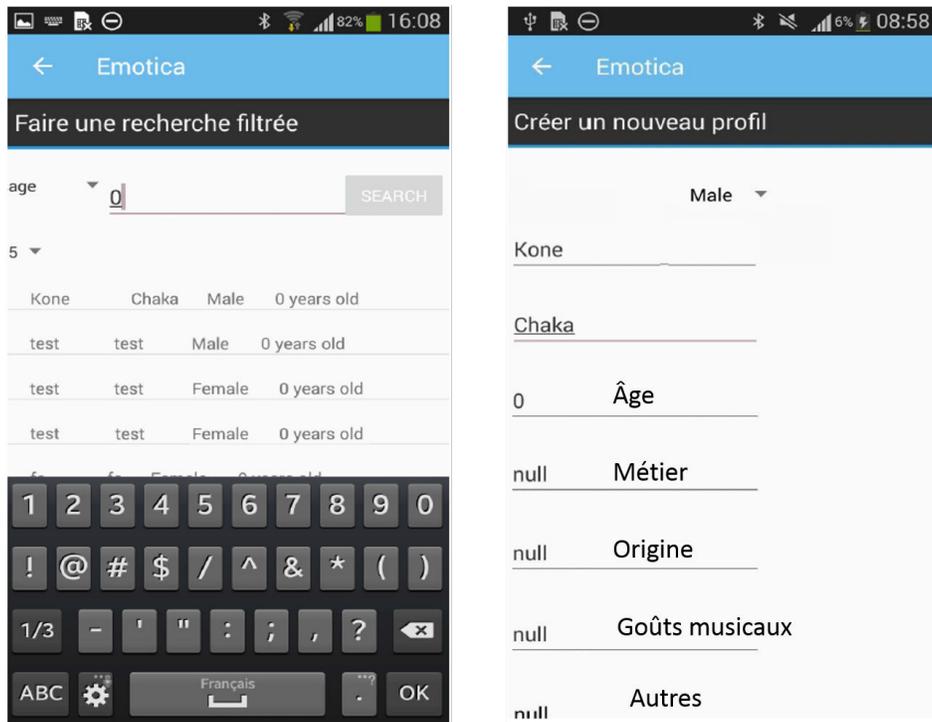


FIGURE 5.7 – Fenêtre Android permettant de créer un nouveau profil et de lancer une recherche

- La restitution des données d'un profil existant
- La modification des données d'un utilisateur

L'architecture matérielle et logicielle d'EMOTICA ayant été décrite, l'objectif est maintenant de présenter l'application pour laquelle nous avons adapté le système Emotica. Par cette application, nous décrirons le protocole d'induction des états émotionnels ainsi que la base de données conçue en utilisant ce protocole.

5.3 Application visée et conception de notre propre base d'apprentissage de signaux physiologiques

La validation de nos algorithmes de détection des émotions a été réalisée à partir de la base de données fournie par le MIT (voir chapitre 2). Cette base de données a été obtenue par des acteurs chargés de jouer un état émotionnel. Compte tenu de ce fait, il nous a paru opportun de créer notre propre base de données à titre comparatif dans un premier temps. Dans cette section, nous présentons la base de données que nous avons développée ainsi que le protocole utilisé pour l'induction des émotions. Afin de se placer dans contexte médical d'étude, nous avons souhaité appliqué la problématique des émotions à la mnémothérapie, par exemple pour des patients atteints de la maladie d'Alzheimer. Ce choix a été fait, suite à de nombreuses discussions avec

5.3 Application visée et conception de notre propre base d'apprentissage de signaux physiologiques

l'équipe médicale de la fondation Noisiez à Sophia Antipolis qui accueille de jours des patients atteints de cette maladie.

Application pour les patients atteints de la maladie d'Alzheimer

Le nombre de patients souffrant d'Alzheimer augmente d'année en année. En effet, en 2005, la maladie d'Alzheimer a touché 26 millions de personnes dans le monde et pourrait atteindre quatre fois plus en 2050, ce qui équivaldrait à 1/85 de la population mondiale [Brookmeyer et al., 2007]. Selon l'Organisation mondiale de la santé, en 2020, un quart des Français de plus de 65 ans seront touchés par la maladie d'Alzheimer. De nombreux centres d'accueil pour les personnes âgées n'acceptent que les patients atteints de démence légère et les patients atteints d'Alzheimer n'y sont accueillis que si leur état leur permet de participer aux activités organisées. Pour palier à ce problème et Vu le nombre croissant de patients, des centres spécialisés pour accueillir ces personnes ont été créés ces dernières années. Lors des premiers stades de la maladie, un grand nombre de famille essaie d'aider le patient, atteint de cette maladie, à rester le plus longtemps possible indépendant. Pour s'occuper de ces patients, les familles, en accord avec les médecins, peuvent utiliser des thérapies non médicamenteuses qui consistent entre autre à organiser des activités ludiques pour stimuler les patients. En effet, une approche non médicamenteuse est très importante dans le soutien de cette maladie et de nombreuses approches ont été proposées.

- La **stimulation cognitive** pour stabiliser les déficiences cognitives des personnes atteintes de la maladie d'Alzheimer. Ce type de stimulation peut ralentir la dégénérescence mémorielle et cognitive grâce aux pratiques d'exercices mentaux. La stimulation cognitive retarde donc l'entrée des patients dans une institution gériatrique.
- La **stimulation sensorielle** est utilisée pour stimuler les différents sens d'une personne. La Zoo-thérapie [Beiger, 2016], la Musicothérapie [Lecourt, 2014], l'Aroma-thérapie [Viau, 2009] sont toutes des thérapies de stimulation sensorielle.
 - Zoo-thérapie : Cette thérapie consiste à utiliser la proximité d'un animal domestique pour réduire le stress ou les conséquences d'un traitement médical et permet aussi de faire sortir les patients de leur état d'apathie.
 - Musico-thérapie : Elle consiste à placer la musique au cœur des thérapies de soutien et d'aide aux patients atteints de cette pathologie. La musicothérapie est composée de trois dimensions : la musicothérapie active est axée sur des productions sonores par le sujet/patient au moyen de sa voix ou d'utilisation d'instruments musicaux. Ceci amène à l'ouverture d'un moyen de communication non verbale avec des personnes en rupture de communication comme les patients d'Alzheimer. La musicothérapie réceptive, quant à elle, est axée sur l'écoute d'extraits musicaux ou sonores. Elle vise à amener le patient dans trois états à savoir l'apaisement, la détente et la relaxation. Enfin, la dernière dimension est celle de la détente psychomusicale. Cette dernière est

5.3 Application visée et conception de notre propre base d'apprentissage de signaux physiologiques

axée elle aussi sur trois phases successives à savoir : l'apaisement, la détente et la ré-dynamisation. Elle se base sur les principes de l'hypnose en remplaçant les suggestions verbales par des inductions musicales.

- Aroma-thérapie : Cette thérapie est l'utilisation des composés aromatiques extraits de plantes, comme dans les huiles essentielles, à des fins médicales.
- La **stimulation mémorielle** : La stimulation de la mémoire permet de conserver l'autonomie et de retarder la dépendance des patients. Différentes techniques sont utilisées en vue de stimuler la mémoire. La Mnémothérapie [MAM, 2017] est l'une de ces techniques. La mnémothérapie est une approche thérapeutique des symptômes de la maladie d'Alzheimer qui consiste à éveiller et faire s'épanouir de façon reproductible la reviviscence des mémoires rétrogrades par stimulation sensorielle indicée. Elle a pour but d'améliorer de façon immédiate et durable l'apathie, la tristesse, les syndromes dépressifs, l'anxiété et les troubles du comportement. A cause de l'amnésie antérograde qui caractérise la maladie d'Alzheimer, la Mnémothérapie est reproductible et permet de réactiver des mécanismes intellectuels et émotionnels du patient atteint de la maladie d'Alzheimer. La reviviscence des mémoires rétrogrades est vécue non pas comme un souvenir mais comme une sensation heureuse de revivre un épisode de sa vie. La mnémothérapie vise donc à faire sortir le patient Alzheimer de son état apathique à l'état de reviviscence.

Nos émotions sont sous un contrôle "inconscient" et quand nous perdons ce contrôle, grâce à une musique, à un geste, à une odeur, nous pouvons nous sentir submergés par une émotion. Par exemple, on peut parler de ce que Proust appelle «les intermittences du cœur» où il explique qu'il n'a été touché par la mort de sa grand-mère qu'après plusieurs jours. Il raconte avoir été submergé par ces émotions le jour où il est retourné dans un endroit qu'il avait fréquenté avec sa grand-mère. Ainsi, dans cet endroit, sa mémoire involontaire a supprimé la frontière entre le présent et le passé permettant ainsi de revivre un moment précis de sa vie avec sa grand-mère, c'est le phénomène de reviviscence.

Les thérapies non médicamenteuses ne sont actuellement pas accessibles pour la plupart des patients atteints de la maladie d'Alzheimer en raison du coût et du faible nombre de centres d'accueil. La problématique actuelle pour ces patients est que l'effet thérapeutique ne peut être automatiquement acquis en raison de la quasi-indisponibilité de système de détection des émotions sur le marché et de l'utilisation en général de capteurs physiologiques invasifs. Pour être acceptée par les patients, ce type de système doit être léger et ne pas constituer une gêne (le moins intrusif possible). Nous allons donc proposer un système de détection des émotions qui répond à ces attentes tout en restant performant en termes de taux de reconnaissance.

Le système Emotica pourrait être utilisé par les familles des patients d'Alzheimer (ou par les équipes médicales) pour améliorer leur qualité de vie. Dans un contexte de mnémothérapie musicale, Emotica présente des avantages, d'un point de vue organisationnel, comme l'obtention rapide d'un indicateur objectif, de l'effet de la reconnaissance d'une musique, sur l'émotion et la mémoire d'un sujet. De plus, afin

5.3 Application visée et conception de notre propre base d'apprentissage de signaux physiologiques

de s'adapter aux différents patients, la caractérisation d'une playlist spécifique au sujet pour une personnalisation de l'environnement sensoriel peut être réalisé aisément. D'un point de vue thérapeutique, EMOTICA présente aussi des avantages comme :

- Grâce à l'amnésie antérograde, on peut donner même à la maison une heure (ou plus) de joie chaque jour à un patient atteint de la maladie d'Alzheimer.
- La stimulation cognitive et la rééducation de la plasticité cérébrale. En faisant écouter des musiques aux patients leur permettant de revivre des instants de leur vie, on entretient le réseau par défaut. Le réseau par défaut peut être défini comme la réorganisation des souvenirs de la mémoire. L'entretien de ce réseau par défaut entraîne un ralentissement de la progression de la maladie.
- Le réapprentissage du bonheur et l'amélioration de l'engagement (incidence pour l'aidant) à travers la mémoire inconsciente. Certains gérontologues démontrent que, lorsque les patients atteints de la maladie d'Alzheimer commencent à venir dans un centre d'accueil de jour, ils sont tristes et leur participation aux activités est très limitée. Cependant tout au long du processus (c'est à dire venir au centre d'accueil la journée et repartir au domicile le soir) même s'ils ne savent pas pourquoi ils viennent dans le centre d'accueil, on peut remarquer qu'ils sont heureux d'être là, c'est ce que nous appelons la «mémoire inconsciente».

Pour répondre à ces attentes, le système EMOTICA a été adapté. Une des principales étapes est de constituer une nouvelle base de donnée qui contiendra entre autre la base d'apprentissage.

Conception de la base de signaux physiologiques

L'une des caractéristiques de la maladie d'Alzheimer est l'apathie qui est le fait de ne pas pouvoir rentrer dans le réseau par défaut (nous rentrons dans ce réseau quand tout à coup, de façon inconsciente, nous nous mettons à penser à divers épisodes de notre vie : ce réseau par défaut peut donc être appelé le réseau du rêve éveillé). Le réseau par défaut va dépérir non pas à cause de lésions mais à cause de sa non utilisation. Il est donc nécessaire de pouvoir rééduquer le réseau par défaut non pas en mettant un patient Alzheimer (ou une personne âgée) dans l'exécutif (c'est-à-dire en lui demandant de faire des exercices mémoriels) mais en utilisant une stimulation sensorielle. L'utilisation d'une musique indicée est un moyen de réaliser une stimulation sensorielle : la mnémothérapie permet donc d'aller stimuler la mémoire pour rentrer directement dans le réseau par défaut. Ceci permettrait, à un patient Alzheimer, de rêver avec la musique indicée permettant ainsi que la musique et le souvenir rentrent en résonance pour amplifier l'intensité de l'émotion ressentie jusqu'à l'amener à l'état de reviviscence.

Dans notre étude, nous avons choisi la mnémothérapie utilisant un stimulus musical permettant la reproduction à l'identique de l'empreinte mémorielle initiale et favorisant la mobilisation de la mémoire par l'émotion suscitée. Notre approche de stimulation peut être utilisée par toute personne saine ou pas, jeune ou âgée et a l'avantage d'activer, d'éduquer ou de rééduquer le réseau par défaut. Dans notre protocole

5.3 Application visée et conception de notre propre base d'apprentissage de signaux physiologiques

d'acquisition, nous avons utilisé les nœuds de capteurs préalablement présentés avec toute l'architecture matérielle/logicielle d'Emotica. Dans le cadre d'une collaboration, nous avons étroitement travaillé avec le gérontologue Jean Claude Broutart du centre d'accueil de jours des patients Alzheimer de la fondation GSF Jean Louis Noisiez pour définir le protocole d'induction des états émotionnels nécessaire pour la conception de la base de données. Dr. Broutart est l'initiateur de cette idée de stimulation mémorielle par les musique indicées. Avec sa participation, nous avons tout d'abord défini les émotions requises (**neutre, plaisir, désagréable, joie, reviviscence**) pour l'application visée. Un nouveau protocole permettant de susciter ces émotions sur les sujets a été défini. Les chercheurs des sciences de la mémoire ont démontré qu'entre 15 et 25 ans nous avons le pic de réminiscence. Ce pic est la période, pendant laquelle, nous gardons le plus de souvenir. En choisissant des musiques de cette période, il est fort probable de tomber sur des musiques liées à des événements précis de la vie d'un patient et ainsi de l'amener à un état de reviviscence.

Chaque patient, ou le plus souvent un accompagnant, construit avant de commencer les sessions d'acquisition une playlist musicale susceptible de lui procurer toutes ses émotions et surtout sélectionne celles qui peuvent l'amener dans cet état de reviviscence. Cette playlist est composée d'au moins 50 morceaux musicaux ; elle est conçue par le sujet et/ou sa famille, qui sélectionnent des musiques suivants ses goûts musicaux, son origine, et son pic de réminiscence. Au total, pour chaque sujet, une playlist d'environ 100 morceaux musicaux a été construit. Au début de chaque session d'acquisition, et surtout pour un nouveau sujet, le protocole d'acquisition est expliqué. Chaque session dure entre 25 et 30 minutes ce qui permet d'avoir une quantité considérable de signaux mais aussi de laisser la personne ressentir profondément chaque musique afin de susciter les vraies émotions. Les principales caractéristiques du déroulement d'une session d'acquisition en vue de la conception de la base de données sont :

- L'acquisition est faite dans une salle isolée afin d'éviter les bruits qui peuvent influencer les émotions/sentiments du sujet.
- Le sujet est assis sur une chaise dorsalée afin qu'il puisse être dans un état confortable.
- Les nœuds de capteurs sont placés sur le corps du sujet pour acquérir les données physiologiques. Le sujet dispose également de la manette lui permettant de donner un feedback sur son état émotionnel.
- A l'aide d'un casque, le sujet écoute les musiques de la playlist aléatoirement afin de garantir l'effet de surprise favorisant ainsi l'arrivée dans le réseau par défaut.
- Le sujet ferme les yeux pour mieux ressentir les émotions à travers la musique.

A noter que dans cette étape de conception de la base de données, le système portable n'a vraiment pas d'intérêt. Son intérêt réside dans le système final avec lequel, on peut acquérir et détecter nos états émotionnels n'importe où, n'importe quand.

Nous avons récoltés un ensemble de signaux physiologiques sur plusieurs mois en raison de la disponibilité de nos sujets et du fait qu'une partie aussi a été récoltée au Vietnam à Danang. Avec une fréquence d'échantillonnage de 100 Hz, nous avons construit une base de données sur 92 sessions d'acquisition avec trois signaux

physiologiques *ECG*, *EMG*, *EDA*. Ces sessions ont été faites sur un ensemble de 10 personnes dont 3 femmes et 7 hommes.

5.4 Résultats et discussions

Après avoir regroupé tous les signaux physiologiques collectés sur les 92 sessions d'acquisition, nous les avons scindés en données d'apprentissage, de validation et de test. Les mêmes techniques de validation des paramètres et de découpage des données en bases d'apprentissage, et de test choisies dans le chapitre 2 ont été utilisées pour valider les hyper-paramètres de nos classifieurs. Ainsi, 3253 vecteurs caractéristiques ont été extraits des signaux pour chaque émotion. Sur ces vecteurs caractéristiques, 1302 (40%) ont été utilisés comme base de test pour tester nos algorithmes et les 60% restants (1951 vecteurs caractéristiques) ont été utilisés comme base d'apprentissage et de validation. 60% de ces 1951 vecteurs caractéristiques, soit 1170 vecteurs caractéristiques constituent a base d'apprentissage et e781 pour la base de validation.

En résumé, les bases d'apprentissage, de validation et de test sont respectivement composés chacun de 5850, 3905 et 6510 vecteurs caractéristiques pour l'ensemble des 5 émotions.

Dans la phase de validation, les hyper-paramètres (le paramètre k pour le KNN et les paramètres c et γ pour le SVM), des classifieurs ont été validés. Dans l'algorithme du SVM, chaque couple de c et γ maximise le taux de reconnaissance d'une émotion donnée et minimise les taux de reconnaissance des autres émotions ce qui garantit finalement la possibilité de faire du multimodal avec les meilleurs résultats possibles. Les figures 5.8 et 5.9 montrent les résultats obtenus pour chaque classifieur implémenté dans le système Emotica.

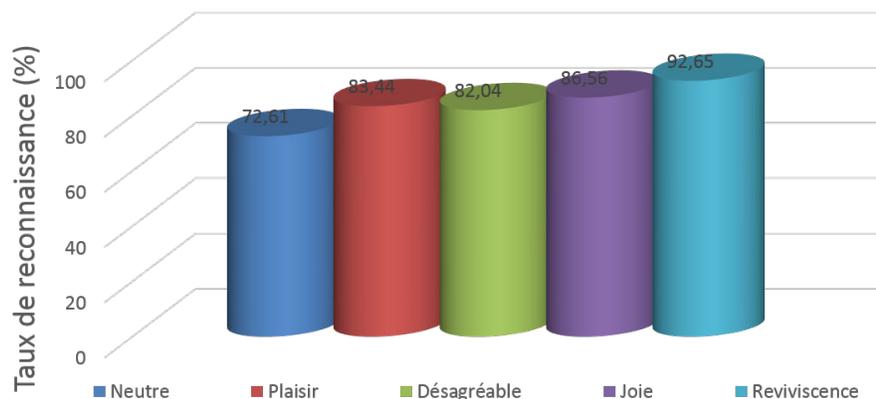


FIGURE 5.8 – Taux de reconnaissance multi-modal des émotions en utilisant l'algorithme du KNN sur notre propre base

5.4 Résultats et discussions

A la lumière des résultats présentés sur la figure 5.8, nous pouvons dire que nos algorithmes de détection des émotions fonctionnent correctement car, ils garantissent un taux de reconnaissance moyen de 83.45%. Chaque émotion est détectée avec un taux de reconnaissance des émotions satisfaisant dont le minimum est obtenu pour l'émotion "neutre" et le maximum pour l'émotion "reviviscence". Nous pouvons expliquer la chute du taux de reconnaissance pour l'émotion "neutre" à cause du fait que cette émotion est à valence neutre tandis que les autres sont à valence positive. Pour le KNN qui n'utilise qu'un seul paramètre (k), il peut être assez délicat de distinguer une seule émotion à valence différente de toutes les autres. Il faut noter que l'émotion la plus importante, lors de la détection pour la pathologie considérée, est la reviviscence et que c'est pour cette émotion que le taux de reconnaissance est maximal.

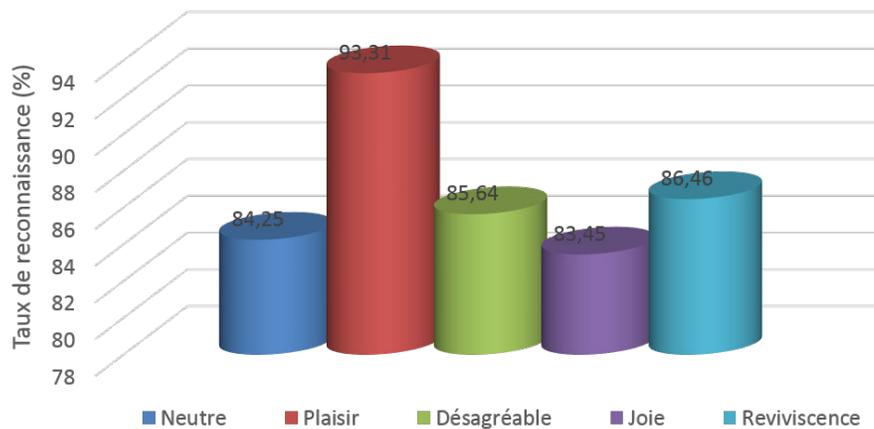


FIGURE 5.9 – Taux de reconnaissance multi-modal des émotions en utilisant l'algorithme du SVM sur notre propre base.

En analysant, la figure 5.9, nous remarquons que le classifieur Machine à Vecteur de Support (SVM) possède, comme le KNN, des taux de reconnaissance élevés. Avec le SVM, nous avons obtenu un taux de reconnaissance moyen de 86.62%. Nous pouvons donc dire que, sur ces données, l'algorithme du SVM est plus performant que celui du KNN. L'analyse des résultats des deux classifieurs émotion par émotion conduit à dire que le KNN est plus adapté pour les émotions "*Joie et Reviviscence*" alors que le SVM est plus adapté pour les émotions "*Neutre, Plaisir et Désagréable*".

Dans le but de comparer la base de données du MIT à celle que nous avons construite durant cette thèse, le tableau 5.1 résume leurs principales caractéristiques. Premièrement, notre base de données s'appuie sur une campagne de mesures plus importante que celle du MIT, ce qui permet d'améliorer les taux de reconnaissance. Deuxièmement, elle est plus générique et globale car ses signaux sont collectés sur plusieurs personnes et sans biais de simulation. Dans la base de donnée du MIT, les signaux sont acquis sur un seul acteur, qui simule certaines émotions, pendant une durée prédéterminée, à partir d'une stimulation liée à la présentation de photos et à l'écoute

	Base du MIT	Notre propre Base
Nombre de sessions	20	92
Nombre de signaux	4 (Resp, GSR, BVP, EMG)	3 (ECG, GSR, EMG)
Nombre d'émotions	8	5
Nombre de sujets	1 (acteur)	10
Type de stimulation	photos et mélodies	Mnémothérapie musicale
Expressions des émotions	simulées	réelles

TABLE 5.1 – Comparaison des deux bases de données

de mélodies. Contrairement à la méthode utilisée dans la conception de la base de données du MIT, dans notre protocole d'acquisition, un temps n'est pas prédéterminé pour exprimer telle ou telle émotion ce qui fait que nous avons des signaux liés à de vraies émotions. Enfin l'utilisation de la musique indicée (mnémothérapie musicale) permet de stimuler le sujet pour exprimer une émotion liée à un événement donné, à un moment de sa vie, ce qui favorise le ressenti par le sujet de vraies émotions donc favorise l'acquisition de vrais signaux.

Nos algorithmes de détection ont été pensés afin qu'ils soient simples puisque nous envisageons de les porter dans des dispositifs mobiles comme le smartphone ou de délocaliser certaines parties du traitement, par exemple, dans l'unité de traitement du système d'acquisition. La fréquence d'échantillonnage étant fixée à 100 Hz et étant donné que le système Emotica a besoin de 230 échantillons pour pouvoir détecter un pic, le temps minimal entre deux détections de pics dans un signal est de 2.3 secondes. Le temps de traitement moyen, qui représente l'exécution de l'application sur Matlab, pour filtrer et analyser le signal afin de détecter des pics, pour calculer et classifier le vecteur caractéristique du pic extrait, pour fusionner les différentes décisions émotionnelles et enfin pour afficher l'état émotionnel résultant de cette analyse est de **2.1482 secondes** pour l'algorithme des plus k plus proches voisins (kNN) et **1.7835 seconds** pour l'algorithme des séparateurs à vaste marge (SVM). Ainsi, nos algorithmes ne demandent respectivement pour le kNN et le SVM environ **4.4 secondes** et **4 secondes** pour acquérir les signaux physiologiques sur une personne et analyser ces signaux afin de détecter l'état émotionnel de la personne.

5.5 Conclusion

Ce chapitre est consacré à la validation du système Emotica dans sa globalité. Cette validation a nécessité de définir un protocole d'acquisition des signaux physiologiques afin de créer notre propre base de données dont une partie est liée à l'apprentissage. Suite à la définition de ce protocole, une campagne de mesures sur des personnes a été réalisée. Il faut noter que, de plus, nous nous sommes placés dans un contexte applicatif réel. Dans un premier temps, nous avons présenté l'architecture matérielle du système complet qui a permis de concevoir une nouvelle base de données de signaux

5.5 Conclusion

physiologiques liée aux émotions. Cette architecture matérielle est composée des nœuds de capteurs non invasifs ayant une faible consommation énergétique, d'une manette permettant l'annotation des signaux et facilitant ainsi l'apprentissage et la validation des paramètres des classifieurs. Afin de paramétrer l'application et d'analyser les résultats, une interface graphique a été développée sous Matlab. D'autre part, dans le but de proposer un système complet de détection des émotions pouvant être utilisé par toute catégorie de personnes, nous avons développé une version mobile. Ce travail nous a donc conduit à convertir les codes Emotica en Java-Android et ensuite à proposer une interface graphique pour ce type d'architecture cible. Par la suite, nous avons construit notre propre base de signaux physiologiques pour la détection des émotions. De ce fait, nous avons défini un ensemble d'émotions pour une application de "Mnémothérapie musicale" et proposé un protocole d'induction des émotions avec la fondation GSF qui est un centre d'accueil de jours des patients atteints de la maladie d'Alzheimer. Ce travail nous a permis de générer une base de données quatre fois plus grande que celle du MIT, et plus fiable car les émotions exprimées sont authentiques et non pas simulées. Enfin, en utilisant les deux algorithmes de classification implantés dans le système Emotica, nous avons déterminé les taux de reconnaissance sur notre base de données. Les algorithmes KNN et SVM donnent respectivement des taux de reconnaissance de 83.45% et 86.62%. Sur les données de la base du MIT ou sur celles de notre propre base de données, l'algorithme du machine à vecteurs de support (SVM) permet d'avoir le meilleur taux de reconnaissance. Cependant il faut noter que l'écart entre les deux demeure faible.

Conclusion générale et perspectives

Conclusion générale

Dans ces travaux de thèse, nous nous sommes focalisés sur la reconnaissance automatique des émotions en utilisant les signaux physiologiques. L'objectif était de proposer un système complet pour la détection des émotions garantissant des performances optimales en termes de taux de reconnaissance et de consommation. Pour y parvenir, il a été défini, dans un premier temps, des algorithmes fiables de détection des émotions, puis une méthodologie d'exploration d'architecture et de configuration d'objets communicants pour la santé. Cette méthodologie d'exploration permet de concevoir des systèmes performants en tenant compte de toutes les contraintes applicatives et utilisateurs. Enfin, l'ensemble du système conçu devait être validé dans un contexte applicatif d'e-santé.

Le premier chapitre de ce manuscrit est consacré à l'état de l'art sur les systèmes de détection des émotions et les algorithmes d'optimisations multicritères. Ce chapitre nous a, dans un premier temps, aidé à comprendre la notion d'émotion, ses modèles de représentation et d'annotation. Afin de définir le support matériel adéquat pour un détecteur d'émotion, différentes techniques existantes d'exploration d'architecture et d'optimisation de la consommation d'énergie ont été présentées. Sur la base de cette étude, nous avons positionné nos travaux par rapport à l'état de l'art et proposé un ensemble d'étapes afin d'atteindre les objectifs visés.

Fort de ces connaissances acquises suite à cette phase d'étude bibliographique, dans le deuxième chapitre, nous avons étudié tout d'abord des algorithmes de détection monomodale des émotions. Nous avons alors proposé des améliorations visant à réduire les temps de traitement tout en augmentant la robustesse et l'efficacité de ces algorithmes afin d'obtenir un taux de reconnaissance supérieur. Deux classifieurs, le "**k plus proches voisins (KNN)**" et le "**machine à vecteurs de support (SVM)**", ont été implémentés. Un accent particulier a été mis sur la validation des paramètres de ces différents classifieurs afin de définir un algorithme avec un fonctionnement optimal. Une comparaison rigoureuse de ces différents classifieurs et de leurs paramètres a ainsi pu être réalisée. Enfin, afin d'améliorer encore le taux de

reconnaissance, nous avons défini et implémenté un nouvel algorithme de détection multimodale en temps réel basé sur la méthode de fusion au niveau des décisions monomodales. Cette approche de détection multimodale a permis d'atteindre de très bons taux de reconnaissance des émotions. En effet, nous avons respectivement obtenu des taux de reconnaissance pour les classifieurs "**k plus proches voisins (KNN)**" et "**machine à vecteurs de support (SVM)**" de 81.69% et 84% pour les benchmarks sélectionnés.

Cette méthodologie de détection des émotions a ainsi été validée dans le chapitre 2 sur la base de données de signaux physiologiques fournie par le MIT. En vue de concevoir notre détecteur d'émotion, dans le chapitre 3, nous avons proposé un modèle analytique d'estimation de la consommation et de l'autonomie pour des objets communicants dédiés aux problématiques de la santé. Ce modèle analytique nous a, par la suite, permis d'étudier l'impact de certains paramètres sur les performances du système. A titre d'exemple, nous avons pu évaluer l'impact de certains capteurs et de la fréquence d'échantillonnage sur le taux de reconnaissance ainsi que l'impact du protocole de communication sur la consommation d'énergie. Concernant l'autonomie du système de détection, elle varie en fonction de l'architecture des capteurs et de leur configuration ; il est alors possible de rechercher le meilleur compromis entre les paramètres architecturaux et le taux de reconnaissance des émotions en fonction des exigences de l'application et de l'utilisateur. Dans le but de valider les modèles proposés, une étude comparative entre les valeurs estimées et mesurées de l'autonomie présente un faible écart, fait, qui démontre la fiabilité de l'approche proposée.

Dans le quatrième chapitre, un environnement (framework) de simulation et d'exploration d'architectures et de ses configurations a été développé en SystemC-TLM. L'innovation dans le modèle proposé est de prendre en compte à la fois les contraintes de performances, qui découlent des contraintes applicatives et utilisateur, et d'architectures, au plus haut niveau de la conception. Les contraintes de performances concernent le taux de reconnaissance des émotions et l'autonomie globale du système tandis que les contraintes architecturales concernent le choix des différents types de protocoles de communication (Bluetooth, Zigbee, Lora, ...), de modules de traitements (microcontrôleurs) et de capteurs. Ce framework, nous a ainsi permis de caractériser différentes architectures. Le cadre de conception, fourni grâce à cet environnement, permet d'étudier la synchronisation entre les différents nœuds capteurs, le taux de reconnaissance et les différences de consommation. Aussi, ce framework représente plus qu'un simulateur car il constitue une aide pour les concepteurs dans la définition d'un système d'acquisition par exemple de signaux physiologiques. Grâce à la prise en compte des contraintes applicatives et à la pertinence des modèles, les performances simulées seront proches de celles obtenues après conception.

Une fois ce cadre de conception développé, nous avons proposé dans le chapitre 5 une architecture logicielle et matérielle permettant à toute personne (même novice dans l'informatique ou l'électronique embarquée) de faire l'acquisition des signaux physiologiques, l'analyse de ces signaux ainsi que la détection des émotions dans les dits signaux. Afin de valider l'ensemble des travaux proposés, nous avons choisi de

nous placer dans un contexte applicatif médical : la mnémothérapie musicale pour les patients atteints de la maladie d’Alzheimer en lien avec le centre d’accueil de jour de la fondation Noisiez. L’objectif de cette thérapie non médicamenteuse est d’améliorer le quotidien des patients atteints de la maladie d’Alzheimer. Elle consiste à éveiller et faire s’épanouir de façon reproductible la reviviscence des mémoires rétrogrades par stimulation sensorielle indicée. Dans notre étude, nous avons choisi la mnémothérapie musicale utilisant un stimulus musical permettant la reproduction à l’identique de l’empreinte mémorielle initiale et favorise la mobilisation de la mémoire par l’émotion suscitée. Un premier travail fut de définir les émotions utiles à ce contexte applicatif, c’est à dire (*Neutre, plaisir, déplaisir, joie, reviviscence*), et d’élaborer une base de signaux physiologiques associée. Pour élaborer cette base de données dont une partie constitue la base d’apprentissage, nous avons défini un nouveau protocole d’induction des émotions et d’annotations appropriées des signaux recueillis en vue de l’exploitation ultérieure. Cette base des données réelle a été faite sur 10 personnes dont 3 femmes et 7 hommes et nous permet pour le moment d’avoir un taux de reconnaissance d’environ 86% pour l’algorithme lié au classifieur Support Vectors Machine (SVM). Il faut noter que la méthodologie proposée pour la création de la base de données se veut être indépendante de l’application visée. Dans le cadre d’une collaboration entre l’université de Nice et celle de Danang (Vietnam), une plus grande base de données est en-cours de conception à Danang avec notre protocole d’induction des états émotionnels.

Enfin, l’ensemble de nos nœuds de capteurs a été pensé de telle sorte qu’il ne heurte guère la sensibilité des sujets sur lesquels les signaux sont collectés. Un critère de non invasivité des nœuds était de rigueur et a été satisfait autant que possible. De plus, les nœuds fournis présentent une autonomie acceptable et le taux de reconnaissance est satisfaisant vis à vis don contexte applicatif.

Perspectives

Ces travaux de thèse ouvrent de nombreuses perspectives. Dans l’objectif d’une validation à plus large échelle, une première perspective, réalisable à moyen terme, consiste à étudier la personnalisation de la détection des émotions. En effet, l’analyse des signaux provenant des acquisitions en-cours à l’université de Danang et ceux des acquisitions faites à Nice sur des jeunes, des moins jeunes, des femmes et des hommes permettra de comprendre si les paramètres comme l’âge, le sexe, l’origine influencent la détection. Ainsi nous pourrions mettre en place une technique de détection multimodale qui sera plus adaptée à un spectre large de sujets. Nous pouvons aussi enrichir notre modèle de représentation des émotions en y ajoutant d’autres contextes applicatifs comme par exemple le *Big Five* (voir l’annexe C) qui permet de définir une personnalité. Pour répondre à cette problématique, nous avons proposé un modèle de classification des émotions en trois classes suivant leurs valences [Koné et al., 2015]. Dans cette approche, nous sommes partis des émotions pour remonter aux caractéristiques du Big Five. Nous savons que l’émotion est de durée limitée contrairement aux traits de caractère qui sont

quasiment inchangeables. Ainsi, une personne plutôt introvertie le restera. Il faut donc partir des signaux physiologiques eux-mêmes pour remonter aux traits de caractères de chaque personne donc à la définition d'un modèle de détection du Big-Five. Cette étude nécessite à ce stade des travaux supplémentaires notamment de validation.

La deuxième perspective de ce travail concerne l'utilisation d'autres algorithmes de classification comme les réseaux bayésiens naïfs et la méthode Fuzzy et de comparer les résultats de classification en termes de performance et de complexité. Il serait aussi intéressant d'étudier l'utilisation des techniques de réduction des données avec des filtres comme celles de kalman. Les techniques comme l'IPCA (Incremental Principal Component Analysis ou analyse en composante principale incrémentale) peuvent aussi être utilisés en vue de la réduction des données. Ainsi plusieurs améliorations pourraient donc être apportées à notre approche de reconnaissance des émotions.

La troisième perspective, consisterait à porter des améliorations à notre environnement de simulation/exploration d'architecture du système d'acquisition. Notamment, il serait nécessaire de réaliser une interconnexion plus rigoureuse entre Octave et SystemC. Cette interconnexion permettrait au noyau SystemC de charger/lancer Octave lors d'une simulation. Plus précisément, l'objectif serait d'améliorer la méthode de connexion que nous avons faite et qui consistait à créer un fichier accessible en écriture et lecture par Octave et SystemC afin de garantir le couplage. Ce couplage est indispensable car il permet de créer les événements nécessaires à la simulation mais aussi d'évaluer en parallèle la performance en tant que taux de reconnaissance des émotions. Dans le cadre de ces travaux, la recherche d'architecture s'est concentrée sur le système d'acquisition. On pourrait étendre cette simulation/exploration à un système complet d'acquisition et de traitement c'est à dire à créer aussi un dispositif de traitement embarqué respectant les contraintes applicatives et d'autonomie.

Enfin, nous proposons de tester notre système sur des patients atteints de la maladie d'Alzheimer dans les centres d'accueil. Nous essayons déjà de travailler dans ce sens avec la fondation Jean Luis Noisiez et la clinique de Mougins. Cette phase de travail est très compliquée car elle nécessite de présenter un protocole devant un comité d'éthique. Cependant il est prévu de commencer des sessions d'acquisition sur des personnes âgées saines dans la clinique de Mougins lors du premier semestre 2018.

Glossaire

Emotica : EMOTIon CAPture : Système développé au cours de cette thèse pour détecter l'état émotionnel d'une personne.

EMG : Activité Électro-MyoGraphique mesurant l'activité électrique des muscles.

ECG : Activité ÉlectroCardioGraphique mesurant le rythme cardiaque.

EDA : Activité ÉlectroDermale mesurant la sudation.

Framework : Outil de travail contenant des modèles, une méthodologie de conception et d'exploration.

Big Five : Ce modèle mis en évidence empiriquement par Goldberg en 1990 est basé sur 5 facteurs pour déterminer la personnalité d'une personne. Il permet de décrire et d'étudier théoriquement les traits de caractère qui sont :

- Extraversion : extériorisation des sentiments et des réactions.
- Névrosisme : sensation des émotions négatives.
- Extraversion : recherche de l'harmonie sociale.
- Ouverture à l'expérience : curiosité intellectuelle, ouverture à l'expérience.
- Agréabilité : organisation et persistance.

SNA : Le Système Nerveux Autonome (SNA) est chargé de la régulation automatique de certaines fonctions de l'organisme comme la digestion, la respiration, les muscles cardiaques . . . En effet, ces fonctions sont sous contrôle inconscient de l'être humain en d'autres termes, l'être humain ne peut contrôler ces fonctions de son organisme.

KNN : K-nearest neighbors (KNN) est une technique de classification basée sur le calcul de distance entre l'échantillon à classer et les échantillons existants dans une base de connaissance. Cette technique récupère donc les k échantillons les plus proches dans la base de connaissance de l'échantillon à classer.

SVM : Support Vector Machine (SVM) est une technique de classification binaire consistant à définir un hyperplan séparant les 2 classes tout en maximisant la distance entre la frontière de séparation et les vecteurs support qui sont les plus proches échantillons.

Conception de la base de données

Les besoins fonctionnels considérés lors de la conception de la base de données sont donnés ci-dessous. L'architecture de la base de données a été entièrement conçue en SQL (Structured query language). SQL est un langage informatique dédié à l'interrogation et au maintien des bases de données relationnelles.

1. **Utilisateur :**

- Créer un nouvel utilisateur (s'il n'existe pas déjà dans la base de données) avec : id unique, nom, prénom, âge, genre, origine, métier, goûts musicaux, détails
- Ajouter les données relatives à cet utilisateur (Les données : âge, origine, détails, métier et goûts musicaux ne seront pas obligatoires, car nous ne voulons pas forcer l'utilisateur à devoir donner des informations qu'il souhaite rester privées)
- Accéder à toutes les données de l'utilisateur
- Accéder à n'importe quel utilisateur de la base de données depuis son id
- Accéder à une liste complète de tous les utilisateurs
- Accéder au dernier utilisateur enregistré (peut être très utile car on souhaite effectuer en général plusieurs tests à la suite sur une même personne)
- Accéder à une liste d'utilisateurs du même âge, origine, métier, goûts musicaux, détails en commun, genre, nom de famille
- Supprimer un utilisateur en fonction de son id, ou supprimer le dernier utilisateur créé

2. **Emotion :**

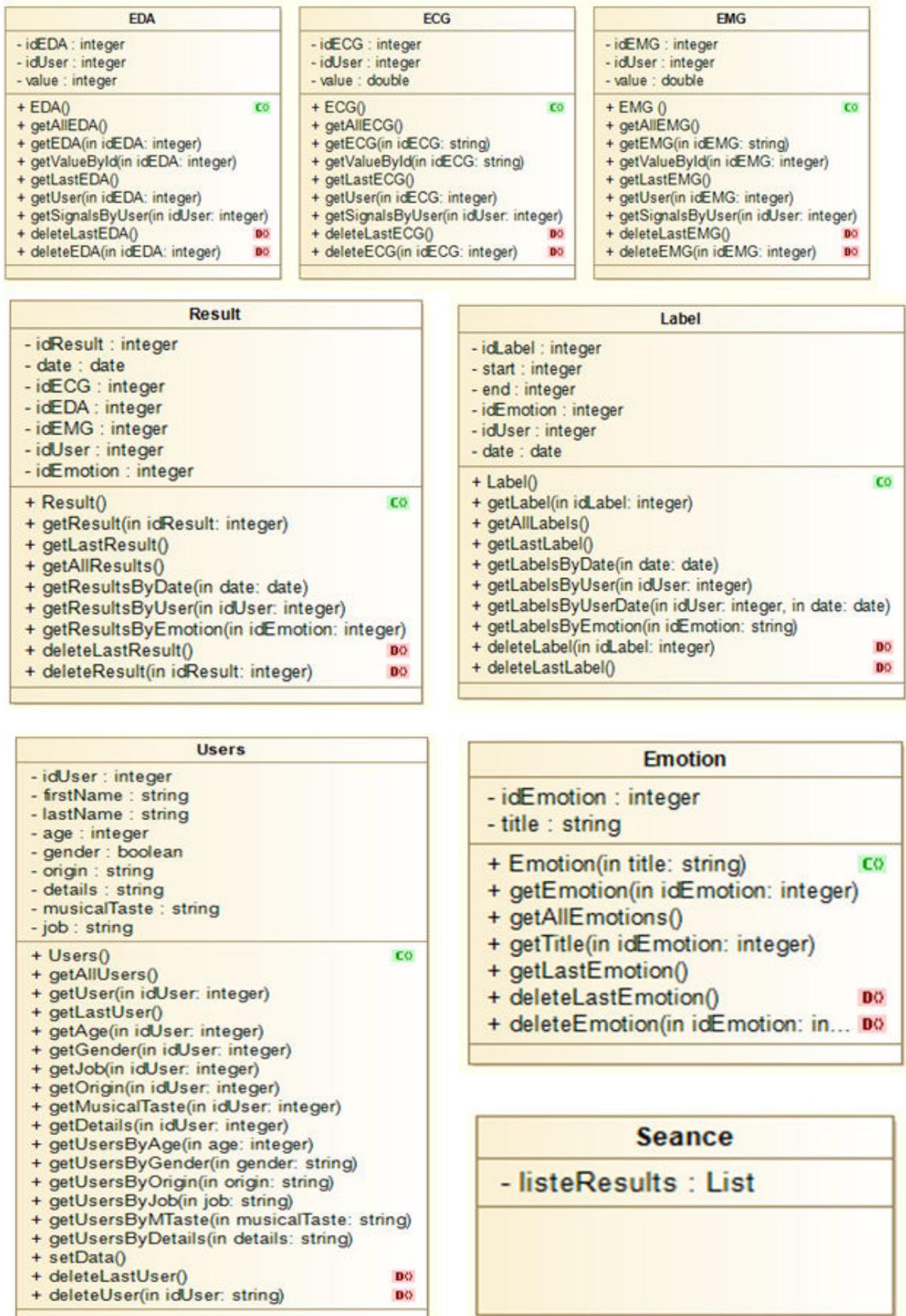
- Créer une nouvelle émotion dans la base de données avec un titre et un id unique
- Accéder à une émotion par son id
- Accéder au titre d'une émotion
- Accéder à une liste de toutes les émotions connues de la base de données
- Accéder à la dernière émotion créée
- Supprimer une émotion par son id ou supprimer la dernière émotion créée

3. **ECG, EMG, EDA (GSR) :** pour chaque table
 - Créer un nouveau signal avec : id, valeur, et l'id de l'utilisateur qui l'a réalisé
 - Accéder à un signal depuis son id
 - Accéder à la valeur d'un signal
 - Accéder à l'utilisateur qui a effectué le test
 - Accéder à une liste de tous les signaux
 - Accéder au dernier signal enregistré
 - Supprimer un signal par son id, supprimer le dernier signal
 - Accéder à une liste de signaux pour un utilisateur
4. **Résultat :**
 - Créer un nouveau résultat avec : id unique, date de création, l'id de l'utilisateur à qui
 - appartient le résultat, l'id de l'ECG, l'id de l'EDA, l'id de l'EMG, l'id de l'émotion trouvée
 - Accéder à une liste de tous les résultats
 - Accéder à une liste de tous les résultats pour une date donnée
 - Accéder au dernier résultat ajouté
 - Accéder à un résultat par son id
 - Accéder à une liste de résultats appartenant à un utilisateur (par son id)
 - Accéder à une liste de résultat pour une émotion choisie (par son id)
 - Supprimer un résultat par son id, supprimer le dernier résultat ajouté
5. **Label :**
 - Créer un nouveau label avec : idLabel, idUser, indiceDebut, indiceFin, idEmotion, date
 - Accéder à la liste de tous les labels
 - Accéder au dernier label enregistré
 - Accéder à un label par son id
 - Accéder à un label en fonction de son indice de début et de fin
 - Accéder à la liste de tous les labels pour un utilisateur choisi (par son id)
 - Accéder à la liste de tous les labels pour un utilisateur choisi (par son id) à une date précise
 - Accéder à la liste de tous les labels selon une émotion choisie (par son id)
 - Accéder à la liste de tous les labels pour une date précise
 - Supprimer un label par son id, supprimer le dernier label enregistré

La figure [B.1](#) montre le digramme de classes de la base données créée permettant de stocker en temps réel l'ensemble des données récoltées lors d'une session d'acquisition.

Les deux classes Label et Séance ont été ajoutées à l'architecture de la base de données dans l'intérêt d'une meilleure exploitation des données pour la réutilisation à l'échelle médicale.

La table label a été ajoutée pour permettre la création de patterns définis allant d'un indice de début à un indice de fin et reliés aux valeurs des signaux associés. Elle sert à pouvoir extraire une nouvelle base de connaissances utilisable pour la



145
 FIGURE B.1 – Diagramme de classes de la base de données

comparaison de signaux la plus fiable possible. Ayant associé à la fois un utilisateur, une séquence de signal (entre indice de début et de fin), l'émotion détectée et les signaux, cela va permettre lors de l'extraction d'une nouvelle base de données de nos connaissances, d'établir une comparaison au plus proche possible. Lors d'un test futur, il sera possible d'analyser pattern par pattern les différences obtenues avec un profil similaire à l'utilisateur testé, dans un laps de temps identique et à l'écoute d'une même chanson. Il s'agit donc de rendre plus fiable la portée du logiciel.

La table Séance a pour but le stockage brut d'une séance complète et permet de regrouper l'ensemble des résultats obtenus au cours de la séance d'acquisition sous forme d'une liste des idResultats récoltés. Cette table constitue une source d'information importante et pourra être réutilisée pour de l'extraction de données.

A chaque lancement de détection, une émotion est recueillie, et un enregistrement est ajouté automatiquement à la table Result. Il contient l'émotion obtenue associée à l'utilisateur testé (par son idUser), les signaux ECG, EMG, EDA recueillis dans le laps de temps séparant la détection actuelle de la précédente (également par leur id), ainsi qu'un champ Date qui indiquera la date et l'heure de l'enregistrement.

Le web service

Il y a de multiples utilités au serveur web. Il peut permettre de stocker un ensemble de données de manière distante, de disposer d'un grand espace de stockage et de posséder une puissance de calcul plus importante dont un utilisateur local n'a pas. Il peut aussi faire profiter plusieurs utilisateurs en même temps.



FIGURE B.2 – Web service

Le service web permet la distribution et l'échange rapide de données entre applications distantes (client - serveur) par l'intermédiaire d'internet. Il est indépendant de toute plateforme d'exécution et procède selon les standards XML et HTTP. Entre autre, il permet d'assurer la communication entre la base de données et l'application mobile.

Les éléments techniques sous Android Studio

1. **ListAdapter** : Android Studio prévoit un "adapter" pour les éléments de type ListView, sous le nom de «ListAdapter». Il s'agit d'une extension du widget

Adapter (également décliné pour les ArrayList ou encore les Spinner). Il permet de faire la liaison entre la liste des données et chaque donnée en elle-même. Mais en général, il est nécessaire de compléter cet adapter et de le personnaliser en fonction des besoins.

Dans le cas d'Emotica, chaque entrée de la liste devra toujours afficher le nom de famille, le prénom, le genre et l'âge et l'origine du patient. Pour cela, nous avons créé un layout représentant seulement un item de la liste. On le nomme en général *'list_item'* et il doit être situé dans l'arborescence au niveau du répertoire «res», dans le sous-répertoire de ressources nommé «layout». Dans notre cas, il doit être composé d'autant de TextView que de champs souhaités (nom, prénom, genre, âge, origine).

```
ListUsersAdapter adapter = new ListUsersAdapter(this.getActivity(), usersList);  
list_users.setAdapter(adapter);  
adapter.addListener(this);
```

Le but de l'Adapter sera justement de relier l'affichage d'un item, à celui de la liste entière. Pour ce faire, on crée une classe ListAdapter et on l'associe à un «Listener». La création du ListAdapter est lancée après avoir récupéré le tableau de résultats souhaité. C'est aussi l'un des avantages de cette méthode, car nous récupérerons les données depuis la base de données et la liste ne se crée que lorsque nous l'appelons à la création de l'adapter. Nous pouvons donc contrôler son affichage et le programmer uniquement à la fin de notre recherche.

2. **Tâches asynchrones** : Dans le but d'effectuer des liaisons entre l'application et la base de données en temps réel sans bloquer l'application et de gagner du temps durant l'exécution, nous avons créé des tâches asynchrones. Nous avons préféré les tâches asynchrones à la création de threads. En effet, par définition, un Thread permet l'exécution d'une tâche de manière parallèle au Thread principal. La tâche asynchrone, consiste en la réalisation d'une tâche de manière parallèle au Thread principal mais sans créer de nouveau Thread. On dit qu'elle résulte de l'utilisation d'un Thread unique. La figure B.3 permet de montrer le processus réel derrière le procédé des tâches asynchrones.

On peut voir que la notion de Thread unique est à nuancer. Aucun Thread n'est créé lors de l'exécution, néanmoins dès que l'on exécute le travail en fond un nouveau Thread est obligatoirement créé. Sa ligne de vie prend fin dès qu'il retourne un résultat et termine son exécution.

La classe abstraite AsyncTask<Params, Progress, Result> possède un ensemble de méthodes conçues pour faciliter la répartition du code («doInBackground()» pour les fonctionnalités que l'on veut exécuter en arrière-plan de l'application, «onPreExecute()» pour le code à réaliser avant l'exécution de l'asynchrone, «onProgressUpdate()» pour rafraîchir l'exécution à chaque progression, et enfin «onPostExecute()» après exécution). Puisqu'elles sont lancées sur le Thread principal, on ne peut exécuter deux tâches asynchrones en même temps. Il faut

ANNEXE B. CONCEPTION DE LA BASE DE DONNÉES

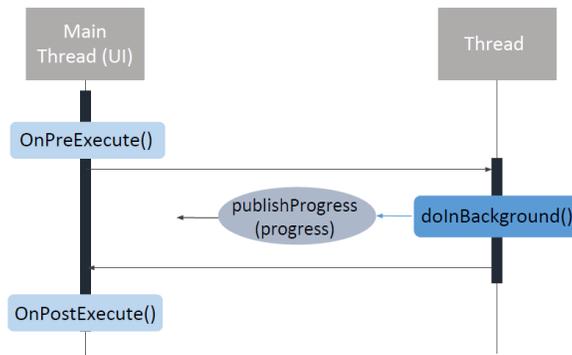


FIGURE B.3 – Processus réel derrière le procédé des tâches asynchrones

donc prévoir de les utiliser pour des tâches de courte durée, dont l'issue est connue.

Le tableau B.1 récapitule les principales différences entre Thread et tâches asynchrones.

Async task	Thread
Utilisation facile (simple constructeur avec passage de paramètres et méthodes préconçues à implémenter)	Utilisation couplée à un Runnable et Handler
Exécuté sur un seul Thread	Création de plusieurs Thread
Une seule tâche asynchrone à la fois	Plusieurs Thread exécutés en même temps (Thread.join() pour attendre la fin d'un Thread)

TABLE B.1 – Différences entre Thread et tâches asynchrones

3. **Spinner** : Un Spinner est une liste déroulante intégrée à un Layout dans notre application. Ils sont utilisés dans le besoin de faire sélectionner à l'utilisateur un attribut particulier parmi une liste plus ou moins longue. Le remplissage d'un Spinner peut se faire de manières différentes mais il implique dans tous les cas l'utilisation d'un ArrayAdapter. L'ArrayAdapter est produit par Android Studio, il s'agit d'un type permettant l'adaptation d'un Array déjà créé à l'intérieur d'un Spinner (ou une ListView, etc) pour un Layout donné. Son utilisation va dépendre de la méthode utilisée B.4, soit :

- Création d'un tableau directement dans la classe (1)

- Création d'une ressource «Array» dans le fichier XML «Strings» contenu dans le répertoire de ressources du projet (2)

```
1. private String[] categories = {"last name", "first name",  
    "gender", "age", "origin",  
    "job", "musical taste", "details"};  
  
2. <string-array name="cat_filter_array">  
    <item>last name</item>  
    <item>first name</item>  
    <item>gender</item>  
    <item>age</item>  
    <item>origin</item>  
    <item>job</item>  
    <item>musical tastes</item>  
    <item>details</item>  
</string-array>
```

FIGURE B.4 – Méthode de création d'un spinner

La différence principale entre ces 2 méthodes est la flexibilité. La première méthode propose plus d'accommodations que la deuxième car elle peut être à la fois un tableau écrit et défini à la main dans la classe, un tableau dont les données ont été récupérées depuis une base de données, ou encore un tableau dont les données ont été passées de classe en classe et récupérées depuis un autre tableau existant. A l'inverse, une ressource doit être écrite en brute dans le fichier XML prévu à cet effet et ne peut être modifiée de manière externe.

L'avantage de cette seconde méthode est qu'elle protège la liste créée de modifications non attendues d'ordre logicielle ou encore humaine. C'est pourquoi, nous l'avons choisi pour cette application, car le projet étant réalisé sur plusieurs années et probablement repris par plusieurs équipes de travail, il nous a semblé important de rendre la liste fixe dans un fichier de ressource afin qu'elle ne soit pas modifiée inutilement.

Au niveau de l'adaptation, il faut également noter que les commandes utilisées ne sont pas identiques, la deuxième méthode implique par exemple l'utilisation de la méthode «CreateFromResource», qui indique que l'adaptation se fait depuis un élément de ressource existant.

De même que pour l'affichage de la ListView précédente, la liste contenue dans le Spinner est entièrement personnalisable. Pour ce faire il faut créer un fichier XML que l'on pourra nommer *spinner_item*, décrivant le contenu d'un item de la liste. Dans notre cas un unique TextView. L'avantage de séparer les items ne se limite pas à l'Adapter mais également à la définition du style. Le texte de la liste n'est pas directement personnalisable depuis le type Spinner du Layout. En créant un item TextView séparé, on peut la styliser.

Après avoir rempli le Spinner, on lui attribue un «Listener», Android Studio produit une méthode prévue à cet effet «SetOnItemClickListener» qui associée au Spinner permet immédiatement d'y insérer le code d'exécution lors d'un clique de l'utilisateur.

Développement d'Emotica sur Java-Android

Pour convertir les codes Emotica développés du Matlab en Java-Android, nous devons d'abord exécuter sur PC Emotica dans un environnement Java pour l'étape de teste. Sur Android il est donc impossible d'utiliser les méthodes natives de MATLAB dans ces contextes pour plusieurs raisons. La première est celle de la licence trop onéreuse de MATLAB. La seconde raison est qu'il est très difficile et coûteux en calculs de faire communiquer les deux langages. Enfin, la bibliothèque Matlab est très gourmande en mémoire, un aspect non négligeable sur mobile. Il fut donc décidé de développer une bibliothèque mathématique en JAVA possédant la même API que la bibliothèque native Matlab qui en plus d'être un langage orienté objet a la faculté d'être un langage hautement portable autrement dit peut être exécuté sous tous les systèmes d'exploitation principaux. Ceci permettrait l'implémentation d'Emotica sur plusieurs systèmes d'exploitations PC, rendant la réutilisation du code très bonne. L'ensemble de ces méthodes mathématiques ont été développées dans une classe que nous avons appelé "Octave".

1. L'ensemble des méthodes développées dans la classe octave en java puis converties en Android :

```
+ <static> BlockRealMatrix : zeros(int, int)
+ <static> BlockRealMatrix : zeros(int)
+ <static> ArrayRealVector : zerosV(int)
+ <static> BlockRealMatrix : ones(int, int)
+ <static> BlockRealMatrix : ones(int)
+ <static> ArrayRealVector : onesAr(int)
+ <static> ArrayRealVector : min(ArrayRealVector)
+ <static> ArrayRealVector : min(BlockRealMatrix)
+ <static> ArrayRealVector : min(ArrayRealVector)
+ <static> ArrayRealVector : min(BlockRealMatrix)
+ <static> ArrayRealVector : max(BlockRealMatrix)
+ <static> ArrayRealVector : max(ArrayRealVector)
+ <static> ArrayRealVector : sum(BlockRealMatrix)
+ <static> ArrayRealVector : sum(BlockRealMatrix, int )
+ <static> double : sum(ArrayRealVector)
+ <static> double : sum(ArrayRealVector, int)
+ <static> double[] : sumtab(ArrayRealVector)
+ <static> ArrayRealVector : floor(ArrayRealVector)
+ <static> double : conv(double[], double[], int, int, int)
+ <static> ArrayRealVector : conv(double[] , double[] )
+ <static> ArrayRealVector : conv(ArrayRealVector, ArrayRealVector)
+ <static> BlockRealMatrix : repmat(ArrayRealVector, int)
+ <static> BlockRealMatrix : repmat(ArrayRealVector, int, int)
+ <static> ArrayRealVector : sort(ArrayRealVector, String) {
+ <static> BlockRealMatrix : sort2(ArrayRealVector, String)
+ <static> ArrayRealVector : hanning(int )
+ <static> double : mean(ArrayRealVector)
+ <static> ArrayRealVector : mean(BlockRealMatrix)
+ <static> double : var (ArrayRealVector)
+ <static> BlockRealMatrix : scalarDivide(BlockRealMatrix, double )
```

```

+ <static> ArrayRealVector : scalarDivideV(ArrayRealVector, double )
+ <static> BlockRealMatrix : matrixMultiply(BlockRealMatrix, BlockRealMatrix)
+ <static> BlockRealMatrix : matrixSquared(BlockRealMatrix)
+ <static> BlockRealMatrix : matrixSquaredPointByPoint(BlockRealMatrix)
+ <static> BlockRealMatrix : appendVectorToMatrix(BlockRealMatrix, ArrayRealVector)
+ <static> BlockRealMatrix : appendVectorToMatrix(BlockRealMatrix, double[])
+ <static> ArrayRealVector : unique(ArrayRealVector)
+ <static> ArrayRealVector : histc(ArrayRealVector, ArrayRealVector)
+ <static> double[] : Scanner(String)
+ <static> double[][] : Scanner(String, String)

```

2. **L'ensemble des méthodes développées dans la classe Emotica en java puis converties en Android** : La classe Emotica est composée des méthodes développées en matlab converties en java pour l'analyse des signaux physiologiques afin de détecter l'état émotionnel d'une personne en temps réel.

```

+ <static> ArrayRealVector : symExt(ArrayRealVector, int)
+ <static> ArrayRealVector : convertToRZ(ArrayRealVector)
+ <static> ArrayRealVector : computeGradient(ArrayRealVector)
+ <static> ArrayRealVector : smooth(ArrayRealVector, int , int)
+ <static> ArrayRealVector : cut(ArrayRealVector, int, StringBuilder)
+ <static> Map < BlockRealMatrix, BlockRealMatrix > : determineEmotion(BlockRealMatrix, int , int)
+ <static> Map < String, BlockRealMatrix > : kNearestNeighbors(BlockRealMatrix, BlockRealMatrix, int )
+ <static> String[] : getSignalNames(Map < String, ArrayRealVector >)
+ <static> Map < String, BlockRealMatrix > : extractionPattern(Map <String,ArrayRealVector>,
    Map<String,Float>,int)
+ <static> BlockRealMatrix : getCharacteristicsVector(ArrayRealVector,float,int,ArrayRealVector,
    ArrayRealVector,int)
+ <static> int : relevantParametersNumber(int[])
+ <static> ArrayList : finalTreatment(Map < String, BlockRealMatrix >,Map <String, ArrayRealVector >,
    int)
+ <static> boolean : isEmpty(String[], Map < String, BlockRealMatrix >)
+ <static> ArrayList<BlockRealMatrix> : analyzeCharacteristics(BlockRealMatrix, BlockRealMatrix,
    int[], int)
+ <static> ArrayRealVector : numberOfMaximums(BlockRealMatrix, int)
- <static> String[] : getResultNames(Map < String, BlockRealMatrix >)
+ <static> ArrayRealVector[] : computeEmotionsProbability(Map<String, BlockRealMatrix>,
    Map<String, ArrayRealVector>)
+ <static> ArrayRealVector : maximumEmotionCharacteristic(BlockRealMatrix, int)

```

B.1 Programmation du microcontrôleur

La figure B.5 montre la machine d'états que nous avons conçu pour le microcontrôleur de notre noeud de capteurs.

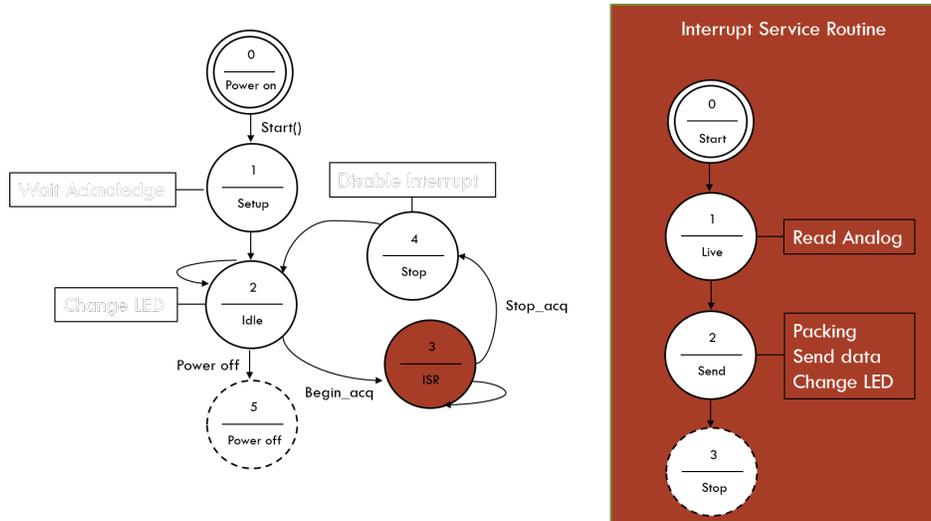


FIGURE B.5 – Machine d'états d'un noeud de capteur

Nous avons conçu la machine d'états de manière à ce qu'elle soit la plus générique et la plus économe en cycle d'exécution. Pour, non seulement, être économe en énergie, mais aussi pouvoir facilement l'adapter aux différents capteurs exploités par notre réseau de capteurs. Pour ce faire, nous avons décidé d'employer l'IDE Atmel Studio 7 car il inclut les GCC C et C++ compilateurs, assembleur, un simulateur, les interfaces intégrées avec des débogueurs et des programmeurs internes. Ce logiciel est particulièrement recommandé pour l'exploitation des MCU du type ATMega328P qui sont conçus par la même société qui a développé cette IDE.

La figure B.6 montre l'arbre des dépendances des différentes fonctions utilisées dans le code MCU associé à la machine d'états proposée. La synchronisation est assurée grâce aux timers que nous utilisons.

B.1 Programmation du microcontrôleur

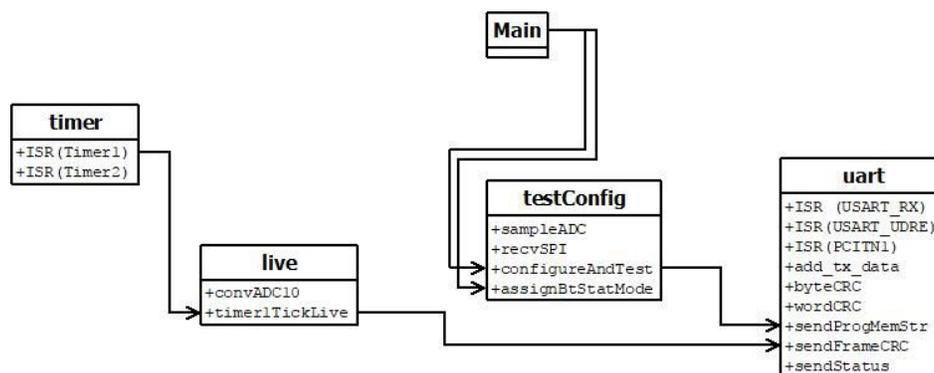


FIGURE B.6 – Arbre des dépendances du code MCU

Classification des émotions et Big five

Les 5 traits du big five permettent d'évaluer 5 fois différemment une personne mais ne permet pas une classification en 5 catégories. On peut déterminer que la personne est plus ou moins extravertie (E); et cela sans donner de jugement sur le fait qu'elle soit agréable ou non (A); et indépendamment de ces aspects relationnels, la personne peut être d'humeur plus ou moins égale (inverse du neuroticisme, N); et tout cela, sans juger de l'ouverture à la nouveauté, aux possibilités (O); et sans que tout ceci dise si elle est consciencieuse ou non (C), qui est encore un autre aspect de la personne. Chaque facteur donc du big five permet d'évaluer une personne indépendamment des autres facteurs.

C.0.1 Modèle de combinaisons des émotions

(B)= (No emotion, anger, hate, grief, love, romantic love, joy and reverence)

$$(e) = \sum_{i=1}^8 \langle E, u_i \rangle u_i \tag{C.1}$$

$$(e) = \alpha_1 * NoEmotion + \alpha_2 * Anger + \dots + \alpha_8 * Reverence \tag{C.2}$$

$$(e) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \alpha_8 \end{pmatrix}_B E_{grief} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha_4 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}_B E_{joy} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha_7 \\ 0 \end{pmatrix}_B \quad \alpha_4 \neq 0, \alpha_7 \neq 0 \quad E_{combinaison} =$$

$$E_{grief} \oplus E_{joy} = (0 \ 0 \ 0 \ \alpha_4 \ 0 \ 0 \ \alpha_7 \ 0)_B$$

C.0.2 Comment classifier les émotions en trois classes suivant leurs valences

Neutral emotions : No emotion, Reverence

Positive emotions : Love, Romantic love, Joy

Negative emotions : Anger, Hate, Grief

$$\begin{aligned} \text{Neutral}_{Emotions} &= E_{NoEmotion} \oplus E_{Reverence} \\ \text{Positive}_{Emotions} &= E_{Love} \oplus E_{Romantic} \oplus E_{Joy} \\ \text{Negative}_{Emotions} &= E_{Anger} \oplus E_{Hate} \oplus E_{Grief} \end{aligned}$$

C.0.3 Interprétation du Big five

$$\begin{aligned} E &= E_{joy} \oplus E_{love} \oplus E_{Romantic} \\ A &= E_{Joy} \oplus E_{Reverence} \\ C &= E_{NoEmotion} \oplus E_{Reverence} \oplus E_{Grief} \\ N &= E_{Anger} \oplus E_{Grief} \oplus E_{Hate} \\ O &= E_{Joy} \oplus E_{Grief} \end{aligned}$$

$$P = \begin{pmatrix} p_1 \\ \cdot \\ \cdot \\ \cdot \\ p_5 \end{pmatrix} \text{ avec } p_i \in [-1, 1]$$

$p_i=1$ présence forte du facteur positif

$p_i=0$ Absence du facteur

$p_i=-1$ présence forte du facteur négatif

$$P = \begin{pmatrix} 0.25 \\ 0.5 \\ -0.5 \\ 0.75 \\ 0.25 \end{pmatrix} \text{ peu extraverti, assez agréable, assez peu conscient, très stable}$$

émotionnellement et peu ouvert

$$P = \begin{pmatrix} -0.33 \\ -0.33 \\ 0.33 \\ 0 \\ -0.66 \end{pmatrix} \text{ introvertie, désagréable, peu conscient, et très fermé}$$

Annexe **E**

Comparaison de différents microcontrôleurs

Le tableau [E.1](#) donne les courants consommés par les différents microcontrôleurs considérés dans notre étude. Ces courants extraits des datasheets des mcus sont donnés suivant l'état dans lequel se trouve les microcontrôleurs.

MCUs	Attiny85	Attiny841	Atmega128	Atmega328p	STM32L476	PIC16F873A	PIC18F45Q10
Mémoire (Octets)	512	512	4096	2048	98304	192	192
Sleep mode current (mA)	0,01	0,004	0,025	0,84	0,0014	0,02	0,0002
PinInput current (mA)	0,001	0,001	0,00005	0,1	0,005	0,001	0,001
Active mode current (mA)	3	2,4	9,5	11,05	0,312	1,6	11,05

TABLE E.1 – Taille de la mémoire et courants consommés par les MCUs à 8 M Hz

Le tableau [E.2](#) permet de faire une comparaison entre quelques microcontrôleurs que nous avons modélisés dans le chapitre [4](#).

ANNEXE E. COMPARAISON DE DIFFÉRENTS MICROCONTRÔLEURS

Microcontrôleurs			Atmega328p	STM32L476	PIC18F45Q10
Architecture du CPU			AVR	ARM Cortex M4	PIC18 Core
Mémoire			32KB Flash, 1KB EEPROM, 2KB SRAM	Up to 1 MB Flash, 128 KB SRAM	32KB Flash, 256Bytes EEPROM, 2KB SRAM
Fréquence de fonctionnement			Up to 20 MHz, 1 MIPS/MHz	Up to 80 MHz, 1.25 DMIPS/MHz	64 MHz, 16 MIPS
Comparaison Énergétique $T = 85^{\circ}\text{C}$	Power supply current	Active @1MHz, $V_{cc}=2\text{V}$	Typ. 0.3 mA - Max. 0.5 mA	39 uA @3.3V	N/A
		Idle @1MHz, $V_{cc}=2\text{V}$	Typ. 0.04 mA Max. 0.5 mA	420 nA with RTC	N/A
	Power-save mode		0.8 uA (32kHz TOSC enabled, $V_{CC} = 1.8\text{V}$)	1.1 uA 1.4 uA (w/ RTC)	N/A
	Power-down mode		(WDT1 enabled/WDT disabled) Typ. 4.2/0.1 uA Max. 15/2 uA	3 0 nA	N/A
	Résolution (Nombre de bits)			8	32
Nombre de timers			2 (8 bits, 1 (16 bits)	2 (16 bits advanced motor control timers), 2 ULP timers, 7 (16 bits), 2 (32 bits)	3 (8 bits), 4 (16 bits)
Protocole de communication série			2 x SPI (Master/S), 1 x USART, 1 x I2C	USB OTG, 2x SAIs (serial audio interface), 3x I2C SMBus/PMBus, 5x USARTs (ISO 7816, LIN, IrDA, modem), 1x LPUART, 3x SPIs (4x SPIs with the Quad SPI) CAN SDMMC SWPMI single wire protocol master I/F IRTIM (Infrared interface)	2 x UART, 2 x SPI, 2 x I2C
Nombre d'entrées analogiques			8 158	10	35
Nombre d'entrées analogiques			23	11	5

TABLE E.2 – comparaison des microcontrôleurs

Personnalisation de la détection des émotions

Après avoir étudié l'état de l'art sur les relations entre les émotions et l'âge et la culture, nous avons voulu en utilisant des langages du web sémantique (RDF, Sparql) faire des requêtes pour charger les ontologies de Profession, le pays d'origine et les différents groupes ethniques. Selon l'état de l'art, ces différentes caractéristiques (profession, origine ...) influencent l'expression des émotions. De ce fait nous avons établi un modèle décrit sur la figure F.1.

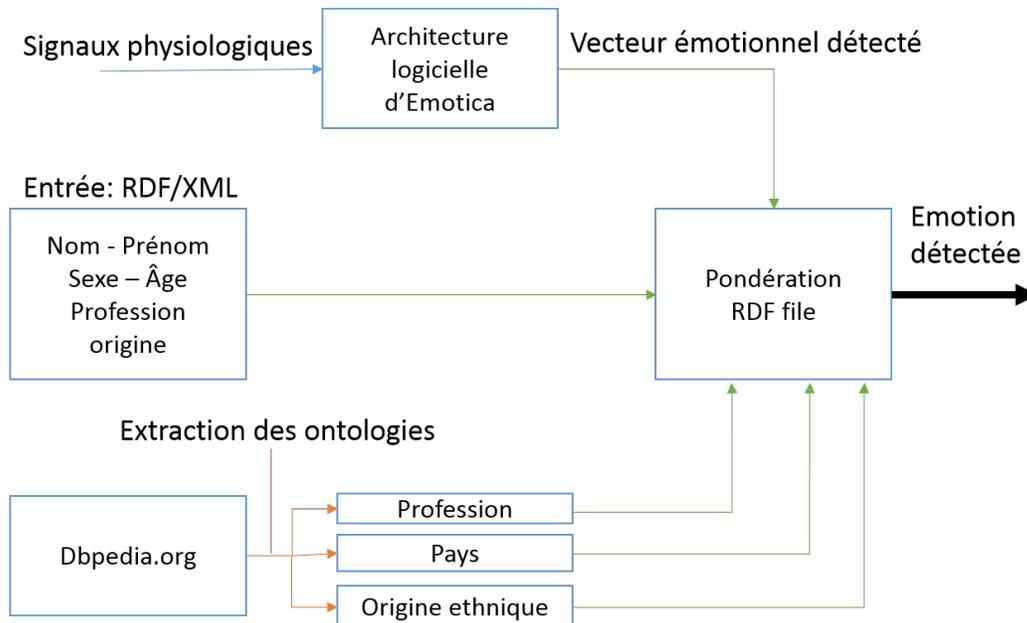


FIGURE F.1 – Modèle de personnalisation de la détection des émotions

- A chaque ontologie, on définit un fichier RDF contenant la pondération correspond à cette variable (`hasImpact=[0 0 0 0.1 0.3 0 0 0 0]`)

ANNEXE F. PERSONNALISATION DE LA DÉTECTION DES ÉMOTIONS

- Les entrées sont automatiquement mises dans un fichier RDF/XML pour avoir toutes les données dans le même langage,
- On détecte d’abord le vecteur émotionnel et ensuite on passe à l’étape de personnalisation (pondération), qui consiste à partir des connaissances de la personne et des ontologies extraites à établir une pondération du vecteur émotionnel calculé dans le système Emotica.

Cette méthodologie permettrait de calibrer le système Emotica afin de rendre la détection plus universelle mais nous n’avons pas pu trouver de modèles de pondération non biaisés. De ce fait, ce travail est suspendu pour le moment afin de comprendre l’effet de la variation dans la détection sur des cas réels. Mais par la suite, nous établirons un modèle de pondération. Le travail qui est donc en cours avec l’université de Danang apportera sans doute des pistes pour la définition d’un modèle de pondération de la détection des émotions en nous basant sur des paramètres comme l’âge, l’origine ethnique et le pays.

Bibliographie

- [Abdat, 2010] Abdat, F. (2010). *Reconnaissance automatique des émotions par données multimodales : expressions faciales et signaux physiologiques*. PhD thesis, UNIVERSITÉ DE METZ. 26, 27, 49
- [Aboulhamid, 2001] Aboulhamid, E. M. (2001). New hardware/software design methodologies. In *Microelectronics, 2001. ICM 2001 Proceedings. The 13th International Conference on*, pages 3–5. IEEE. 98
- [Acharya et al., 2006] Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., and Suri, J. S. (2006). Heart rate variability : a review. *Medical and biological engineering and computing*, 44(12) :1031–1051. 23
- [ADEUNIS, 2015] ADEUNIS, R. (2015). Lorawan product range 2015. *LORA boot camp Grenoble, 17 décembre 2015 : présentation Adeunis RF*. 77
- [Allwein et al., 2000] Allwein, E. L., Schapire, R. E., and Singer, Y. (2000). Reducing multiclass to binary : A unifying approach for margin classifiers. *Journal of machine learning research*, 1(Dec) :113–141. 63
- [Ang et al., 2002] Ang, J., Dhillon, R., Krupski, A., Shriberg, E., and Stolcke, A. (2002). Prosody-based automatic detection of annoyance and frustration in human-computer dialog. In *INTERSPEECH*. Citeseer. 22
- [Arnold, 1960] Arnold, M. B. (1960). Emotion and personality. 14
- [Batliner et al., 2000] Batliner, A., Fischer, K., Huber, R., Spilker, J., and Nöth, E. (2000). Desperately seeking emotions or : Actors, wizards, and human beings. In *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*. 22
- [Baud and Sèdes, 2009] Baud, J.-F. and Sèdes, A. (2009). Vers une approche physiologique de l’interface corporelle pour les arts interactifs. xvi, 27
- [Beiger, 2016] Beiger, F. (2016). *L’enfant et la médiation animale : Une nouvelle approche par la zoothérapie*. Dunod. 130
- [Benini et al., 2001] Benini, L., Bruni, D., Chinosi, M., Silvano, C., Zaccaria, V., and Zafalon, R. (2001). A power modeling and estimation framework for vliw-based embedded systems. In *Proc. Int. Workshop on Power And Timing Modeling, Optimization and Simulation PATMOS*, volume 1, pages 2–3. 79

BIBLIOGRAPHIE

- [Benovoy et al., 2008] Benovoy, M., Cooperstock, J. R., and Deitcher, J. (2008). Biosignals analysis and its application in a performance setting. In *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 253–258. 28
- [Berder and Sentieys, 2010] Berder, O. and Sentieys, O. (2010). Powwow : Power optimized hardware/software framework for wireless motes. In *Architecture of Computing Systems (ARCS), 2010 23rd International Conference on*, pages 1–5. VDE. 96
- [Berthier, 2012] Berthier, N. (2012). *Programmation synchrone de pilotes de périphériques pour un contrôle global de ressources dans les systèmes embarqués*. PhD thesis, Université de Grenoble. 42
- [Bonamy, 2013] Bonamy, R. (2013). *Modélisation, exploration et estimation de la consommation pour les architectures hétérogènes reconfigurables dynamiquement*. PhD thesis, Université Rennes 1. 39
- [Bousefsaf, 2014] Bousefsaf, F. (2014). *Mesure sans contact de l'activité cardiaque par analyse du flux vidéo issu d'une caméra numérique : extraction de paramètres physiologiques et application à l'estimation du stress*. PhD thesis, Université de Lorraine. 23
- [Broad, 1913] Broad, C. (1913). Cournot, a.-essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique. 60
- [Brookmeyer et al., 2007] Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of alzheimer's disease. *Alzheimer's & dementia*, 3(3) :186–191. 130
- [Burch et al., 1993] Burch, R., Najm, F. N., Yang, P., and Trick, T. N. (1993). A monte carlo approach for power estimation. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 1(1) :63–71. 43
- [Cacioppo and Tassinari, 1990] Cacioppo, J. T. and Tassinari, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45(1) :16. 24
- [Campos, 1983] Campos, J. J. (1983). The importance of affective communication in social referencing : A commentary on feinman. *Merrill-Palmer Quarterly (1982-)*, pages 83–87. 16
- [Campos et al., 1989] Campos, J. J., Campos, R. G., and Barrett, K. C. (1989). Emergent themes in the study of emotional development and emotion regulation. *Developmental Psychology*, 25(3) :394. 16
- [Carnot, 2011] Carnot, A. I. (Mis en ligne le 21 juin 2011). “livre blanc : Objets cummunicants & internet des objets”. xvi, 31
- [Chang and Lin, 2011] Chang, C.-C. and Lin, C.-J. (2011). LIBSVM : A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2 :27 :1–27 :27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 62

- [CHELIUS and FRABOULET, 2007] CHELIUS, G. and FRABOULET, A. (2007). Worldsens : outil de simulation pour la conception d'applications de réseaux de capteurs. *Techniques de l'ingénieur. Télécoms, (RE94)*. 96
- [Cheng and Liu, 2008] Cheng, B. and Liu, G.-Y. (2008). Emotion recognition from surface emg signal using wavelet transform and neural network. In *Proceedings of the 2nd international conference on bioinformatics and biomedical engineering (ICBBE)*, pages 1363–1366. 28
- [Choy and Wilton, 2006] Choy, N. C. K. and Wilton, S. J. (2006). Activity-based power estimation and characterization of dsp and multiplier blocks in fpgas. In *2006 IEEE International Conference on Field Programmable Technology*, pages 253–256. IEEE. 43
- [Clavel et al., 2008] Clavel, C., Vasilescu, I., Devillers, L., Richard, G., and Ehrette, T. (2008). Fear-type emotion recognition for future audio-based surveillance systems. *Speech Communication*, 50(6) :487–503. 21
- [Clynes, 1988] Clynes, M. (1988). Generalised emotion how it may be produced, and sentic cycle therapy. In *Emotions and psychopathology*, pages 107–170. Springer. 49, 50
- [Clynes, 2013] Clynes, M. (2013). *Music, mind, and brain : The neuropsychology of music*. Springer Science & Business Media. 49, 50
- [Conn et al., 2008] Conn, K., Liu, C., Sarkar, N., Stone, W., and Warren, Z. (2008). *Towards affect-sensitive assistive intervention technologies for children with autism*. INTECH Open Access Publisher. 28
- [Cosnier, 2003] Cosnier, J. (2003). Douleur, émotion et communication. *Douleur et analgésie*, 16 :23–28. 10
- [da Silva et al., 2014] da Silva, H. P., Fred, A., and Martins, R. (2014). Biosignals for everyone. *IEEE Pervasive Computing*, 13(4) :64–71. 83
- [Damaševičius and Štuikys, 2007] Damaševičius, R. and Štuikys, V. (2007). Estimation of power consumption at behavioral modeling level using systemc. *EURASIP Journal on Embedded Systems*, 2007(1) :1–8. 44
- [Damm et al., 2010] Damm, M., Moreno, J., Haase, J., and Grimm, C. (2010). Using transaction level modeling techniques for wireless sensor network simulation. In *Proceedings of the Conference on Design, Automation and Test in Europe*, pages 1047–1052. European Design and Automation Association. 96
- [Darwin, 1872] Darwin, C. (1872). The expression of the emotions in man and animals. *Oxford University Press Inc, 3 sub edition*. 10, 13
- [Dellaert et al., 1996] Dellaert, F., Polzin, T., and Waibel, A. (1996). Recognizing emotion in speech. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, volume 3, pages 1970–1973. IEEE. 22
- [Deng et al., 2005] Deng, H.-B., Jin, L.-W., Zhen, L.-X., and Huang, J.-C. (2005). A new facial expression recognition method based on local gabor filter bank and pca plus lda. *International Journal of Information Technology*, 11(11) :86–96. 67

BIBLIOGRAPHIE

- [Descartes, 1990] Descartes, R. (1990). 1649. les passions de l'âme. *Oeuvres de Descartes*. 13
- [Desset et al., 2012] Desset, C., Debaillie, B., Giannini, V., Fehske, A., Auer, G., Holtkamp, H., Wajda, W., Sabella, D., Richter, F., Gonzalez, M. J., et al. (2012). Flexible power modeling of lte base stations. In *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2858–2862. IEEE. 44
- [Devillers et al., 2005] Devillers, L., Vidrascu, L., and Lamel, L. (2005). Challenges in real-life emotion annotation and machine learning based detection. *Neural Networks*, 18(4) :407–422. xiv, 20, 22
- [Du et al., 2011] Du, W., Mieveville, F., Navarro, D., and Connor, I. O. (2011). Ideal1 : A validated systemc-based system-level design and simulation environment for wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2011(1) :143. 96
- [Duda et al., 2012] Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons. 55
- [Duffy, 1941] Duffy, E. (1941). An explanation of “emotional” phenomena without the use of the concept “emotion”. *The Journal of General Psychology*, 25(2) :283–293. 13
- [Ekman, 1992] Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6. 10
- [Ekman, 1999] Ekman, P. (1999). Basic emotions. *Handbook of Cognition and Emotion*, Edited by T. Dalgleish and M. Power. 10
- [Ekman, 2006] Ekman, P. (2006). Commentaries : Duchenne and facial expression of emotion. *Studies in Emotion and Social Interaction*. Cambridge University Press, pages 270–284. 24
- [Ekman and Friesen, 1978] Ekman, P. and Friesen, W. V. (1978). *Manual for the facial action coding system*. Consulting Psychologists Press. 18
- [Ekman et al., 1972] Ekman, P., Friesen, W. V., and Ellsworth, P. (1972). *Emotion in the Human Face : Guide-lines for Research and an Integration of Findings : Guidelines for Research and an Integration of Findings*. Pergamon. 10
- [Ekman et al., 1982] Ekman, P., Friesen, W. V., and Ellsworth, P. (1982). What are the similarities and differences in facial behavior across cultures. *Emotion in the human face*, 2 :128–144. 14
- [Fehr and Russell, 1984] Fehr, B. and Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology : General*, 113 :464–486. 9
- [Fernandez and Picard, 1997] Fernandez, R. and Picard, R. (1997). Signal processing for recognition of human frustration. 28
- [film, 2017] film, S. (2017). Outil de production des états émotionnels avec des films. <http://www.ipsp.ucl.ac.be/recherche/filmstim>. Accessed : 2017-09-22. 49

- [Flamary et al., 2015] Flamary, R., Rakotomamonjy, A., and Sebag, M. (2015). Interfaces cerveau-ordinateur : méthodes, applications et perspectives. xvii, 55, 56, 57
- [Forbes and Litman, 2004] Forbes, R. K. and Litman, D. J. (2004). Predicting emotion in spoken dialogue from multiple knowledge sources. In *HLT-NAACL*, pages 201–208. Citeseer. 22
- [Fornaciari et al., 1998] Fornaciari, W., Gubian, P., Sciuto, D., and Silvano, C. (1998). Power estimation of embedded systems : A hardware/software codesign approach. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 6(2) :266–275. 79
- [Friedman et al., 2001] Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin. 55, 57
- [Frijda, 1986] Frijda, N. H. (1986). *The emotions, chapter Physiology of Emotion*. PhD thesis. 27
- [FSC, 2017] FSC (2017). Fondation Suisse pour les Cyberthèses. <http://www.fsc-sfc.org/>. Accessed : 2017-10-29. 33
- [Fummi et al., 2008] Fummi, F., Quaglia, D., and Stefanni, F. (2008). A system-based framework for modeling and simulation of networked embedded systems. In *Specification, Verification and Design Languages, 2008. FDL 2008. Forum on*, pages 49–54. IEEE. 96
- [Gao et al., 2003] Gao, Y., Leung, M. K., Hui, S. C., and Tananda, M. W. (2003). Facial expression recognition from line-based caricatures. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 33(3) :407–412. 19
- [Hamdi, 2012] Hamdi, H. (2012). *Plate-forme multimodale pour la reconnaissance d'émotions via l'analyse de signaux physiologiques : application à la simulation d'entretiens d'embauche*. PhD thesis, Université d'Angers. 67
- [Hassan et al., 2009] Hassan, H. A., Anis, M., and Elmasry, M. (2009). Total power modeling in fpgas under spatial correlation. *IEEE transactions on very large scale integration (VLSI) systems*, 17(4) :578–582. 43
- [Healey and Picard, 1998] Healey, J. and Picard, R. (1998). Digital processing of affective signals. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 6, pages 3749–3752. IEEE. 28
- [Healey and Picard, 2000] Healey, J. and Picard, R. (2000). Smartcar : detecting driver stress. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 4, pages 218–221. IEEE. 28
- [Healey, 2000] Healey, J. A. (2000). *Wearable and Automotive Systems for A ect Recognition from Physiology*. PhD thesis, Citeseer. 23, 49, 50, 68
- [Henderson et al., 2008] Henderson, T. R., Lacage, M., Riley, G. F., Dowell, C., and Kopena, J. (2008). Network simulations with the ns-3 simulator. *SIGCOMM demonstration*, 14(14) :527. 96

BIBLIOGRAPHIE

- [Herbelin et al., 2005] Herbelin, B., Benzaki, P., Riquier, F., Renault, O., Grillon, H., and Thalmann, D. (2005). Using physiological measures for emotional assessment : a computer-aided tool for cognitive and behavioural therapy. *International Journal on Disability and Human Development*, 4(4) :269–278. 28
- [Hiner et al., 2010] Hiner, J., Shenoy, A., Lysecky, R., Lysecky, S., and Ross, A. G. (2010). Transaction-level modeling for sensor networks using systemc. In *Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010 IEEE International Conference on*, pages 197–204. IEEE. 96
- [Hönig et al., 2007] Hönig, F., Batliner, A., and Nöth, E. (2007). Real-time recognition of the affective user state with physiological signals. *Proceedings of the Doctoral Consortium, Affective Computing and Intelligent Interaction*. 28
- [Hsu et al., 2011] Hsu, C.-W., Liao, J.-L., Fang, S.-C., Weng, C.-C., Huang, S.-Y., Hsieh, W.-T., and Yeh, J.-C. (2011). Powerdepot : integrating ip-based power modeling with esl power analysis for multi-core soc designs. In *Proceedings of the 48th Design Automation Conference*, pages 47–52. ACM. 44
- [Huang et al., 1998] Huang, T. S., Chen, L. S., Tao, H., Miyasato, T., and Nakatsu, R. (1998). Bimodal emotion recognition by man and machine. In *ATR Workshop on Virtual Communication Environments*, volume 31. 67
- [info, 2017] info, D. (2017). Protocole Zigbee pour la domotique. <http://www.domotique-info.fr/technologies-domotique/zigbee/>. Accessed : 2017-09-22. 77
- [Istrate, 2003] Istrate, D. (2003). *Détection et reconnaissance des sons pour la surveillance médicale*. PhD thesis, Institut National Polytechnique de Grenoble-INPG. 20
- [Izard and Malatesta, 1980] Izard, C. and Malatesta, C. (1980). Cross cultural perspectives on emotion and emotion communication. *Handbook of cross-cultural psychology, Boston, MA : Allyn and Bacon*, 1. 10
- [Izard, 1971] Izard, C. E. (1971). The face of emotion. 14
- [Jacques et al., 2011] Jacques, D., Zdanowicz, N., Reynaert, C., et al. (2011). Les bases anatomiques et physiologiques du système limbique et leur implication dans le traitement du bégaiement. *Les bégaiements de l'adulte. Wadre : Mardaga*. p45-49. 23
- [Jadoon, 2013] Jadoon, J. K. (2013). *Evaluation of power management strategies on actual multiprocessor platforms*. PhD thesis, Université Nice Sophia Antipolis. 41
- [Jrad et al., 2011] Jrad, N., Congedo, M., Phlypo, R., Rousseau, S., Flamary, R., Yger, F., and Rakotomamonjy, A. (2011). sw-svm : sensor weighting support vector machines for eeg-based brain-computer interfaces. *Journal of neural engineering*, 8(5) :056004. 55
- [Jurik et al., 2009] Jurik, A. D., Bolus, J. F., Weaver, A. C., Calhoun, B. H., and Blalock, T. N. (2009). Mobile health monitoring through biotelemetry. In *Proceedings of the Fourth International Conference on Body Area Networks*, page 12.

- ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). [xvi](#), [26](#)
- [Kahneman, 1973] Kahneman, D. (1973). *Attention and effort*. Citeseer. [25](#)
- [Kang et al., 2014] Kang, S.-M., Leblebici, Y., and Kim, C. (2014). Cmos digital integrated circuits : analysis & design. Technical report, McGraw-Hill Higher Education. [38](#)
- [Kim et al., 2004] Kim, K. H., Bang, S., and Kim, S. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical and biological engineering and computing*, 42(3) :419–427. [66](#), [71](#)
- [Klunnert et al., 1983] Klunnert, M., Campos, J., Sorce, J., Emde, R., and Svejda, M. (1983). Social referencing : Emotional expressions as behavior regulators. *Emotion : Theory, research and experience*, 2 :57–86. [16](#)
- [Koné et al., 2015] Koné, C., Meftah, I. T., Le Thanh, N., and Belleudy, C. (2015). Detecting and classification of the basic emotions using a multimodal approach for emotions detection. In *The Fourth International Conference on Global Health Challenges, GLOBAL HEALTH 2015*. [140](#)
- [Kozhaya and Najm, 1997] Kozhaya, J. N. and Najm, F. N. (1997). Accurate power estimation for large sequential circuits. In *Proceedings of the 1997 IEEE/ACM international conference on Computer-aided design*, pages 488–493. IEEE Computer Society. [43](#)
- [Kreibig et al., 2007] Kreibig, S. D., Wilhelm, F. H., Roth, W. T., and Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear-and sadness-inducing films. *Psychophysiology*, 44(5) :787–806. [28](#)
- [Kuorilehto et al., 2008] Kuorilehto, M., Hännikäinen, M., and Hämäläinen, T. D. (2008). Rapid design and evaluation framework for wireless sensor networks. *Ad Hoc Networks*, 6(6) :909–935. [96](#)
- [Lang, 1995] Lang, P. J. (1995). The emotion probe : Studies of motivation and attention. *American psychologist*, pages 372–372. [24](#)
- [Lang, 2005] Lang, P. J. (2005). International affective picture system (iaps) : Digitized photographs, instruction manual and affective ratings. *Technical Report*. [49](#)
- [Larose, 2005] Larose, D. T. (2005). An introduction to data mining. *Traduction et adaptation de Thierry Vallaud*. [xi](#), [60](#)
- [Lecourt, 2014] Lecourt, E. (2014). *La musicothérapie : Une synthèse d'introduction et de référence pour découvrir les vertus thérapeutiques de la musique*. Editions Eyrolles. [130](#)
- [Lee et al., 2002a] Lee, C. M., Narayanan, S. S., and Pieraccini, R. (2002a). Classifying emotions in human-machine spoken dialogs. In *Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on*, volume 1, pages 737–740. IEEE. [22](#)

BIBLIOGRAPHIE

- [Lee et al., 2002b] Lee, C. M., Narayanan, S. S., and Pieraccini, R. (2002b). Combining acoustic and language information for emotion recognition. In *INTERSPEECH*. 20
- [Levenson, 1992] Levenson, R. W. (1992). Autonomic nervous system differences among emotions. *Psychological science*, 3(1) :23–27. 25
- [Levis et al., 2003] Levis, P., Lee, N., Welsh, M., and Culler, D. (2003). Tossim : Accurate and scalable simulation of entire tinyos applications. In *Proceedings of the 1st international conference on Embedded networked sensor systems*, pages 126–137. ACM. 96
- [Li and John, 2003] Li, T. and John, L. K. (2003). Run-time modeling and estimation of operating system power consumption. In *ACM SIGMETRICS Performance Evaluation Review*, volume 31, pages 160–171. ACM. 79
- [Linden, 1995] Linden, D. (1995). Handbook of batteries. In *Fuel and Energy Abstracts*, volume 4, page 265. 42
- [Liu et al., 2007a] Liu, C., Conn, K., Sarkar, N., and Stone, W. (2007a). Affect recognition in robot assisted rehabilitation of children with autism spectrum disorder. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 1755–1760. IEEE. 28
- [Liu et al., 2008] Liu, C., Conn, K., Sarkar, N., and Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics*, 24(4) :883–896. 28
- [Liu et al., 2007b] Liu, Y., Dick, R. P., Shang, L., and Yang, H. (2007b). Accurate temperature-dependent integrated circuit leakage power estimation is easy. In *Proceedings of the conference on Design, automation and test in Europe*, pages 1526–1531. EDA Consortium. 37
- [Looije et al., 2007] Looije, R., Neerinx, M., and Kruijff, G.-J. M. (2007). Affective collaborative robots for safety & crisis management in the field. *Intelligent Human Computer Systems for Crisis Response and Management (ISCRAM 2007)*, Delft, Netherlands (May 2007), 1. 21
- [Lorandel, 2015] Lorandel, J. (2015). *Étude de la consommation énergétique de systèmes de communications numériques sans fils implantés sur cible FPGA*. PhD thesis, Université Européenne de Bretagne. 38, 43, 44
- [Majdoub, 2012] Majdoub, T. (2012). *Technique de modélisation transactionnelle en vue de l’amélioration de la simulation des modèles de performances des architectures électroniques dans le domaine automobile*. PhD thesis, UNIVERSITE DE NANTES. 34
- [MAM, 2017] MAM (2017). Association MAM : Mémoire, Alzheimer, Mnémothérapie. <http://association-mam.fr/la-mnemothérapie/la-mnemothérapie-2/>. Accessed : 2017-11-06. 131
- [Marrin and Picard, 1998] Marrin, T. and Picard, R. (1998). The conductor’s jacket : A device for recording expressive musical gestures. In *Proceedings of the International Computer Music Conference*. 24

- [Mbarek, 2013] Mbarek, O. (2013). *An electronic system level modeling approach for the design and verification of low-power systems-on chip*. PhD thesis, Université Nice Sophia Antipolis. 41
- [Ménard, 2008] Ménard, M. (2008). *Exploration objective de l'audition à partir des auditory steady-state responses et adaptation à l'implant cochléaire*. PhD thesis, Université Claude Bernard-Lyon I. xvi, 33
- [Mercier, 2003] Mercier, H. (2003). Analyse automatique des expressions du visage. *Application à la langue des signes. Master report, UPS Toulouse*. 19
- [Métrailler, 2005] Métrailler, P. (2005). Système robotique pour la mobilisation des membres inférieurs d" une personne paraplégique. 33
- [Milgram et al., 2006] Milgram, J., Cheriet, M., and Sabourin, R. (2006). "one against one" or "one against all" : Which one is better for handwriting recognition with svms? In *Tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft. 63
- [Mueller et al., 2001] Mueller, W., Ruf, J., Hoffmann, D., Gerlach, J., Kropf, T., and Rosenstiehl, W. (2001). The simulation semantics of systemc. In *Design, Automation and Test in Europe, 2001. Conference and Exhibition 2001. Proceedings*, pages 64–70. IEEE. 99
- [Najoua et al., 2012] Najoua, C., Mohamed, B., and Hedi, B. M. (2012). Analytical dynamic power model for lut based components. In *Design & Technology of Integrated Systems in Nanoscale Era (DTIS), 2012 7th International Conference on*, pages 1–6. IEEE. 44
- [Nakasone et al., 2005] Nakasone, A., Prendinger, H., and Ishizuka, M. (2005). Emotion recognition from electromyography and skin conductance. In *Proc. of the 5th International Workshop on Biosignal Interpretation*, pages 219–222. Citeseer. 24
- [Narayanan, 2002] Narayanan, S. (2002). Towards modeling user behavior in human-machine interaction : Effect of errors and emotions. In *Proc. ISLE Workshop on Dialogue Tagging for Multi-modal Human Computer Interaction*. 22
- [National Health, 2017] National Health (2017). The numbers count : Mental disorders in america, 2013. <http://www.nimh.nih.gov/health/publications/the-numbers-count-mental-disorders-in-america/index.shtml>. Accessed : 2017-09-17. 16
- [Neto et al., 2003] Neto, E. S., Neidecker, J., and Lehot, J. (2003). Comprendre la variabilité de la pression artérielle et de la fréquence cardiaque. In *Annales françaises d'anesthésie et de réanimation*, volume 22, pages 425–452. Elsevier. 23
- [OMS, 2017] OMS (2017). Organisation mondiale de la santé. <http://www.who.int/mediacentre/factsheets/fs369/fr/>. Accessed : 2017-09-12. 3, 16
- [Pan and Jain, 2008] Pan, J. and Jain, R. (2008). A survey of network simulation tools : Current status and future developments. *Email : jp10@cse.wustl.edu*, 2(4) :45. 96

BIBLIOGRAPHIE

- [Panksepp, 1982] Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain sciences*, 5(03) :407–422. [14](#)
- [Pantic and Rothkrantz, 2003] Pantic, M. and Rothkrantz, L. J. (2003). Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, 91(9) :1370–1390. [66](#)
- [Petrushin, 1999] Petrushin, V. (1999). Emotion in speech : Recognition and application to call centers. In *Proceedings of Artificial Neural Networks in Engineering*, volume 710. [22](#)
- [Piolat and Bannour, 2008] Piolat, A. and Bannour, R. (2008). Emotions et affects : Contribution de la psychologie cognitive. In *P. Nagy and D., Boquet (Eds.), Le sujet des émotions au Moyen Age. Paris : Beauchesne Editeur*, pages 53–84. [10](#)
- [Plutchik, 1980] Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of emotion*, 1 :3–31. [xvi](#), [xvii](#), [14](#), [15](#), [29](#), [47](#), [59](#)
- [Polley et al., 2004] Polley, J., Blazakis, D., McGee, J., Rusk, D., and Baras, J. S. (2004). Atemu : a fine-grained sensor network simulator. In *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, pages 145–152. IEEE. [96](#)
- [priya Muthusamy, 2012] priya Muthusamy, R. (2012). Seminar paper : Emotion recognition from physiological signals using bio-sensors. *University of Fribourg, Switzerland*. [66](#)
- [Rainville et al., 2006] Rainville, P., Bechara, A., Naqvi, N., and Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International journal of psychophysiology*, 61(1) :5–18. [28](#)
- [Rani et al., 2007] Rani, P., Sarkar, N., and Adams, J. (2007). Anxiety-based affective communication for implicit human-machine interaction. *Advanced Engineering Informatics*, 21(3) :323–334. [28](#)
- [Rani et al., 2003] Rani, P., Sarkar, N., Smith, C. A., and Adams, J. A. (2003). Affective communication for implicit human-machine interaction. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, volume 5, pages 4896–4903. IEEE. [28](#)
- [Rethinagiri, 2013] Rethinagiri, S. K. (2013). *System-Level Power Estimation Methodology for MPSoC based Platforms*. PhD thesis, Université de Valenciennes et du Hainaut-Cambresis. [41](#)
- [Riliskis and Osipov, 2015] Riliskis, L. and Osipov, E. (2015). Symphony : A framework for accurate and holistic wsn simulation. *Sensors*, 15(3) :4677–4699. [96](#)
- [Rivière and Godet, 2003] Rivière, A. and Godet, B. (2003). L’affective computing : rôle adaptatif des émotions dans l’interaction homme-machine. *Report, Charles de Gaulle University, Lille, France*, pages 9–12. [16](#)
- [Roy et al., 2003] Roy, K., Mukhopadhyay, S., and Mahmoodi-Meimand, H. (2003). Leakage current mechanisms and leakage reduction techniques in deep-submicrometer cmos circuits. *Proceedings of the IEEE*, 91(2) :305–327. [37](#)

- [Russell, 1980] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, pages 1161–1178. 12
- [Russell, 1997] Russell, J. A. (1997). chapitre how shall an emotion be called? *American Psychological Association, Washington, DC, US*. 12
- [Salz and Horowitz, 1989] Salz, A. and Horowitz, M. (1989). Irsim : An incremental mos switch-level simulator. In *Design Automation, 1989. 26th Conference on*, pages 173–178. IEEE. 43
- [Scherer and Zentner, 2001] Scherer, K. R. and Zentner, M. R. (2001). Emotional effects of music : Production rules. *Music and emotion : Theory and research*, pages 361–392. xiv, 10, 11
- [Schlosberg, 1954] Schlosberg, H. (1954). Three dimensions of emotion. *Psychological review*, 61(2) :81. 12
- [Semtech, 2017] Semtech (2017). <http://www.semtech.com/wireless-rf/loran.html>. Accessed : 2017-09-12. xvii, 78
- [Sentieys, 1997] Sentieys, O. (1997). Réduction de consommation d'énergie en électronique embarquée. *Journée scientifique électronique embarquée du*, 24. 40
- [Shafiee et al., 2015] Shafiee, H., Asghar, W., Inci, F., Yuksekkaya, M., Jahangir, M., Zhang, M. H., Durmus, N. G., Gurkan, U. A., Kuritzkes, D. R., and Demirci, U. (2015). Paper and flexible substrates as materials for biosensing platforms to detect multiple biotargets. *Scientific reports*, 5 :8719. 32
- [Shafran et al., 2003] Shafran, I., Riley, M., and Mohri, M. (2003). Voice signatures. In *Automatic Speech Recognition and Understanding, 2003. ASRU'03. 2003 IEEE Workshop on*, pages 31–36. IEEE. 22
- [Shah, 2009] Shah, K. (2009). Power grid analysis in vlsi designs. 43
- [Sharma et al., 1998] Sharma, R., Pavlovic, V. I., and Huang, T. S. (1998). Toward multimodal human-computer interface. *Proceedings of the IEEE*, 86(5) :853–869. 67
- [Shelby et al., 2005] Shelby, Z., Pomalaza-Raez, C., Karvonen, H., and Haapola, J. (2005). Energy optimization in multihop wireless embedded and sensor networks. *International journal of wireless information networks*, 12(1) :11–21. 79
- [Shu et al., 2011] Shu, L., Hauswirth, M., Chao, H.-C., Chen, M., and Zhang, Y. (2011). Nettopo : A framework of simulation and visualization for wireless sensor networks. *Ad Hoc Networks*, 9(5) :799–820. 96
- [Steidl et al., 2005] Steidl, S., Levit, M., Batliner, A., Nöth, E., and Niemann, H. (2005). " of all things the measure is man " : Automatic classification of emotions and inter-labeler consistency. In *ICASSP (1)*, pages 317–320. Citeseer. 22
- [Sunwoo et al., 2010] Sunwoo, D., Wu, G. Y., Patil, N. A., and Chiou, D. (2010). Presto : An fpga-accelerated power estimation methodology for complex systems. In *2010 International Conference on Field Programmable Logic and Applications*, pages 310–317. IEEE. 44

BIBLIOGRAPHIE

- [Takahashi, 2004] Takahashi, K. (2004). Remarks on emotion recognition from multi-modal bio-potential signals. In *Industrial Technology, 2004. IEEE ICIT'04. 2004 IEEE International Conference on*, volume 3, pages 1138–1143. IEEE. 28
- [Takahashi and Tsukaguchi, 2003] Takahashi, K. and Tsukaguchi, A. (2003). Remarks on emotion recognition from multi-modal bio-potential signals. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, volume 2, pages 1654–1659. IEEE. 28
- [Tan et al., 2002] Tan, T. K., Raghunathan, A., and Jha, N. K. (2002). Embedded operating system energy analysis and macro-modeling. In *Computer Design : VLSI in Computers and Processors, 2002. Proceedings. 2002 IEEE International Conference on*, pages 515–522. IEEE. 79
- [Tato, 1999] Tato, R. (1999). *Emotion Recognition in Speech Signal*. PhD thesis, University of Manchester. xiv, 15
- [Tayari et al., 2012] Tayari, I. M., Le-Thanh, N., and Amar, C. B. (2012). Detecting depression using a multidimensional model of emotional states. *Global Health*. 105
- [Tayari, 2013] Tayari, M. I. (2013). *Modélisation, détection et annotation des états émotionnels à l'aide d'un espace vectoriel multidimensionnel*. PhD thesis, Université Nice Sophia Antipolis. xiv, 12, 15, 16, 18, 29, 46, 71
- [Tayari et al., 2013] Tayari, M. I., Le-Thanh, N., and Chokri, B. A. (2013). Multimodal approach for emotion recognition using a formal computational model. *International Journal of Applied Evolutionary Computation (IJAEC)*, 4 :11–25. 14, 59, 67
- [Tellier, 2008] Tellier, M. (2008). Dire avec des gestes. *Le français dans le monde. Recherches et applications*, (44) :40–50. 1
- [Titzer et al., 2005] Titzer, B. L., Lee, D. K., and Palsberg, J. (2005). Aurora : Scalable sensor network simulation with precise timing. In *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, pages 477–482. IEEE. 96
- [Todorovich et al., 2005] Todorovich, E., Boemo, E., Angarita, F., and Vails, J. (2005). Statistical power estimation for fpgas. In *International Conference on Field Programmable Logic and Applications, 2005.*, pages 515–518. IEEE. 43
- [Tomkins, 1962] Tomkins, S. S. (1962). *Affect imagery consciousness : Volume I : The positive affects*, volume 1. Springer publishing company. 14
- [van den Broek et al., 2011] van den Broek, E. L., Janssen, J. H., van der Zwaag, M. D., Westerink, J. H., and Healey, J. A. (2011). Prerequisites for affective signal processing (asp)-part v : A response to comments and suggestions. xiv, 28
- [Vandyke, 2013] Vandyke, D. (2013). Depression detection & emotion classification via data-driven glottal waveforms. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 642–647. IEEE. 29
- [Viau, 2009] Viau, B. (2009). La maison d'alzheimer idéale. *La revue canadienne de la maladie d'Alzheimer et autres démences*, pages 22–24. 130

- [Villon and Lisetti, 2007] Villon, O. and Lisetti, C. (2007). A user model of psychophysiological measure of emotion. In *International Conference on User Modeling*, pages 319–323. Springer. 28
- [Wagner et al., 2011] Wagner, J., Andre, E., Lingensfelder, F., and Kim, J. (2011). Exploring fusion methods for multimodal emotion recognition with missing data. *IEEE Transactions on Affective Computing*, 2(4) :206–218. 67
- [Wang, 2005] Wang, L. (2005). *Support vector machines : theory and applications*, volume 177. Springer Science & Business Media. xi, 19, 61
- [Woodworth and Schlosberg, 1938] Woodworth, R. S. and Schlosberg, H. (1938). Experimental psychology holt. *New York*, pages 327–328. 12, 14
- [Wundt, 1896] Wundt, W. (1896). *Grundriss der Psychologie*. W. Engelmann. 12
- [Zhang et al., 2006] Zhang, H., Berg, A. C., Maire, M., and Malik, J. (2006). Svm-knn : Discriminative nearest neighbor classification for visual category recognition. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 2126–2136. IEEE. 61
- [Zhang et al., 2003] Zhang, Y., Parikh, D., Sankaranarayanan, K., Skadron, K., and Stan, M. (2003). Hotleakage : A temperature-aware model of subthreshold and gate leakage for architects. *University of Virginia Dept. of Computer Science Technical Report*. 39
- [Zhi et al., 2001] Zhi, Q., Kaynak, M. N., Sengupta, K., Cheok, A. D., and Ko, C. C. (2001). Hmm modeling for audio-visual speech recognition. In *ICME*. 67
- [Zhou and Wang, 2005] Zhou, J. and Wang, X. (2005). Multimodal affective user interface using wireless devices for emotion identification. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 7155–7157. IEEE. 28
- [Zhu et al., 2004] Zhu, D., Melhem, R., and Mossé, D. (2004). The effects of energy management on reliability in real-time embedded systems. In *Computer Aided Design, 2004. ICCAD-2004. IEEE/ACM International Conference on*, pages 35–40. IEEE. 79
- [Zong and Chetouani, 2009] Zong, C. and Chetouani, M. (2009). Hilbert-huang transform based physiological signals analysis for emotion recognition. In *2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 334–339. IEEE. 68, 71