

Université de Montréal

Analyse des émotions dans un jeu vidéo.

par Doumbouya René Lanciné

Département d'informatique et de recherche opérationnelle, Université de Montréal

Faculté des arts et sciences

Mémoire présenté

en vue de l'obtention du grade de M.sc

en Informatique

option Mémoire

Décembre 2016

© Doumbouya René Lanciné, 2016

Résumé

Tout le long d'une partie de jeu vidéo, nous voyons des personnes devenir agressives, pousser des cris de joie et parfois pleurer. Ces réactions observées nous poussent à nous demander quelles sont les émotions majeures créées chez un joueur et les éléments responsables. Dans ce but, notre étude propose une méthode et un outil d'analyse émotionnelle du jeu ainsi qu'un accès facile aux informations pour les utilisateurs. Cette analyse fournit une illustration graphique des transitions émotionnelles suivant le parcours du joueur au cours de la partie. Pour élaborer cette méthode nous avons réalisé une expérience sur un groupe de 21 personnes et utilisé 5 outils de mesure physiologique. Nous l'avons axée principalement sur la combinaison de 2 d'entre eux, la reconnaissance d'expressions faciales et le suivi du regard. Nos résultats montrent l'efficacité de la méthode à travers une amélioration de la précision tant dans l'identification des émotions que de leurs déclencheurs. De cette étape nous est venue l'envie de prédire les émotions des joueurs. C'est dans cette optique que nous présentons dans ce mémoire 2 approches basées sur le modèle théorique d'évaluation de l'émotion Ortony, Clore et Collins et utilisant des algorithmes d'apprentissage machine, en particulier une version des *k plus proches voisins* et les *forêts d'arbres de décision*. Les résultats obtenus étant prometteurs, une telle réalisation permettrait d'élargir les possibilités dans la conception de jeux vidéo.

Mots-clés : Émotions dans les jeux vidéo, Analyse émotionnelle, Intelligence émotionnelle, Informatique affective.

Abstract

Throughout a video game, we see people becoming aggressive, shouting for joy and sometimes crying. These observed reactions prompt us to ask what are the major emotions created in a player during a game session, as well as to identify the responsible elements. To this end, our study proposes a method of emotional analysis of a game as well as a tool whose implementation uses this method in order to give easy access to information for users. This analysis submits a graphic illustration of emotional transition along the course of the player during the game. To develop this method we performed an experiment on a group of 21 people and used 5 physiological measurement tools. We focused mainly on the combination of 2 of them, the recognition of facial expressions and the follow-up of the gaze. Our results show the effectiveness of this method through an improvement of the precision both in the identification of the emotions and their triggers. From this step came the desire to predict the emotions of the players. It is with this in mind that we present in this paper 2 approaches based on the theoretical model of emotion evaluation Ortony, Clore and Collins and Machine learning algorithms, in particular a version of the *k nearest neighbors* and the *Random forest classifiers*. The results obtained being promising, such an achievement would widen the possibilities in the design of video games.

Keywords : Emotions in video games, Emotional analysis, Emotional intelligence, Affective computing.

Table des matières

RÉSUMÉ	I
ABSTRACT	II
TABLE DES MATIÈRES	III
LISTE DES TABLEAUX	VI
LISTE DES FIGURES.....	VII
LISTE DES ABRÉVIATIONS	IX
REMERCIEMENTS	XII
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 LES ÉMOTIONS DANS LES JEUX VIDÉO	6
2.1 LES TYPES DE JEUX VIDÉO.....	6
2.2 LA GÉNÉRATION D'ÉMOTIONS DANS LES JEUX VIDÉO.....	8
2.3 COMMENT MESURER LES ÉMOTIONS ?.....	10
2.3.1 Les mesures subjectives	10
2.3.2 Les mesures objectives.....	11
2.4 REVUE DE LITTÉRATURE ET AUTRES TRAVAUX SUR LE SUJET	16
2.5 LA PRÉDICTION DES ÉMOTIONS DANS UN JEU	20
CHAPITRE 3 L'ANNOTATION AFFECTIVE DES ÉLÉMENTS DANS UN JEU	22
3.1 L'ANALYSE DE LA STRUCTURE DES JEUX VIDÉO	22
3.2 L'OCULOMÉTRIE	24
3.3 LES TYPES D'ÉMOTIONS.....	25
3.4 LE GRAPHE ÉMOTIONNEL.....	29
CHAPITRE 4 L'EXPÉRIENCE	31
4.1 CONTEXTE DE L'EXPÉRIENCE.....	31
4.1.1 <i>Outlast</i>	31
4.1.2 <i>Les participants</i>	36
4.2 MESURE DES ÉMOTIONS	38
4.2.1 <i>La plateforme de recherche biométrique iMotions</i>	38

4.2.2 L'oculomètre Tobii Pro TX300	39
4.2.3 Le module de reconnaissance d'expressions faciales FACET.....	40
4.2.4 Le casque Emotiv ÉPOC.....	40
4.2.5 Le bracelet Affectiva Q-sensor 2.0	41
4.2.6 L'électrocardiogramme fourni par BIOPAC.....	42
CHAPITRE 5 "ÉMOGRAPH" : LA MÉTHODE ET L'APPLICATION.....	44
5.1 LA MÉTHODE	44
5.1.1 L'analyse du suivi du regard	44
5.1.2 L'analyse des expressions faciales	46
5.1.3 La recherche de l'émotion dominante	50
5.2 L'APPLICATION WEB DYNAMIQUE : ÉMOGRAPH.....	52
CHAPITRE 6 LES RÉSULTATS	55
6.1 LA MÉTHODE	55
6.2 LA WEB APPLICATION ÉMOGRAPH.....	70
6.2.1 Le module d'analyse émotionnelle (Émotrigger)	70
6.2.2 Le module de visualisation.....	72
6.3 LES INTERPRÉTATIONS DES RÉSULTATS	74
CHAPITRE 7 LA PRÉDICTION ÉMOTIONNELLE	77
7.1 LE MODÈLE OCC	79
7.1.1 Les types de situations.....	79
7.1.2 Le lien entre les types de situations et les émotions	81
7.1.3 Les variables d'évaluation.....	83
7.1.4 La description formelle des scènes.....	86
7.2 L'APPRENTISSAGE MACHINE.....	88
7.2.1 L'apprentissage supervisé.....	90
7.2.2 Les phases de l'apprentissage.....	91
7.2.3 Les algorithmes utilisés.....	94
7.3 LES APPROCHES PROPOSÉES.....	97
7.3.1 La construction des ensembles d'entraînements.....	100
7.3.2 L'approche par analyse individuelle.....	101
7.3.3 L'approche par l'analyse générale	102

7.4 LES RÉSULTATS ET LEURS INTERPRÉTATIONS	102
7.4.1 <i>La prédiction</i>	103
7.4.2 <i>Le module de prédiction (Nostradamus)</i>	106
CHAPITRE 8 LES CONCLUSIONS	112
BIBLIOGRAPHIE	I
ANNEXE A FORMULAIRE DE CONSENTEMENT	10
JEUX VIDÉO CITÉS	II

Liste des tableaux

Tableau 1. Genres de jeux vidéo.....	6
Tableau 2. Description des muscles du visage en fonction des émotions	14
Tableau 3. Parcours de la séquence de jeu.....	33
Tableau 4. Interprétation des valeurs de FACET.....	46
Tableau 5. Émotions de base reconnues chez divers participants	59
Tableau 6. Occurrences d'émotions dominantes par catégories de joueurs.....	63
Tableau 7. Captures de jeu.....	67
Tableau 8. Relations entre les émotions et les types de situations.....	82
Tableau 9. Les variables d'évaluation du modèle OCC	84
Tableau 10. Variables centrales et locales impliquées à chaque type de situation.	85
Tableau 11. Échelle des valeurs prises par les variables centrales et locales.	87
Tableau 12. Les variables globales, centrales et locales et les valeurs associées.	87
Tableau 13. Récapitulatif des résultats de l'apprentissage machine.	105
Tableau 14. Matrice de confusion.....	106

Liste des figures

Figure 1. Population mondiale de joueurs en 2015 (en millions).....	1
Figure 2. Le <i>Self-Assessment Manikin</i> utilisé pour évaluer les dimensions affectives respectivement valence, excitation et dominance.....	11
Figure 3. Les relations entre le joueur, l'interface utilisateur et les mécanismes de base dans un jeu.....	23
Figure 4. Illustration du concept de modes de gameplay.....	24
Figure 5. Exemple du modèle circomplexe tiré de (J. A. Russell, 1980).	27
Figure 6. Un exemple des émotions de bases décrites par Ekman	28
Figure 7. Construction de la représentation graphique	29
Figure 8. Couverture du jeu Outlast.....	33
Figure 9. Exemple d'erreur dans la collecte des données.	37
Figure 10. Distribution des participants.....	38
Figure 11. Unité de suivi du regard	39
Figure 12. Oculomètre Tobii Pro TX300.....	40
Figure 13. Le casque Emotiv EPOC.....	41
Figure 14. Le placement des électrodes sur la tête	41
Figure 15. Le bracelet Affectiva Q-sensor 2.0.....	42
Figure 16. Le système Biopac MP150	42
Figure 17. Dispositif expérimental	43
Figure 18. Annotation de la vidéo d'un participant	45
Figure 19. Extrait du fichier "Moving_AOI"	46
Figure 20. Extrait du fichier exporté de FACET	47
Figure 21. Effet de la correction par la valeur de référence.....	48
Figure 22. Seuil de colère relatif.....	49
Figure 23. Exemple de recherche d'émotions fortes (Colère).....	51
Figure 24. Architecture globale de base d'Émograph	53
Figure 25. Émotions dominantes du participant P21	55
Figure 26. Capture de jeu du participant P21	56
Figure 27. Temps de jeu du participant P21	56

Figure 28. Émotions dominantes du participant P08	57
Figure 29. Capture du jeu du participant P08	58
Figure 30. Temps de jeu du participant P08	58
Figure 31. Graphe de transition émotionnelle du participant P21.	61
Figure 32. Statistiques globales de la séquence de jeu	62
Figure 33. Répartition totale des segments selon le genre.....	62
Figure 34. Proportion de dégoût par zones d'intérêt.	64
Figure 35. Proportion de peur par zones d'intérêt.	64
Figure 36. Proportion de tristesse par zones d'intérêt.	65
Figure 37. Proportion de joie par zones d'intérêt.	65
Figure 38. Proportion de Surprise par zones d'intérêt.....	66
Figure 39. Proportion de Colère par zones d'intérêt.	67
Figure 40. Fenêtre d'accueil d'Émograph.....	70
Figure 41. Interface du module d'analyse émotionnelle par participants	71
Figure 42. Interface du module d'analyse émotionnelle par zones d'intérêt	72
Figure 43. Interface du module de visualisation.....	73
Figure 44. Image du jeu.	74
Figure 45. Modèle Ortony, Clore et Collins (Ortony et al., 1988)	81
Figure 46. L'art de l'apprentissage.	92
Figure 47. Procédé de la technique <i>Leave-one-out cross validation</i>	93
Figure 48. Architecture finale du système.	98
Figure 49. Rapport de classification	106
Figure 50. Interface du module de prédiction pour l'objectif 1.....	109
Figure 51. Interface du module de prédiction pour l'objectif 2.	110

Liste des abréviations

AOI : Area Of Interest – Zones d'intérêt

ANS : Autonomic Nervous System

JV. : Jeux vidéo

EEG : Électroencéphalogramme

ECG/EKG : Électrocardiographie

GSR : Galvanic Skin Response

HR : Heart Rate

UX : Interface Utilisateur

À ma mère, aucune dédicace ne saurait exprimer mon amour, mon respect et toute la considération pour les sacrifices que tu as consenti pour mon éducation et mon bien être. Je te remercie pour toutes tes prières et bénédictions que tu me portes depuis ma naissance, j'espère qu'elles m'accompagneront toute ma vie. Que ce modeste accomplissement soit l'exaucement de tes prières, de tes innombrables sacrifices. Puisse le bon Dieu, le Très Haut t'accorder santé, bonheur, longue vie et faire en sorte que jamais je ne te déçoive.

À ma grand-mère et mon grand frère pour leur amour infini, leur soutien inconditionnel et leurs prières à mon égard.

À la mémoire de mon grand-père, j'aurais tant aimé que tu puisses lire ce mémoire. Mais je sais que tu veilles sur moi depuis là-haut. Que le bon Dieu ait ton âme dans sa sainte miséricorde.

Remerciements

J'aimerais adresser ma sincère gratitude aux personnes qui m'ont aidé dans la réalisation de ce mémoire pour leur présence et leur soutien.

En premier lieu, je remercie M. Claude Frasson, mon directeur de recherche pour sa pédagogie, sa grande expérience et pour m'avoir donné l'opportunité de participer à une recherche si enrichissante. Son encadrement, sa disponibilité et son support financier ont joué un rôle majeur dans l'aboutissement de ce mémoire.

En second lieu, mes remerciements s'adressent au Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) - Subventions de recherche et développement coopérative (RDC) et à Mr Yann Cyr, président de la compagnie Beam Me Up Games pour leur financement sans lequel cette recherche n'aurait pu se faire.

Enfin mes remerciements vont à mes parents, amis et collègues de l'université. Leur présence, conseils et encouragements contre vents et marées ont été la source de ma motivation qui a rendu ce travail possible.

CHAPITRE 1 INTRODUCTION

Créer et jouer à des jeux pour se divertir est dans la nature des êtres humains. Les grecs anciens, les vikings, et vraisemblablement même nos ancêtres des cavernes avaient des jeux organisés à partir d'un système à base de règles simples. Les choses sont bien différentes de nos jours, les jeux sont sophistiqués et ont été revus très à la hausse avec des éléments choisis, construits, composés pour captiver l'esprit du joueur.

Le divertissement numérique que nous connaissons aujourd'hui a commencé avec un projectile blanc envoyé dans l'espace en faisant un sifflement strident dans un sous-sol du MIT en 1961. Le projectile s'écrasait sur le vaisseau ennemi sans aucun effet sonore particulier, aucun bruit d'explosion ou d'impact. Ce fut la première salve de torpilles de ce qui serait connu plus tard comme *Spacewar!* (S. Russell, Graetz, & Witaenem, 1962). L'humanité ne se doutait pas que les jeux vidéo étaient nés. En effet *Spacewar!* a marqué l'ouverture d'un chapitre crucial de notre ère.

De nos jours, on estime à plus de 1.2 milliards le nombre de joueurs à travers le monde comme montré sur la figure 1 (SuperData, 2015). La demande des joueurs sans cesse en quête de jeux divertissants a conduit les compagnies à produire de plus en plus de jeux vidéo. Vu la croissance de la production de jeux, la compétition dans l'industrie du jeu est devenue intense. En effet cette popularité entraîne un aspect financier très présent qui se traduit par de grosses sommes d'argent investies. Le rapport mondial du marché des jeux pour l'année 2014 estime le chiffre d'affaires de l'industrie du jeu à \$81.5 milliard (Newzoo, 2014).

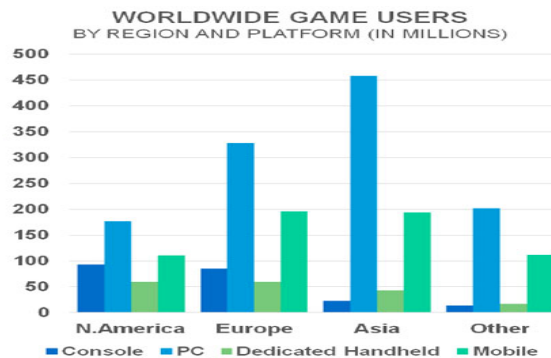


Figure 1. Population mondiale de joueurs en 2015 (en millions)

Cette compétition a poussé les créateurs de jeux vidéo à fournir plus d'efforts pour produire des jeux qui satisferont au maximum les consommateurs. En effet pour être compétitifs, les jeux vidéo doivent faire intervenir le côté humain en créant un lien affectif avec le joueur. Ils attirent l'attention à travers des éléments culturels et esthétiques. Actuellement ces éléments sont conçus avec les technologies les plus avancées pour être au plus proche de la réalité. L'avènement des jeux comme *Grand Theft Auto*, *Uncharted*, *The Matrix* a permis d'influencer significativement la manière de divertir des joueurs. Cela signifie qu'ils trouvent plus de sens dans l'interaction et la non-linéarité de systèmes dont les résultats dépendent de leurs choix. Ce qui engage une certaine dimension personnelle avec le jeu, une dimension émotionnelle. En effet la perception de ces éléments immerge le joueur dans un environnement virtuel ce qui génère chez lui des réactions diverses et reflètent des émotions. Si certaines personnes en présence d'un monstre poussent des cris de terreur, d'autres sont dégoûtées ou surprises.

Il était inévitable que cette forte popularité mêlée à l'envie d'en savoir plus sur la question émotionnelle de l'expérience du joueur attire une nouvelle génération de chercheurs qui en ont fait leur centre d'intérêt. La recherche dans le monde du jeu vidéo a émergé au point de devenir une discipline scientifique à part entière avec des journaux dédiés, des subventions de recherche, des conférences à travers le monde.

Pour trouver quels sont les facteurs responsables de ces expressions il faut comprendre comment sont conçus et structurés les jeux vidéo et quels sont les éléments qui entrent en interactions avec les joueurs. Tous les jeux vidéo proposent un ensemble de défis et d'actions qui divertissent. Les joueurs apprécient les défis aussi longtemps qu'ils ont une attente raisonnable d'être en mesure de l'accomplir. Les aspects esthétiques et cinématographiques jouent un rôle important dans la perception du joueur. En effet les jeux vidéo peuvent être considérés comme une succession organisée de scènes. Il est donc nécessaire d'extraire d'un jeu vidéo des séquences particulières afin de les soumettre à une analyse.

Du côté humain, avec les progrès scientifiques et techniques des nouvelles technologies notamment dans les domaines du traitement des images, des vidéos, de la voix, des signaux physiologiques reliés aux émotions, détecter les émotions chez une personne n'est plus une tâche qui sort de l'ordinaire. Ainsi, un système (un ordinateur, un robot, un agent intelligent, etc.) peut détecter l'émotion d'un individu à partir des canaux habituels utilisés par les humains,

soit l'expression du visage, l'intonation de la voix, la posture du corps, les gestes et les mouvements. Il peut également les détecter en analysant les réponses physiologiques comme par exemple le rythme cardiaque, la tension artérielle, le pouls, la conductance de la peau et la respiration. Ces différentes informations sont recueillies à l'aide d'outils particuliers, de plus en plus accessibles aux chercheurs, comme les caméras détectant avec précision chaque geste, les oculomètres permettant de suivre et d'enregistrer le mouvement des yeux, les capteurs placés sur la peau, etc. En effet ces outils permettent d'effectuer la capture des données physiologiques et la reconnaissance automatique des émotions des joueurs.

Dans notre étude nous nous concentrons sur la reconnaissance des expressions faciales combinée aux données recueillies par la caméra qui a fait le suivi du regard. Il est pertinent d'avoir une vision claire de l'évolution des émotions du joueur tout le long de la partie. C'est dans ce sens que notre étude propose une illustration graphique qui trace le parcours émotionnel du joueur. C'est à dire que nous pouvons identifier quelle était l'émotion pertinente du joueur pendant les scènes du jeu.

La majeure partie des systèmes de reconnaissance automatique des émotions utilisent des algorithmes d'apprentissage, en particulier ceux supervisés. Ils sont utilisés pour classer les émotions des individus à partir d'informations telles que les signaux électriques du cerveau, l'intonation de la voix, l'expression du visage, qui aident à les reconnaître. Pour ce faire, ces algorithmes nécessitent un entraînement préalable sur plusieurs occurrences du phénomène. Ces différents cas sont rassemblés dans ce qu'on appelle un ensemble d'entraînement ou d'apprentissage. En parallèle, la dimension émotionnelle a de plus en plus suscité les intérêts des chercheurs. En effet l'objectif était de produire des systèmes qui simulent le raisonnement et comportement humain. Plusieurs psychologues appuient l'importance de détenir une information sur la situation ou le contexte avant de déduire l'émotion d'une personne à partir de son comportement, ses gestes, ses signaux physiologiques. Ainsi, dans un but d'humaniser un système, l'incorporation de l'émotion dans la plupart des travaux portant sur les systèmes intelligents s'est faite pour simuler le processus humain de génération de l'émotion à l'aide de modèles computationnels inspirés des modèles théoriques établis par des psychologues. Dans la littérature relative à l'émotion, il y a principalement trois approches qui diffèrent selon la spécification des composantes principales intervenant dans le processus de génération (la

cognition, le processus physiologique, etc.) et dans les interactions entre ces différentes composantes (Gratch, Marsella, & Petta, 2009). Néanmoins, les théories les plus appliquées dans les systèmes intelligents et qui représentent la troisième approche sont les théories de l'évaluation cognitive comme dans les travaux de (Scherer, 1986) et de (Lazarus, 1991). Ces théories se focalisent sur la perception cognitive des situations ou des événements engendrant des réponses émotionnelles.

Dans ce mémoire nous proposons l'introduction des théories d'évaluation cognitive en combinaison à la reconnaissance automatique des émotions et aux algorithmes d'apprentissage machine, pour répondre à la nouvelle problématique qui a vu le jour au fil de l'avancée de notre recherche. À savoir, serions-nous en mesure d'anticiper les émotions des joueurs ? L'objectif étant d'intégrer un aspect prédictif dans notre recherche, qui se base sur les réponses émotionnelles des joueurs en face des scènes du jeu. Pour répondre à cette question, il faut tout d'abord être en mesure de traduire les situations émotionnelles rencontrées dans le jeu en langage compréhensible par une machine. Nous avons utilisé un modèle théorique d'évaluation cognitive des émotions que nous avons complété par une spécification appropriée des aspects sociodémographiques et psychologiques des joueurs. La suite, consiste à doter notre outil d'un pouvoir de prédiction convenable par l'intégration d'algorithmes d'apprentissages capables d'exploiter les données recueillies. La performance prédictive est évaluée à l'aide des données collectées.

Ainsi, ce mémoire s'articule autour de 8 chapitres. Il débute par la description des rapports étroits entre les émotions et les jeux vidéo, on retrouve également un bref état des connaissances sur la prédiction des émotions dans le domaine des jeux vidéo dans le chapitre 2. Suivi de la nécessité de la décomposition du jeu en scènes ainsi que leur annotation dans le chapitre 3. En ce qui concerne le chapitre 4, nous présenterons la phase d'expérimentation qui a été menée sur 21 participants et qui a conduit au développement d'une méthode d'analyse dynamique et d'un outil tous les deux nommés 'EMOGRAPH'. Le chapitre 5 présente en détail la méthode créée ainsi que l'outil qui l'utilise avant d'en présenter et interpréter dans le chapitre 6 les résultats obtenus. Compte tenu du fait de l'apparition tardive de la problématique de prédiction, nous l'avons placé dans l'avant dernier chapitre. De ce fait, le chapitre 7, explique en détail la problématique de la prédiction, son intégration dans notre étude et présente les

résultats obtenus et son effet sur l'application développée. Ce mémoire s'achève avec le chapitre 8, qui expose les contributions de notre recherche et les améliorations possibles pour de futures recherches.

CHAPITRE 2 LES ÉMOTIONS DANS LES JEUX VIDÉO

2.1 Les types de jeux vidéo

Les réactions d'un être humain en train de jouer sont différentes d'un jeu à l'autre. Il existe des études complexes sur la conception de jeux vidéo afin de générer les effets souhaités chez les joueurs. De ce fait il est important d'identifier les principaux genres de jeux vidéo.

Il faut savoir au préalable que le gameplay d'un jeu en définit son genre. Nous pouvons alors définir le genre d'un jeu vidéo comme étant une catégorie dans laquelle les jeux qui s'y trouvent sont liés par des caractéristiques communes de gameplay. Il n'est pas évident de définir le terme gameplay car il n'a pas d'équivalent en français. Il fait référence à la jouabilité, la maniabilité et l'ergonomie d'un jeu. Il représente la façon dont un jeu se joue. Le genre d'un jeu n'est pas toujours défini par son contenu ou sa difficulté mais plutôt par son challenge. En effet un jeu de rôle de l'époque médiévale est différent d'un jeu de guerre de l'époque médiévale. Le genre de jeu peut comprendre lui-même une grande variété de jeux vidéo conduisant à une classification plus spécifique appelée sous-genre.

Le tableau suivant dresse une liste de genres classiques de jeux vidéo avec leurs sous-genres éventuels ainsi que quelques exemples.

Tableau 1. Genres de jeux vidéo

Genre	Sous-genre	Exemples
Action : la plupart des challenges sont orientés vers les aptitudes physiques et la coordination des joueurs.	Tirs, Plateforme, Combat, Puzzle.	<i>Doom, Street fighter, Tetris.</i>

<p>Aventure : représente une histoire interactive au sujet d'un personnage incarné par le joueur.</p>		<p><i>Ace Attorney, The Longest Journey, Fahrenheit.</i></p>
<p>Action–Aventure : représente une combinaison des genres qui composent son nom.</p>	<p>Première/troisième personne, Carte ouverte, Infiltration, Jeu d'horreur.</p>	<p><i>Metroïd Prime, Grand Theft Auto, Metal Gear Series, Outlast.</i></p>
<p>Stratégie : les challenges sont représentés par des conflits stratégiques et le joueur peut choisir parmi une variété d'actions à plusieurs endroits du jeu.</p>		<p><i>League of Legends, Final Fantasy Tactics, Tom Clancy's End War.</i></p>
<p>Jeux de rôles : le joueur contrôle un ou plusieurs personnages qu'il guide à travers une série de quêtes prédéterminées.</p>		<p><i>Final Fantasy IV, The Lord of the Rings Online.</i></p>
<p>Simulation : conçus pour simuler de manière la plus proche les aspects réels ou fictifs de la réalité.</p>	<p>Construction et gestion, Véhicule, Vie.</p>	<p><i>SimCity, SimLife, Spy Hunter.</i></p>
<p>Sports : simulation sportive contre un adversaire humain ou l'intelligence artificielle.</p>	<p>Course, Combat, Sports.</p>	<p><i>Fight night Champion, FIFA 17, F1 2013.</i></p>

Ce mémoire porte sur un jeu de genre action-aventure, précisément de sous-genre jeu d'horreur (survival-horror en anglais). La sensation de pouvoir ou de sécurité que donnent les avatars de beaucoup de jeux n'existe pas dans ce genre de jeux, l'avatar tel que présenté souligne l'impuissance. Les joueurs vivent l'expérience d'être victime de la violence plutôt que de créer de la violence. Cette catégorie utilise l'avancée des technologies graphiques pour créer des environnements horribles et avec extrêmement de sang. L'objectif de ce genre de jeu est de créer la peur, la surprise, le divertissement ou le rire chez les joueurs. Nous donnerons plus de détail sur le jeu choisi dans le chapitre 4.

2.2 La génération d'émotions dans les jeux vidéo

Imaginez ce scénario dans un jeu vidéo : vous êtes dans un couloir faiblement éclairé, avec des planches de bois sur les fenêtres. Vous avancez dans le couloir tranquillement, quand une main sortie de nulle part brise les planches de bois et vous étrangle. Les mots "Fin de la partie" s'affichent à l'écran. Que ressentez-vous ? De la surprise mêlée à la colère ? De la joie ? De la peur ? Lorsqu'il s'agit d'émotions, les jeux vidéo touchent nos instincts les plus basiques, une colère qui pousse à jeter la manette contre un mur dans le scénario précédent ou avoir le sourire d'une oreille à l'autre après avoir marqué un but dans *FIFA 17*. Il s'agit dans la conception de jeux vidéo de distinguer principalement la nature des émotions créés chez les joueurs et non la complexité de l'expérience émotionnelle.

Les émotions chez les êtres humains viennent de leurs expériences et de leur perception. Les jeux vidéo étant des simulations de plus en plus proches de la réalité permettent de faire vivre aux joueurs tout type d'expériences.

L'histoire dans laquelle évolue le personnage a toute sa part d'influence. Elle permet de garder le joueur intéressé longtemps dans le jeu. Les émotions plus profondes viennent lorsque le joueur se sent lié au personnage cela se fait par une histoire bien écrite. Elle immerge émotionnellement le joueur, celui-ci séjourne par la pensée dans un monde virtuel dans lequel il ne court aucun danger réel. Les émotions peuvent être caractérisées par une forte envie d'agir d'une façon ou d'une autre jusqu'à ce que l'épisode soit clos à cause d'un changement de situation (Tan, 1996). Cette tendance à l'action est uniquement virtuelle au niveau du récit. Le joueur ne peut généralement rien changer au déroulement de l'histoire. L'histoire est d'ailleurs

racontée par l'intermédiaire de scènes cinématiques qui surviennent à des moments précis du jeu sous forme de vidéo. Elles déterminent les actions prises par les PNJ (personnages non joueurs) et l'avatar, ainsi que la façon dont ces actions sont présentées dans le jeu. Les histoires mettant en avant des scénarios qui sont réalisés par des protagonistes avec des traits humains et dont les buts et le sort intéressent le joueur, génèrent des émotions telles que l'espoir, la peur, l'angoisse, la sympathie, la pitié, le soulagement, la gratitude, l'admiration, la honte, la colère, la terreur, la joie et la tristesse (Tan, 1996).

Jouer à jeu ne se résume pas à comprendre son histoire mais plutôt à résoudre des problèmes, faire face à des situations, affronter des adversaires et tant d'autres actions. L'histoire doit être en concordance avec le gameplay. Comme mentionné dans la section 2.1, il n'est pas évident d'exprimer une définition complète du gameplay. Plusieurs concepteurs de jeu ont tenté de trouver une définition à travers les années. Par exemple Sidney K. « Sid » Meier le créateur de jeux vidéo canadien, le définit comme "une suite de choix intéressants" (Rollings & Morris, 2003). Pour Ernest Adams, concepteur de jeux et écrivain américain, il s'agit des défis et actions qui divertissent (Adams, 2009). Les défis que le joueur doit surmonter et les actions que le jeu offre pour le faire. Malgré le fait que les émotions ressenties soit influencées et suggérées par la conception du jeu, la notion de gameplay est essentielle dans la définition de l'expérience émotionnelle d'un jeu et suscite des réactions émotionnelles. Le joueur n'a pas peur pour sa propre survie dans un jeu mais étant donné que ses actions sont synchronisées à celles de son avatar du monde virtuel, il peut avoir peur que le monstre l'attrape par exemple, il ressent cette crainte comme si c'était lui qui était en danger.

L'esthétique est un élément clé de la génération d'émotions dans les jeux vidéo autant dans la conception des personnages que dans celle du monde virtuel. La conception des personnages est un aspect important pour susciter des réponses émotionnelles. Le personnage incarné par le joueur ou ceux avec lesquels il interagit donne au monde fictif plus de crédibilité à ses yeux. La plupart des jeux établissent leur gameplay autour de personnages. Les joueurs ont besoin de personnages attractifs pour s'y identifier et développer une relation avec eux et avec leurs problèmes. L'apparence des personnages a donc un énorme impact sur la manière dont les joueurs se sentent par rapport à eux. Cette apparence comprend une variété de caractéristiques. L'apparence physique (humains, robots, animaux, etc.), le genre, les habits, les

armes, les objets symboliques, le rôle dans le jeu, la personnalité, l'attitude, le son de la voix, la langue parlée. Outre l'avatar, le décor artistique dans lequel il navigue tient également sa place. Il doit être en adéquation avec l'histoire. Le plaisir esthétique se traduit par la forme ou la structure avec laquelle le joueur perçoit l'élément affiché sans aucune relation avec ce qui se déroule dans le monde irréel. Il est évident qu'un jeu ayant une esthétique soignée sera émotionnellement plus intéressant qu'un jeu "vilain" et ce malgré la qualité de l'histoire ou du gameplay.

Pour avoir un effet sur le joueur, les éléments cités ont besoin d'un moyen d'être perçu. Il s'agit de l'interface utilisateur. C'est cet élément qui crée l'expérience du joueur, rend la combinaison de toutes les mécaniques génératrices d'émotions, visible, audible et jouable. Son impact est énorme quant à la perception du joueur par rapport au jeu. L'interface utilisateur est la couche située entre le joueur et les éléments internes du jeu. Elle comprend les éléments visuels et audio du jeu mais aussi ceux de contrôle par le joueur pour illustrer les conséquences de ses actions du monde réel dans le jeu. Par exemple lorsque le joueur appuie sur un bouton. Parmi les éléments visuels et audio on cite, la perspective (première personne ou troisième personne), le paysage, les indicateurs (cartes), la bande sonore, les effets sonores, les vibrations. C'est pourquoi il est intéressant dans notre recherche de comprendre quels sont les parties précises de l'interface que le joueur regarde pendant qu'il vit son expérience vidéoludique. D'où l'introduction des techniques d'oculométrie dans cette étude.

2.3 Comment mesurer les émotions ?

La reconnaissance des états émotionnels des utilisateurs est une des composantes les plus importantes de l'informatique affective et celle ayant fait l'objet du plus grand nombre de travaux depuis les 15 dernières années. La littérature dans ce domaine contient d'ailleurs une très grande variété de techniques. Il existe de nos jours plusieurs outils et techniques de mesure des émotions chez un être humain. On distingue les mesures objectives des mesures subjectives.

2.3.1 Les mesures subjectives

Les mesures subjectives consistent à l'auto-évaluation des émotions de la personne, les utilisateurs notent eux-mêmes leurs émotions. Différents types de mesures subjectives sont

utilisables : les questionnaires et échelles, les entretiens. À titre d'exemple célèbre nous pouvons citer le *Self-Assessment Manikin* (Bradley & Lang, 1994), présenté sur la figure 2.

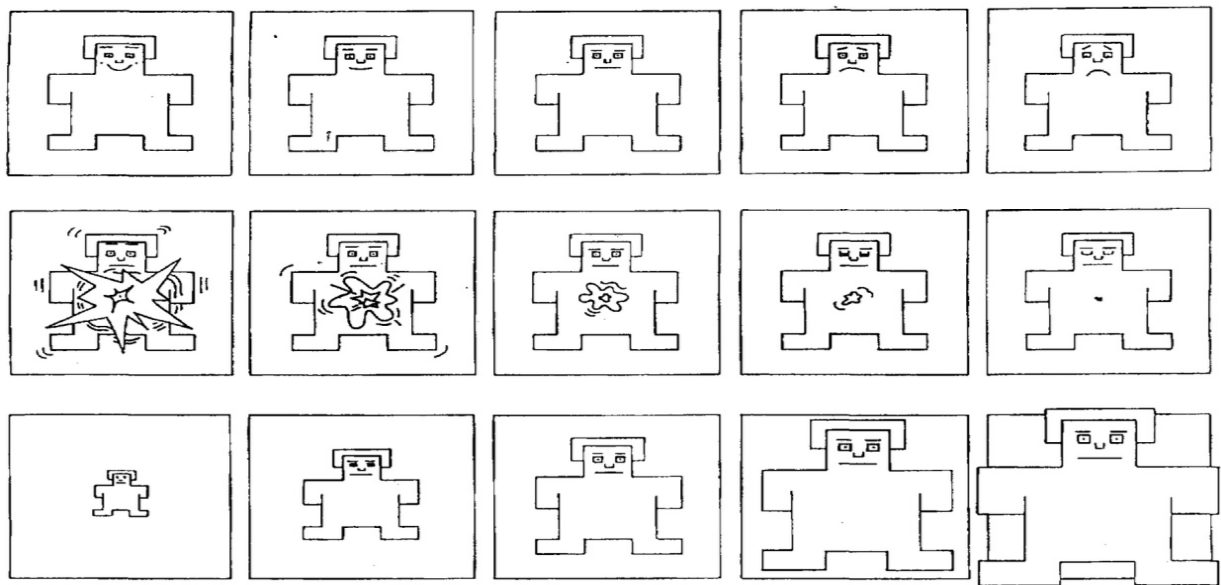


Figure 2. Le *Self-Assessment Manikin* utilisé pour évaluer les dimensions affectives respectivement valence, excitation et dominance

2.3.2 Les mesures objectives

Les travaux utilisent souvent plusieurs de ces méthodes. Les mesures objectives consistent à capturer et analyser les signaux émis par le corps du joueur sans prendre en considération son avis ou jugement. Plusieurs signes du corps humain témoignent de la présence d'émotions. Le ton de la voix change, le rythme cardiaque s'accélère, la sueur apparaît, les muscles du visage traduisent l'émotion ressentie. Les émotions sont également déterminables par des signaux physiologiques tels que les impulsions électriques émises par le cerveau, la conductance de la peau, le rythme cardiaque.

La reconnaissance des émotions par la voix

La reconnaissance automatique des émotions par la voix est une approche qui a pris une certaine ampleur ces dernières années avec l'avènement de systèmes qui proposent des interactions vocales associées à de l'intelligence artificielle. La majorité des travaux sont basés sur l'extraction des caractéristiques de communication non verbale qui aide à renforcer le

langage comme le ton, le rythme. Les travaux de (Lee & Narayanan, 2005) ont montré que la combinaison de ces caractéristiques à l'information du langage (les mots) permet d'obtenir de bons résultats de classification des émotions de 40,7% pour les hommes et 36,4% pour les femmes. Les caractéristiques peuvent également être spectrales notamment avec le MFCC (Mel Frequency Spectral Coefficient) comme l'a montré cette étude de (Kwon, Chan, Hao, & Lee, 2003).

La reconnaissance des émotions par l'analyse des signaux physiologiques

Procéder à une analyse du système nerveux autonome (Autonomic Nervous System en anglais) permet d'identifier les émotions d'une personne. La plupart des activités mesurées concernent les réponses électro-dermiques et cardiovasculaires comme le rythme cardiaque, la pression sanguine, la conductance de la peau et la résistance périphérique totale (Mauss & Robinson, 2009). Cependant, ces activités ne sont pas exclusivement reliées au processus émotionnel, de ce fait il est incertain de savoir si elles reflètent les émotions ou un processus du ANS (Stemmler, 2004).

En dirigeant l'analyse des signaux physiologiques vers les ondes électriques du cerveau par un EEG (électroencéphalogramme), les chercheurs ont prouvé leurs relations aux émotions d'une personne. En effet cette étude faite par (Benlamine, Chaouachi, Frasson, & Dufresne, 2016) montre que les expressions faciales sont reconnaissables par les signaux physiologiques provenant d'un EEG avec une précision élevée. En utilisant des techniques d'apprentissage machine ils ont réussi à identifier les expressions faciales évaluées par le EEG et faire le lien avec les émotions ressenties. En plus de rendre les émotions mesurables, cette étude a permis d'identifier avec précision les électrodes qui entrent en considération pour la mesure de chaque émotion.

La reconnaissance des émotions par les expressions faciales

Les expressions faciales sont les mouvements des muscles du visage. Charles Darwin a émis la théorie selon laquelle l'utilisation des expressions faciales était une caractéristique habituelle et ignorée qui était connectée aux processus émotionnels et de communication (Darwin, 1872). Depuis les recherches dans le domaine se sont multipliées en confirmant et élargissant la recherche faite par Darwin. (Ekman & Friesen, 1971) ont fait une classification




des 6 émotions de base ainsi que leurs expressions faciales associées. (Ekman, 1972) va plus loin en suggérant que les expressions faciales des émotions sont universellement reconnaissables malgré l'influence du facteur culturel. En effet, elles donnent des informations sur les émotions, l'humeur et le comportement interpersonnel d'une personne. En général, l'analyse des expressions faciales est importante pour mieux comprendre les êtres humains, leurs actions et leurs comportements (De la Torre & Cohn, 2011).





Les méthodes existantes dans le domaine utilisent l'électromyographie faciale (fEMG) (facial electromyography en anglais), le "Facial Action Coding System" (FACS) et les logiciels d'analyse automatique des expressions faciales. En ce qui concerne la fEMG, il s'agit de l'utilisation d'électrodes reliées aux muscles du visage autour des sourcils, de la bouche et des pommettes pour détecter les impulsions électriques. Cette méthode est précise mais nécessite une grande quantité d'équipements et n'est pas commode du fait des capteurs placés sur le visage.

Le FACS est un système de classification créé par (Ekman & Friesen, 1978) qui donne aux experts humains les outils nécessaires pour décomposer les expressions faciales en "Action Unit" (AUs). Ces unités font état des plus petits mouvements du visage qui peuvent être visuellement isolés par un humain. Ce système est une méthode fiable, non intrusive et précise pour analyser les expressions faciales. L'inconvénient principal est qu'il s'agit d'un processus très coûteux en temps. L'analyse d'une minute d'une vidéo peut prendre jusqu'à 100 minutes pour un expert bien entraîné (iMotions, 2016).

Allant au-delà de ces limitations, la nouvelle génération de technologie d'expression faciale se veut automatique et informatisée. En effet l'analyse se fait par un système qui utilise une caméra pour reconnaître les traits du visage et fait la reconnaissance automatique des émotions de base en temps réel. De tels systèmes sont performants comme illustré par (Littlewort, Bartlett, Fasel, Susskind, & Movellan, 2006) et plus tard avec le CERT (Littlewort et al., 2011). Le tableau suivant fait une description des muscles du visage impliqués dans la génération d'émotions que Darwin considère universelles (Lewis, Haviland-Jones, & Barrett, 2010).

Tableau 2. Description des muscles du visage en fonction des émotions

Émotion	Muscles du visage (description de Darwin)	Illustration
Colère	Les narines soulevées, la bouche comprimée, le front plissé, les yeux grands ouverts, la tête haute.	
Mépris	Fermeture partielle des paupières, les yeux se détournent, la lèvre supérieure relevée, les lèvres saillantes et le nez froncé.	
Dégout	La lèvre inférieure tournée vers le bas, la lèvre supérieure relevée, expiration, la bouche ouverte, cracher, souffler avec les lèvres en avant, la lèvre inférieure et la langue poussées vers l'avant.	

Peur	Les yeux ouverts, la bouche ouverte, les lèvres rétractées et les sourcils levés.	
Joie	Les muscles, zygomatique et orbiculaire sont contractés, la lèvre supérieure est relevée, Les rides du sourire ou sillons nasogéniens se forment.	
Tristesse	Les coins de la bouche descendent, les coins internes des sourcils se soulèvent.	
Surprise	Les sourcils levés, la bouche ouverte, les lèvres saillantes.	

Dans notre étude nous collectons les données provenant de 5 capteurs faisant intervenir la reconnaissance des expressions faciales, les signaux physiologiques et les mesures subjectives par l'utilisation d'un questionnaire d'auto-évaluation des émotions.

2.4 Revue de littérature et autres travaux sur le sujet

Les travaux que nous présentons sont en liens étroits avec notre sujet. Ils s'intéressent aux concepts d'éléments générateurs d'émotions dans les jeux vidéo, de mesure de données biométriques/physiologiques et de suivi du regard.

Why we play games: Four keys to more emotion without story. (Lazzaro, 2004)

L'équipe de XEODesign¹ a dirigé son intérêt vers la recherche de déclencheurs d'émotions autres que l'histoire dans les jeux vidéo. Pour cela ils ont mené leur recherche sur 45 personnes (15 joueurs extrêmes, 15 joueurs modérés et 15 non joueur) et 37 jeux vidéo de 5 genres différents. Ils ont collecté 3 types de données : les enregistrements vidéo de ce que les joueurs ont fait et dit, les réponses aux questionnaires soumis et les indices émotionnels verbaux et non-verbaux pendant le jeu. Leurs résultats montrent qu'il y'a 4 éléments clés dans les jeux vidéo qui créent des émotions, le *hard fun* (challenge, stratégie, et résolution de problèmes), le *easy fun* (intrigue et curiosité), les états modifiés (excitation ou soulagement des pensées), le facteur humain (les expériences sociales des jeux en groupe).

Using physiological measures in conjunction with other usability approaches for better understanding of the player's gameplay experiences. (Mirza-Babaei & McAllister, 2011)

Cette étude propose une méthode qui se base sur l'utilisation de données physiologiques pour identifier les moments pertinents dans un jeu qui suscitent un impact émotionnel important chez les joueurs. En combinant cette méthode à des méthodes d'évaluation d'utilisabilité comme les interviews post-expérience et des discussions avec les joueurs, elle aide à résoudre certains problèmes concernant l'expérience utilisateur. Ils ont mené leur étude sur 6 participants masculins âgés entre 21 et 30 ans qui ont joué à 2 jeux de tirs à la première personne, *Call of*

¹ <http://www.xeodesign.com>

duty Modern Warfare 2 et *Haze*. La plateforme de jeu était la console Playstation 3 de Sony. Les participants n'avaient jamais joué aux jeux auparavant. Ils devaient jouer les deux premières parties des jeux. Une moitié jouait à *Call of Duty Modern Warfare 2* et l'autre à *Haze*. Les interviews ont été faites dès après les sessions de jeu. Elles étaient basées sur des événements précis du jeu. Ces événements ont été choisis en fonction de l'observation des changements physiologiques des participants, leur rapports verbaux, leur position assise, leur expressions faciales et l'affichage du jeu. Les données physiologiques ont été collectées en utilisant les capteurs et logiciels de la suite BIOPAC². Ces données sont la conductance de la peau (GSR) et l'électrocardiographie (EKG). Le rythme cardiaque a été calculé à partir de l'EKG. Leurs résultats identifient quels sont les types d'évènements qui affectent l'expérience du participant.

- Les moments qui se répètent comme le début et la fin d'une mission ou la mort du personnage.
- Les messages d'instructions ou de direction dans le jeu.
- Les cinématiques.
- Les armes préférées des joueurs.
- Les nouveautés du jeu par rapport aux anciens.
- La jouabilité.
- L'environnement du jeu.
- Les endroits qui permettent de se détendre avant ou pendant les combats.

Understanding User Interaction in a Video Game by using Eye Tracking and Facial Expressions Analysis. (Iqbal, 2015)

Le but de cette étude est de réaliser un test d'utilisabilité d'un jeu vidéo en utilisant deux données, le suivi du regard et l'analyse des expressions faciales. Cette étude a été menée sur 8 participants répartis équitablement sur 4 catégories de joueurs, de novice à experts. Le jeu choisi était *Call of Duty Modern Warfare 2* sur PC. L'expérience a été effectuée en utilisant le Tobii T60 eye tracker comme technologie de suivi du regard. Des enregistrements vidéo de l'écran du

² www.biopac.com

joueur et du joueur ont été capturés par une webcam placée au-dessus du Tobii T60 eye tracker avec le logiciel Tobii Studio. Les participants devaient compléter une liste de tâches au cours du jeu. Leurs résultats concernant l'analyse du suivi du regard ont montré que les novices et joueurs moyens portent leur attention sur la partie centrale de l'écran de jeu et ils portent attention aux autres parties de l'écran uniquement lorsqu'ils sont coincés. À l'opposé des joueurs avancés et experts qui observent tout l'écran dès le départ. Leur analyse émotionnelle a montré que les novices ressentaient 82% de dégoût, les joueurs de niveau moyen 46% de joie, les experts 47% de joie.

Le suivi du regard dans les jeux vidéo

(SMI, 2011) La compagnie de recherche en utilisabilité Key Lime Interactive³ en partenariat avec l'éditeur de jeu THQ Inc⁴ a fait une étude d'utilisabilité d'un jeu en 3D par le suivi du regard. THQ a fait le test d'utilisabilité à base du protocole 'Think-aloud' et d'observations. Ils ont uniquement trouvé que les joueurs rencontraient des challenges mais ne comprenaient pas pourquoi ils les rencontraient. De ce fait, Key Lime Interactive a refait le test en utilisant leur propre technologie de suivi du regard. L'expérience a été faite sur 8 participants avec un jeu de tir à la troisième personne également appelé jeu de tir en vue objective. C'est un type de jeu dans lequel le personnage jouable est vu de manière externe (en intégralité) et dont le gameplay consiste principalement à tirer. Chaque partie a duré 60 minutes. L'analyse était basée sur les zones d'intérêts mise en évidence pour analyser le nombre de fois où chacune d'elles étaient visitées et pendant combien de temps. Le système de suivi du regard à distance SensoMotoric instruments (SMI) RED a été utilisé pour l'étude. Ils ont trouvé 3 éléments.

- Les données du suivi du regard ont été classifiées en fonction des repères visuels et des objets 3D dans le jeu.
- Le temps moyen passé par les joueurs sur une zone d'intérêt mise en évidence.
- Le suivi du regard a aidé les concepteurs du jeu à replacer et refaire les éléments du jeu pour les rendre plus efficaces.

³ www.keylimeinteractive.com

⁴ www.thq-games.com

(Tobii, 2015) La compagnie allemande Guerrilla Games a utilisé le suivi du regard comme outil pour tester l'expérience utilisateur de leur jeu *Killzone 3*. Il s'agit d'un jeu de tir qui prend place sur une planète du futur. Avec plusieurs millions d'euros investi, s'assurer de la qualité et de l'expérience utilisateur du jeu sont des éléments cruciaux. Ils ont travaillé avec une entreprise d'experts en utilisabilité nommé Valsplat. Leurs expériences ont été menées sur 6 à 8 participants hommes âgés entre 18 et 30 ans qui ont joué le jeu au complet pendant que 4 observateurs regardaient les indicateurs de confusion, d'ennui ou de frustration. Chaque test a duré environ 8 heures divisées en sessions de 3 heures de jeu avec des questionnaires. Ils ont utilisé le Tobii T60 XL comme technologie de suivi du regard et le jeu était sur la plate-forme Playstation 3 de Sony. L'expérience concernait les aspects individuels et coopératif du jeu. Leurs résultats ont montré 4 éléments principaux concernant la jouabilité du jeu vidéo qui n'auraient pas été accessibles sans le suivi du regard.

1. Trouver son chemin n'est pas toujours une tâche évidente pour le joueur. Le suivi du regard révèle ce que le joueur regarde et non ce qu'il devrait regarder. Cette information permet aux développeurs d'ajuster les objets qui détournent l'attention du joueur.
2. Beaucoup d'efforts et de temps sont investis sur des éléments qui rendent l'environnement plus animé. Le suivi du regard illustre quels sont ceux vus par le joueur et, plus important, ceux non vus. Le développeur peut alors améliorer leur disposition.
3. Les instructions de jeu sont essentielles pour un joueur. Le suivi du regard permet aux développeurs de trouver des méthodes qui permettent d'afficher des informations aux joueurs de manière optimale.
4. L'observation des interactions des participants dans le cadre d'une partie coopérative permet de savoir si le joueur regarde l'écran de son partenaire et quels sont les éléments qu'il regarde.

2.5 La prédiction des émotions dans un jeu

La prédiction des émotions des joueurs dans un jeu est une piste de recherche très récente. De ce fait il y a très peu d'œuvres qui existent sur le sujet. Les travaux s'intéressent à mesurer les émotions des joueurs en temps réel afin de produire des environnements adaptatifs. Comme le montre cette étude (Shaker, Yannakakis, & Togelius, 2010). Cette étude montre que des niveaux personnalisés peuvent être automatiquement générés pour des jeux de plate-forme. Le but de cette étude étant d'adapter de manière dynamique en temps réel les paramètres de génération de niveau en se basant sur les modèles de l'expérience du joueur avec pour objectif d'optimiser l'expérience de jeu du joueur. Pour se faire il est nécessaire de prédire les émotions des joueurs. Ils ont collecté 3 types de données sur 327 joueurs. Les données collectées concernent les aspects du jeu, le style de jeu du joueur, ses habilités de jeu et son expérience de jeu. Dans un premier temps ils ont utilisé un perceptron à une couche pour prédire les préférences du joueur en fonction des caractéristiques qui maximise cette prédiction. Ils ont prédit le fun, la frustration, le challenge à 64%, 84.7% et 70% respectivement. Dans un second temps ils ont évolués la topologie et les poids des perceptrons multicouches pour correspondre aux préférences prédites. Ils ont été capables de prédire les états émotionnels des joueurs avec une grande précision et une bonne adaptation du jeu aux caractéristiques de jeu du joueur.

En ce qui concerne l'aspect de prédiction des émotions, nous l'abordons dans notre étude suivant deux approches :

- La première consiste à utiliser les informations propres à un joueur ou à une scène en particulier pour effectuer la prédiction.
- La seconde, consiste à prédire les émotions d'un joueur en se basant sur l'historique des émotions exprimées par tous les joueurs dans toutes les scènes du jeu.

Les scènes sont décrites par un procédé qui s'appuie un modèle théorique d'évaluation de l'émotion tandis que les informations du joueur proviennent de formulaires qui sont remplis par les joueurs avant la session de jeu. L'apprentissage est fait sur ces informations avec plusieurs algorithmes testés. Nous évaluons la performance des algorithmes d'apprentissage machine avec la méthode de sélection de modèle par des techniques de validation croisée afin

de choisir celui qui offre une meilleure précision. L'intégration de cet aspect dans notre application implique 2 aspects :

- La prédiction de l'émotion d'un joueur par rapport à un objectif émotionnel d'une scène.
- La prédiction de l'émotion d'un joueur par rapport à une scène déjà regardée par d'autres joueurs.

Nous reviendrons plus en profondeur sur ce processus dans le chapitre 7.

CHAPITRE 3 L'ANNOTATION AFFECTIVE DES ÉLÉMENTS DANS UN JEU

3.1 L'analyse de la structure des jeux vidéo

Avant d'entrer en profondeur dans la structure des jeux vidéo, il est nécessaire de savoir qu'est-ce qu'un jeu et comment il fonctionne. (Adams, 2009) dans son livre définit un jeu comme un type d'activité ludique, menée dans le cadre d'une prétendue réalité, où les participants tentent au moins d'atteindre un but arbitraire, non trivial en agissant conformément aux règles. Les éléments essentiels d'un jeu sont le divertissement, la simulation, le but à atteindre et les règles.

Autant les personnes sont différentes autant elles prennent du plaisir à partir de choses différentes. C'est pourquoi les concepteurs de jeux vidéo sont capables de faire des jeux qui divertissent de plusieurs façons. Le gameplay est l'élément principal responsable de l'aspect divertissant d'un jeu vidéo. Nous avons vu dans le chapitre 2 que l'aspect esthétique jouait un rôle clé dans la création d'émotions. Les considérations esthétiques vont au-delà du monde du jeu. L'interface graphique complète ces considérations pour créer une belle expérience de jeu. Il s'agit des boutons, des images, du style de l'écriture etc. Les animations avec lesquelles le jeu répond sur la pression d'un bouton sont considérées comme éléments esthétiques. Peu importe le degré d'innovation ou de technologie un "vilain" jeu vidéo est un mauvais jeu vidéo. D'autres éléments entrent en compte dans le divertissement comme l'harmonie, l'histoire, les risques et récompenses, la nouveauté, la progression, l'aspect social avec le mode multijoueur.

Pour fournir du divertissement un jeu vidéo est composé de 2 principaux éléments.

- 1- Les mécanismes de base: il s'agit d'un modèle symbolique et mathématique qui représente les règles générales du jeu et que l'on peut implémenter algorithmiquement. Elles sont le cœur du fonctionnement du jeu car c'est elles qui génèrent le gameplay. Elles définissent les challenges que le jeu offre et les actions que les joueurs peuvent faire pour rencontrer ces challenges. Également elles déterminent quels sont les effets des actions du joueur dans le monde virtuel et établit les conditions de réussite ou d'échec du jeu. Dans un jeu vidéo ces mécanismes sont

cachées aux joueurs. Une qualité des mécanismes de base est le degré de réalisme. Le réalisme symbolise une simulation des situations du monde réel.

- 2- L'interface utilisateur : ce concept est familier pour un lecteur dans le domaine de l'informatique mais dans un jeu vidéo il joue un rôle plus complexe. L'interface utilisateur joue le médiateur entre ces mécanismes du jeu et le joueur comme l'illustre la figure 3. Elle prend les défis générés par les mécanismes de base et les transforme en images et sons pour le joueur. Elle prend en charge aussi les boutons appuyés sur la manette ou le clavier par le joueur et les convertit en actions associées au contexte du jeu. Le joueur finit par associer les boutons aux actions naturellement et n'a plus à penser " je dois presser le Y pour sauter" mais plutôt " je dois sauter" et appuie instinctivement le bouton Y de la manette ou du clavier. L'interface utilisateur interprète la pression sur le bouton comme une action de saut et informe les mécanismes de base ou du noyau. Celles-ci déterminent l'effet du saut et envoient l'instruction à l'interface utilisateur en lui indiquant quel résultat afficher au joueur.

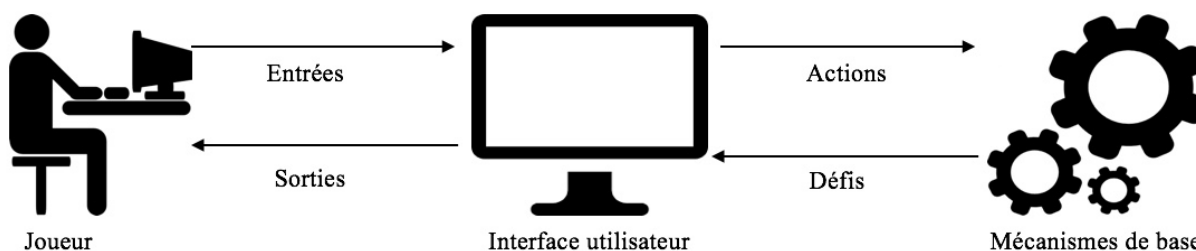


Figure 3. Les relations entre le joueur, l'interface utilisateur et les mécanismes de base dans un jeu.

Toutefois pour qu'un jeu soit cohérent, les défis et actions disponibles pour le joueur doivent être conceptuellement liés peu importe le moment. Par exemple au cours d'un combat au corps à corps dans un jeu de combat comme *Naruto Shippūden Ultimate Ninja Storm 4* le joueur peut se déplacer, attaquer, se protéger. Il ne doit pas être capable d'ouvrir la carte du monde ou son inventaire d'objets. Pourtant il s'agit d'actions qu'il peut réaliser dans un autre contexte du jeu. La détermination de quand et comment surviennent les changements de gameplay et l'adaptation de l'interface utilisateur se fait par la structure du jeu. La structure de jeu est faite principalement des modes de gameplay. Le concept de modes de gameplay est

central dans la conception de jeu vidéo. Il consiste à présenter un sous ensemble particulier de tout le gameplay du jeu qui est disponible à un moment précis. C'est l'interface utilisateur qui présente ce sous ensemble au joueur. La figure 4 reprend la figure 3 pour illustrer ce concept en montrant les intervenants qui sont l'interface utilisateur et le gameplay du jeu.

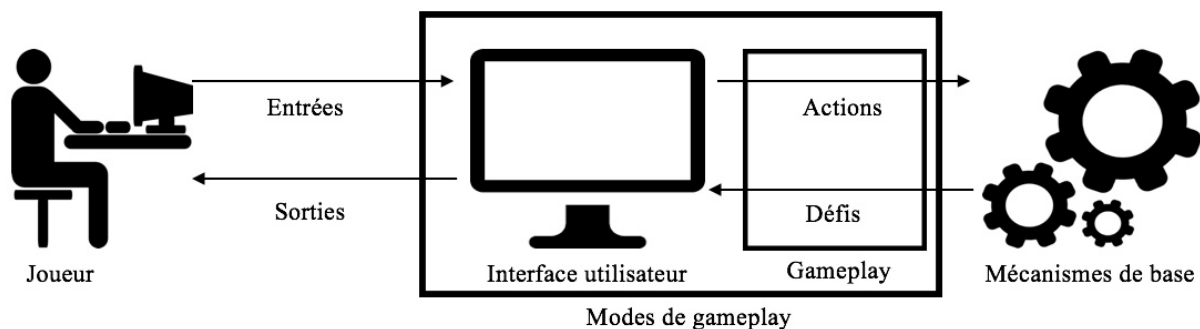


Figure 4. Illustration du concept de modes de gameplay

Le contenu visuel étant alors un élément central dans la structure d'un jeu vidéo et responsable des réactions et comportements du joueur, nous proposons dans notre étude l'intégration de l'analyse des mouvements oculaires du joueur afin d'investiguer en profondeur sur les éléments visuels du jeu déclencheurs d'émotions.

3.2 L'oculométrie

Les yeux et leurs mouvements jouent un grand rôle en termes de communication. Le contact visuel ou la direction du regard sont des sources d'informations dans la communication humaine qui permettent entre autres d'indiquer la cible de notre intérêt visuel (Kleinke, 1986).

L'oculométrie est une technique pour suivre la trace du regard d'une personne. Avec les technologies actuelles il est relativement simple d'observer les endroits qu'un utilisateur a regardé sur un écran. Les équipements d'oculométrie peuvent être montés sur un écran et le logiciel permet de garder une trace de ce qui est affiché à l'écran pendant que l'utilisateur le regarde (Nielsen & Pernice, 2010). Comme les paramètres de calibration sont différents d'un utilisateur à l'autre alors l'équipement a besoin d'être calibré avant chaque utilisation. Le progrès en oculométrie est remarquable et son utilisation est pertinente dans le domaine du jeu

vidéo comme le montre les études dans la section 2.4 afin d'améliorer l'expérience utilisateur du joueur.

L'utilisation des zones d'intérêt (Area Of Interest) comme représentation des mouvements oculaires du joueur permet d'en extraire une valeur sémantique. Elles permettent de délimiter les parties d'une interface ayant un contenu intéressant et d'en extraire les métriques correspondantes. Les zones d'intérêt ont un contenu sémantique et une organisation logique en fonction des actions dans le jeu. La majeure partie des outils permettent de créer facilement des AOI et de les modifier selon la vidéo. Elles deviennent des AOI mobiles (Moving AOI) et nous permettent de suivre le parcours du joueur dans le temps en fonction des événements ou de l'évolution du jeu. Celles-ci sont définies manuellement par le chercheur.

Dans notre étude nous voulons identifier les objets du jeu regardés par le joueur afin d'en extraire leur effet émotionnel. Nous utilisons comme équipement l'oculomètre sur écran Tobii Pro TX300⁵ et les AOI sont marquées par le logiciel iMotions⁶. Nous présenterons plus en détails le procédé dans les chapitres 4 et 5.

3.3 Les types d'émotions

Les émotions sont des composantes fondamentales du genre humain. La définition des émotions est souvent sujette à de nombreuses polémiques, et il n'est pas exagéré de considérer qu'il y'a autant de définitions que de chercheurs travaillant sur le sujet. La littérature actuelle se limite à la reconnaissance d'émotions "simples" comme la joie, la colère tandis que les émotions plus complexes comme la frustration, la honte, sont le résultat de plusieurs activités cognitives de plus haut niveau, les rendant difficilement reconnaissable par l'analyse du système nerveux périphérique. Les auteurs des travaux sur l'interface utilisateur se réfèrent à trois principales théories: l'approche cognitive, l'approche multidimensionnelle et l'approche catégorielle (Gauducheu, 2009).

⁵ <http://www.tobii.com/product-listing/tobii-pro-tx300/>

⁶ www.imotions.com

L'approche cognitive

Selon l'approche cognitive, l'émotion est une réaction suite à l'évaluation d'évènement stimulus interne ou externe considéré comme central/pertinent pour la personne (Scherer, 2001). De ce fait des personnes différentes peuvent ressentir des émotions différentes en fonction des évènements. Le type et l'intensité de l'émotion suscitée sont le résultat d'un processus cognitif d'évaluation des stimuli. Cette évaluation porte sur les dimensions suivantes : la pertinence de l'évènement, les éléments comme la nouveauté de l'évènement et la valence (positive ou négative) sont impliqués dans cette dimension. L'implication de la personne par rapport à l'évènement c'est à dire que les causes et conséquences doivent être définies, le lien avec les buts et besoins. La possibilité de faire face à l'évènement (contrôle) et l'importance de la compatibilité avec les normes culturelles et personnelles (Scherer, 2001). Le résultat de cette évaluation s'exprime à travers différentes composantes de l'émotion : les réactions physiologiques, les sentiments, les expressions faciales, etc. Les théories par rapport à l'évaluation cognitive peuvent être utilisées pour générer un modèle ou prédire les émotions d'un utilisateur en temps réel (C. Elliott & Brzezinski, 1998).

L'approche multidimensionnelle

L'approche multidimensionnelle, correspond à un ensemble de modèles qui tentent de conceptualiser l'expérience émotionnelle humaine sur 2 ou 3 dimensions. Plusieurs auteurs se situent dans cette perspective mais il existe des désaccords sur les dimensions retenues et leur nombre. Wilhelm Max Wundt, le père de la psychologie moderne, a proposé que trois axes principaux se distinguent: "agréable par rapport désagréable/déplaisant", "excitation ou déprimant" et "stress ou relaxation" (Wundt, 1907). Pour Russel, développeur du modèle circomplexe, les états émotionnels peuvent être caractérisé par le caractère plaisant/déplaisant(valence) de l'expérience subjective, et le degré d'excitation(degree-of-arousal en anglais) (J. A. Russell, 1980). La figure 6 présente un exemple du modèle circomplexe, il s'agit d'une solution pour la représentation de 28 mots affectifs à l'échelle multidimensionnelle. La valence étant sur l'axe horizontal et l'excitation sur celui vertical. L'évaluation en termes de plaisir(valence) et excitation(arousal) correspondrait en fait à l'affect fondamental primaire (*core affect* en anglais) (J. A. Russell, 2003). L'affect fondamental est universel et son expérience subjective est simple, un individu est toujours dans un tel état

affectif, d'intensité variable. Il correspond au ressenti affectif le plus élémentaire accessible à la conscience. La plupart des modèles tridimensionnels incorporent valence et excitation ou intensité. Par exemple les modèle de vecteur d'émotion (Bradley, Greenwald, Petry, & Lang, 1992) et l'activation positive – activation négative (PANA) (Watson & Tellegen, 1985). La définition des émotions proposée par cette approche est très utilisée dans les travaux sur l'interface utilisateur (Isbister, Höök, Laakso, & Sharp, 2007) (Mandryk, Atkins, & Inkpen, 2006). Certains auteurs utilisent une seule dimension, par exemple, la valence (Hazlett & Benedek, 2007) ou l'activation (Rajae-Joordens, 2008).

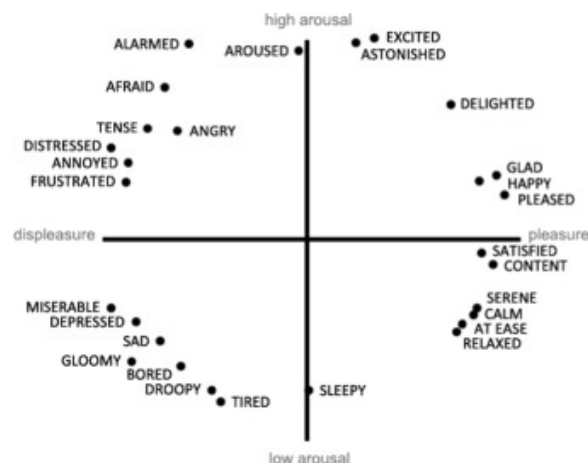


Figure 5. Exemple du modèle circomplexe tiré de (J. A. Russell, 1980).

L'intérêt des auteurs pour l'approche multidimensionnelle s'explique assez facilement puisque celle-ci permet de définir simplement l'expérience émotionnelle. Néanmoins, cette approche présente, de notre point de vue, certaines limites pour l'évaluation des systèmes interactifs. En effet, le contrepoint de cette simplicité est que les informations que l'on peut recueillir sur les émotions sont en fait assez ambiguës. La dimension plaisir/déplaisir, si elle est mesurée à la fin de la session d'utilisation du système, risque d'être assez proche de la dimension satisfaction de l'utilisabilité. Il sera en effet difficile de savoir si l'émotion a été suscitée par une évaluation globale du système, la réalisation de la tâche ou encore les dernières interactions de l'utilisateur avec le système. Par ailleurs, la dimension activation semble difficile à interpréter : quelle signification attribuer à un état d'excitation par rapport à un état calme ? Il risque donc d'être difficile de pouvoir tirer des conclusions des émotions observées pour la conception. Ce problème semble d'autant plus critique que nous ne disposons pas d'un modèle qui permettrait

d'expliquer pourquoi un individu déclare ressentir un affect positif, ou être plutôt dans un état d'excitation (et inversement), à la suite de l'utilisation d'un système interactif.

L'approche catégorielle

En ce qui concerne l'approche catégorielle, Paul Ekman le fameux psychologue américain et pionnier dans l'étude des émotions et leurs relations aux expressions faciales, considère qu'il existe des émotions dites de base, universelles, qui ont des caractéristiques spécifiques et auxquelles correspondent des configurations faciales et des composantes physiologiques spécifiques (Ekman, 1992). Comme mentionné dans le chapitre 2, (Ekman & Friesen, 1971) ont catégorisé 6 émotions de base qui sont la colère, la peur, la tristesse, la joie, le dégoût et la surprise comme le montre la figure 5. D'autres études ont montrées que le mépris est également universellement reconnu (Matsumoto, 1992). Ces émotions ne représentent pas des états affectifs singuliers mais plutôt des ensembles ayant des caractéristiques communes (Ekman & Friesen, 1975). Dans la littérature, on trouve essentiellement la référence à l'approche catégorielle dans les travaux qui mesurent les émotions à partir des expressions faciales comme le montre cette étude (Hazlett & Benedek, 2007). En effet, les auteurs utilisent le FACS (Facial Action Coding System), conçu par Ekman notamment, qui est un outil de codage des expressions faciales sur la base des mouvements des muscles du visage (Ekman & O'Sullivan, 1991).

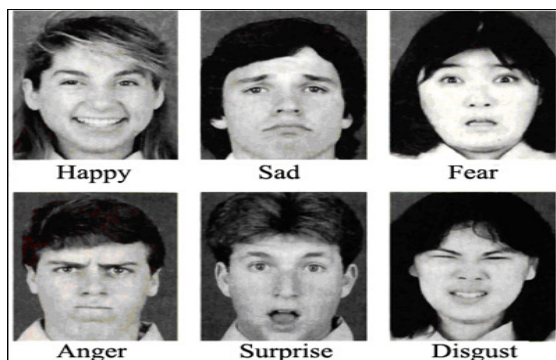


Figure 6. Un exemple des émotions de bases décrites par Ekman

La question du recueil des données ne dépend pas uniquement du choix d'une définition des émotions mais aussi du choix des méthodes employées. Par conséquent, il semble plus pertinent d'utiliser des théories qui permettent d'avoir une représentation plus fine de

l'expérience émotionnelle, en particulier de ses causes. Dans notre cadre, l'approche catégorielle avec les émotions *basiques* est la plus adaptée. Dans la mesure où elles existent et peuvent être identifiées mais aussi elles ont l'avantage de la similarité entre cultures et sont facilement discriminables. De ce fait il est plus accessible d'étudier les réponses émotionnelles des personnes vis à vis des éléments graphiques de l'interface utilisateur.

3.4 Le graphe émotionnel

En combinant les techniques d'oculométrie et de classification des émotions du joueur par la reconnaissance des expressions faciales selon l'approche catégorielle nous pouvons garder une trace des émotions le long de son parcours dans un jeu. En effet, tout d'abord nous extrayons les émotions pertinentes ressenties durant les étapes de la partie. Ensuite grâce au suivi du regard nous identifions les objets/scènes du jeu chronologiquement. Nous établissons une relation entre les éléments majeurs observés à un instant donné et les mesures d'émotions effectuées afin de trouver les variations des émotions les plus fortes dans le temps en fonction des zones d'intérêts. La figure 7 reprend la figure 3 afin de présenter où se situe l'étape de construction du graphe ainsi que les parties qui interviennent dans cette étape. Cette figure illustre la combinaison des techniques mentionnées plus haut.

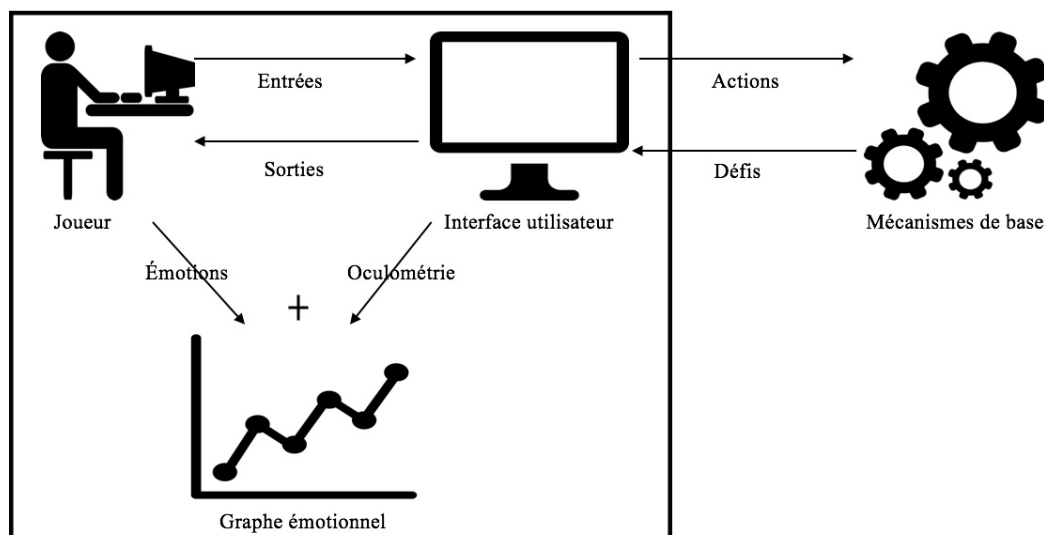


Figure 7. Construction de la représentation graphique

Dans ce mémoire nous nous focaliserons sur l'influence émotionnelle produite par l'interface utilisateur sur le joueur en trouvant les déclencheurs d'émotions fortes. Outre cet objectif, nous représentons cette influence sous forme graphique dans le souci d'avoir une meilleure visualisation. Nous présenterons dans le prochain chapitre la phase d'expérimentation à partir de laquelle nous avons développé la méthode et l'application que nous verrons plus loin dans ce mémoire.

CHAPITRE 4 L'EXPÉRIENCE

Dans cette partie du mémoire, nous allons examiner les détails de l'expérience que nous avons réalisée tels que, le jeu choisi, les participants, les équipements utilisés. Rappelons que l'objectif de cette expérience est de collecter les données nécessaires au développement d'une méthode dynamique d'identification d'émotions fortes et de leurs déclencheurs. Cette méthode doit permettre de déterminer les émotions dominantes des joueurs ainsi qu'en fournir précisément l'/les élément(s) graphique(s) associé(s). L'outil nommé "Émograph" découle du développement de la méthode qui porte le même nom. Dans le prochain chapitre nous présenterons la méthode ainsi que l'outil.

4.1 Contexte de l'expérience

Nous avons mené l'expérience au laboratoire des Jeux Vidéo Intelligents (JVI) situé au local V-13-1 du pavillon Roger-Gaudry de l'Université de Montréal. Cette expérience implique la collecte des données à travers des capteurs physiologiques et le suivi du regard. De ce fait le jeu choisi doit avoir un impact fort sur l'état émotionnel d'un joueur afin d'avoir des valeurs caractéristiques utilisables pour atteindre l'objectif fixé. Il doit être capable d'influencer les réponses émotionnelles des participants et générer les expressions faciales de base pour les enregistrer et les analyser. L'expérience a été réalisée suivant un protocole expérimental établi sur une durée allant de 55 minutes à 1 heure. Il faut compter une période d'installation des capteurs et de démarrage du jeu de 15 minutes qui va servir aussi de phase de calibration. L'expérience débute une fois que tous les capteurs sont placés sur le participant, fonctionnent correctement et que le jeu est lancé. La phase d'expérience s'est déroulée sur 3 semaines. Nous avons mis à disposition des joueurs une plate-forme de jeu vidéo ; un ordinateur de bureau muni d'une webcam tournant sous Windows 7 sur lequel est installé et diffusé le jeu. Aussi un écran secondaire qui va servir au chercheur qui joue le rôle de superviseur pour surveiller le bon fonctionnement des équipements et le déroulement de l'expérience.

4.1.1 Outlast

Les jeux vidéo et les émotions sont intimement liés comme nous l'avons vu dans les sections précédentes. Étant donné que les jeux vidéo sont de genres différents (section 2.1) ils

génèrent des émotions différentes. Par exemple, il a été discuté dans ces études (Anderson & Dill, 2000) que jouer à des jeux violents était lié à l'accroissement des comportements agressifs chez les joueurs. Les résultats de l'étude de Bailey et West sur les jeux d'action montrent que ce genre de jeux a tendance à altérer les émotions et le traitement visuel. Ils ont un impact sur la transformation des expressions faciales positives (Bailey & West, 2013).

Dans notre cadre, nous avons choisi le jeu *Outlast* (la figure 8 montre la couverture du jeu) sur PC. Comme mentionné dans la section 2.1 Il s'agit d'un jeu d'horreur (survival-horror) à la première personne, développé et publié par la compagnie Red Barrels Games⁷, sorti sur Windows le 4 septembre 2013, via la plateforme de téléchargement Steam⁸. La perspective de vue à la première personne (ou vision subjective) désigne une façon d'apercevoir l'environnement de jeu comme si l'on incarnait l'avatar en voyant le jeu à travers ses yeux. *Outlast* permet aux joueurs d'incarner Miles Upshur, un journaliste d'investigation qui est envoyé pour trouver la 'vérité' au sujet la compagnie Murkoff Corporation. Il va se lancer dans l'exploration des activités de l'hôpital psychiatrique délabré de Mount Massive Asylum, lequel prétend être une œuvre de charité pour le soin et le traitement des maladies mentales. Le joueur va naviguer à travers un décor sombre tout le long du jeu et des scènes d'horreur à l'intérieur de l'asile. Un personnage non joueur (PNJ) malicieux, violent qui poursuit le journaliste et les seuls moyens de survie sont : ' fuir, se cacher ou ... mourir'' avec pour seule arme sa caméra. Le joueur peut passer par-dessus certains obstacles, se baisser, sauter et se faufiler à travers des espaces étroits. Il est toujours accompagné de sa caméra, qu'il peut activer ou non, sachant qu'elle consomme de l'énergie, des batteries de secours peuvent néanmoins être trouvées durant le jeu. Elle est utile pour mieux comprendre l'histoire via des notes prises par le journaliste, ainsi que pour son option de vision nocturne qui permet de voir dans certains lieux plongés dans le noir. Ce jeu positionne le joueur en tant que témoin, victime de violence et donne un sentiment d'impuissance. L'impuissance dans laquelle le joueur est placé est un élément essentiel à la structure narratrice du jeu et du gameplay, elle joue une grande part dans les expressions émotionnelles des joueurs.

⁷ <http://www.redbarrelsgames.com>

⁸ <http://store.steampowered.com/>

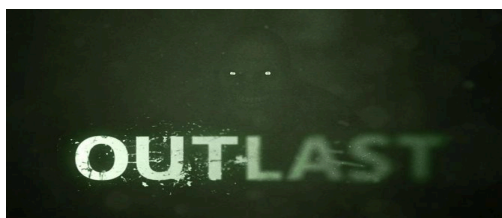


Figure 8. Couverture du jeu Outlast

Pour jouer au jeu, les touches du clavier et la souris sont les commandes. Les instructions sont données au fil du jeu en haut de l'écran. L'environnement de jeu étant vaste le joueur peut terminer la séquence en utilisant plusieurs stratégies en fonction de sa personnalité ou de son profil, mais il faut préciser que le circuit d'avancement du jeu est fixe. Cela signifie que le joueur ne peut emprunter qu'un seul chemin pour remplir les objectifs de la séquence choisie. Les étapes de la séquence choisie sont présentées dans le tableau 3. Elles ont été choisies à la main par nous en fonction des événements, actions ou challenges du jeu. Certains participants peuvent ne pas passer par toutes les étapes dépendamment du joueur.

Nous avons choisi la première séquence du jeu pour notre expérience. Elle consiste pour Miles à trouver un chemin pour entrer dans l'asile avant de débiter son enquête dans un lieu lugubre, dégradé et ensanglanté. Une des scènes choquantes dans cette séquence de jeu implique la découverte d'un agent de sécurité empalé qui dans son dernier souffle de vie l'avertit des choses étranges qui se déroulent dans l'asile et le somme de fuir immédiatement. Il s'agit de la scène nommée "Cadavre qui parle".

Tableau 3. Parcours de la séquence de jeu

Scène	Description	Actions/Challenges
Bureau à l'entrée (non obligatoire)	Situé à la barrière d'entrée de l'asile, il contient quelques dessins d'enfants et du mobilier de bureau.	Explorer le bureau situé avant les grilles d'entrée de l'asile.
Grille d'entrée	Grille permettant d'accéder à l'asile.	Ouvrir les grilles pour entrer dans la cour de l'asile.

Premier jardin	Grand jardin de l'asile avec des camions à l'entrée du manoir.	Traverser le jardin et trouver un chemin d'accès à l'asile.
Camions	Camions se situant dans le premier jardin, aux portes de l'asile.	Aucune action.
Porte fermée (premier jardin) - (non obligatoire)	Porte impossible à ouvrir bloquant l'entrée au manoir	Tenter d'ouvrir la porte et constater qu'elle est fermée.
Petit passage	Chemin de transition vers l'arrière-cour. Petit espace dans le mur	Se baisser pour passer dans le trou d'une grille qui conduit à l'arrière-cour de l'asile.
Deuxième jardin	Arrière-cour de l'asile avec une fontaine, une porte bloquée et un échafaudage	Explorer le jardin et utiliser l'échelle pour monter sur l'échafaudage.
Porte fermée (deuxième jardin) - (non obligatoire)	Porte métallique qui empêche l'entrée dans l'asile.	Tenter d'ouvrir la porte et constater qu'elle est fermée.
Échelle	Échelle qui donne accès à l'échafaudage.	Monter à l'échelle pour atteindre l'échafaudage, c'est le seul moyen d'entrer dans l'asile.
Salon obscur	Salon sans lumière avec des meubles en désordre et des taches de sang sur le sol.	Se déplacer dans le noir ou activer la vision nocturne de la caméra pour traverser le salon.

Salon 2 (tv)	Allumage brusque de la tv avec un grésillement et des meubles de salon.	Traverser le salon.
Sang couloir	Traces de sang sur les murs du couloir.	Traverser le couloir et enregistrer les traces de sang dans la caméra.
Bureau 1 (non obligatoire)	Mobilier de bureau, documents concernant les évènements sombres de l'asile	Explorer le bureau pour trouver des indices.
Bureau 2 (non obligatoire)	Mobilier de bureau, documents concernant les évènements sombres de l'asile, on y trouve des piles de rechange pour la camera.	Explorer le bureau pour trouver des indices.
Cuisine	Des traces de sang un peu partout, table à manger, un évier très sale et une conduite d'aération cassée.	Explorer la cuisine et accéder à la conduite d'aération.
Sang coulant de la conduite d'aération	Gouttes de sang dégoulinant de la conduite d'aération.	Enregistrer la scène dans la caméra.
Conduite d'aération	Conduite brisée et sale reliant la cuisine à la bibliothèque.	Passer par la conduite d'aération pour rejoindre la bibliothèque.
Bibliothèque	Salle obscure remplie d'étagères de livres et de morceaux de corps humains.	Explorer la bibliothèque.

Cadavre suspendu	Le cadavre surgit de l'obscurité suspendu par les jambes sur l'ouverture de la porte de la bibliothèque et tombe sur le sol.	Aucune action.
Cadavre qui parle	Un agent de sécurité de l'asile pendu au-dessus de morceaux de corps humain prend la parole.	Enregistrer la mise en garde de l'agent dans la caméra.
Porte fermée (fin) - (non obligatoire)	Porte par laquelle on aperçoit l'ombre du "variant" passer mais impossible à ouvrir.	Tenter d'ouvrir la porte pour suivre le variant.
Toilettes (non obligatoire)	WC cassés avec un membre humain à l'intérieur, lavabo crasseux et une pile de rechange pour la camera.	Explorer les toilettes et enregistrer les informations dans la caméra.
Monstre (PNJ)	Le monstre appelé "variant" dans le jeu attaque et projette le personnage par-dessus le balcon.	Aucune action.
Personnage alternatif (PNJ)	Un vieux monsieur qui réveille Miles après la chute.	Écouter le monsieur.
Sang fin séquence	Phrases écrites avec du sang	Enregistrer les phrases dans la caméra.

4.1.2 Les participants

L'expérience a été conduite sur 21 participants tous âgés entre 18 et 40 ans (moyenne = 24.05 et écart-type = 4.59) tous étudiants à l'Université de Montréal, la figure 10 montre la

distribution des participants de l'expérience. Les participants ont été recrutés à travers des annonces faites sur le campus. Ils ont signé un formulaire de consentement approuvé par le comité d'éthique de la recherche en arts et en sciences (CÉRAS) de l'université au préalable de l'expérience et ont reçu une compensation monétaire pour leur participation à l'expérience. Cet échantillon était constitué de 12 hommes et 9 femmes. Les données collectées sur les femmes n'étaient pas d'excellente qualité. Cela est dû au fait que celles-ci peuvent bouger le haut du corps ou la tête à la vue de certaines images du jeu ou au fait qu'elles ne soient pas à l'aise avec les commandes du jeu alors elles baissent la tête pour regarder le clavier. La figure 9 montre un exemple, la participante se couvre les yeux en voyant les têtes humaines sur les étagères. Compte tenu de la faible qualité des données collectées sur 3 participantes nous les avons extraites de l'échantillon. Également il y'a eu des problèmes techniques lors de l'enregistrement de la vidéo de jeu d'un participant. Nous n'avons pas pu l'annoter alors nous l'avons retiré de l'échantillon. Au final nous disposons des données valides pour 17 participants. Nous avons catégorisé les joueurs selon le nombre d'heures de jeu par semaine. De ce fait on dénombre, 5 joueurs extrêmes et 6 intermédiaires et 8 novices.



Figure 9. Exemple d'erreur dans la collecte des données.

Afin de recueillir le maximum d'informations pour compléter notre étude nous avons soumis des formulaires aux participants avant et après le jeu. Avant le jeu, il s'agit d'un formulaire de collecte des données sociodémographiques (l'âge, le genre, l'origine ethnique), du niveau scolaire ainsi que des heures de jeu par semaine. Après le jeu, deux formulaires sont soumis, le premier est le questionnaire "Big Five" pour l'évaluation des traits de personnalité du participant. Celui utilisé dans notre étude est basé sur le "Big-Five Factor Markers from the

International Personality Item Pool'' développé par (Goldberg, 1992). Nous donnerons plus de précisions quant à l'utilité de ce formulaire dans le chapitre 7. Le second est une auto-évaluation des émotions par les participants pendant des étapes de la scène du jeu.

Les données collectées sur les participants servent à dresser un profil du participant qui va permettre la prédiction des émotions que nous verrons dans le chapitre 7. Ces informations personnelles sont anonymes et protégées dans la base de données. Au cours des phases de calibration et d'expérimentation les participants ne doivent pas porter de chapeaux ou de lunettes de soleil ni avoir de gomme à mâcher dans la bouche ou boire au risque de rajouter du bruit dans les données.

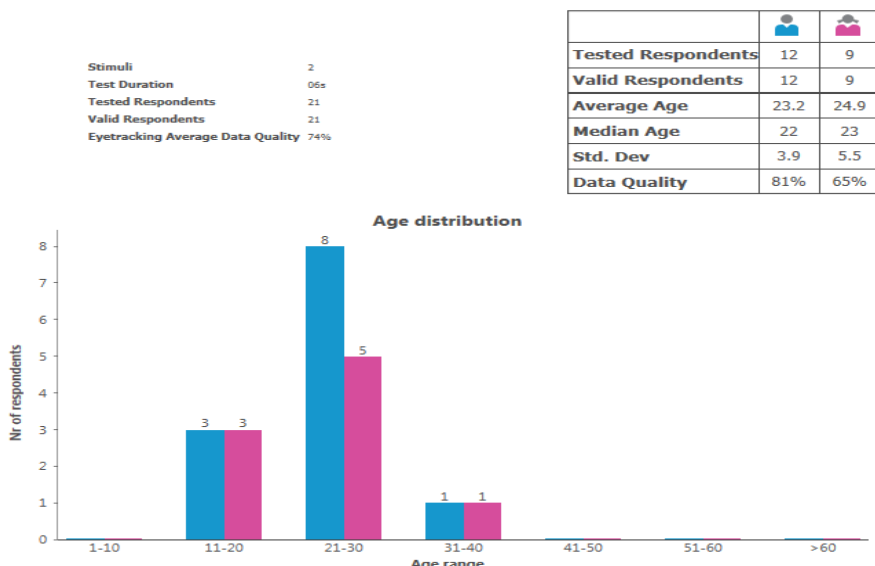


Figure 10. Distribution des participants

4.2 Mesure des émotions

Cette section aborde les outils que nous avons utilisés pour collecter les données physiologiques et réaliser l'oculométrie pendant l'expérience.

4.2.1 La plateforme de recherche biométrique iMotions

Les outils de suivi du regard, de reconnaissance faciale et d'EEG sont intégrés et synchronisés sur la plateforme iMotions. iMotions Biometric Research Platform est une solution logicielle facile à utiliser pour la conception d'études, l'intégration et la synchronisation de

plusieurs capteurs, la collecte et l'analyse de données. Cette plateforme supporte plus de 50 principaux capteurs biométriques, y compris l'analyse des expressions faciales, le suivi du regard, l'EEG ainsi que les technologies à base de formulaires pour la recherche multimodale sur le comportement humain. Elle est approuvée par les plus grandes institutions universitaires et industries comme Yale⁹, Stanford¹⁰ et Unilever¹¹, faisant fonctionner les labos de recherche à travers le monde. Pour une meilleure précision le logiciel passe par une étape de calibration qui permet de définir des valeurs de référence pour les senseurs intégrés. Cette plateforme est installée sur l'ordinateur utilisé pour l'expérience. Les résultats sont présentés sous forme de probabilité d'une expression faciale à un moment donné.

4.2.2 L'oculomètre Tobii Pro TX300

L'oculomètre Tobii Pro TX300 et le logiciel iMotions Attention Tool (version 6.0) ont été utilisés pour faire le suivi du regard. Cet oculomètre est encapsulé à un écran de 23 pouces Full HD, échantillonnant à 300Hz. Il peut également être utilisé en mode "barre", adaptable à des écrans encore plus larges ou de la vidéo projection. Cette solution est reconnue pour sa stabilité pour des études exigeantes et complexes basées sur le suivi du regard. L'unité de suivi du regard détecte automatiquement les yeux du participant. Celui-ci doit être placé dans le périmètre de détection de l'équipement. Les figures 11 et 12 montrent cet équipement.



Figure 11. Unité de suivi du regard

⁹ <http://www.yale.edu>

¹⁰ <http://stanford.edu>

¹¹ <https://www.unilever.ca/fr/>



Figure 12. Oculomètre Tobii Pro TX300

4.2.3 Le module de reconnaissance d'expressions faciales FACET

Pour analyser et reconnaître les réponses émotionnelles provenant du visage nous avons utilisé le moteur Emotient FACET qui est intégré dans la plateforme iMotions.

FACET (iMotions, 2013) est un logiciel d'analyse et de reconnaissance des expressions faciales. Il est le successeur de CERT. Il analyse les réponses émotionnelles des utilisateurs en temps réel, image par image. Il détecte et reconnaît les sept expressions faciales de base. Il fournit une grande précision de classification indépendamment de diverses conditions telles que le sexe, l'âge, les poils du visage, les lunettes, la position de la tête et l'éclairage, en raison du grand nombre de données dont il dispose. Le participant doit être placé face à la caméra, dans des bonnes conditions de luminosité et à une distance optimale. Pendant l'étape de calibration, les participants doivent être détendus, rester silencieux et neutre afin d'optimiser la précision de la reconnaissance faciale.

4.2.4 Le casque Emotiv ÉPOC

Nous avons utilisé le casque EEG Emotiv EPOC. C'est un casque électroencéphalogramme de haute-résolution qui comporte 14 capteurs permettant d'interpréter les signaux électriques produits par le cerveau humain et d'analyser en temps réel les pensées, sentiments et émotions de l'utilisateur. Pour installer le casque sur la tête du participant, il faut en premier lieu appliquer une solution saline sur les 14 capteurs pour permettre une bonne

conductivité des signaux électriques et ensuite le déposer tranquillement sur la tête du participant. Finalement, il faut s'assurer que tous les capteurs ont une bonne réception des signaux électriques, ce qui est fait à l'aide du panneau de contrôle fourni par la plateforme iMotions. La figure 13 montre le casque et la figure 14 illustre le bon placement et fonctionnement des électrodes sur la tête.



Figure 13. Le casque Emotiv EPOC

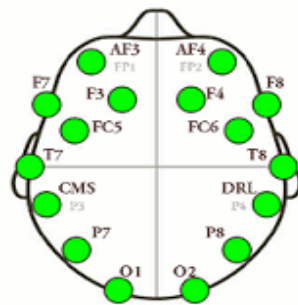


Figure 14. Le placement des électrodes sur la tête

4.2.5 Le bracelet Affectiva Q-sensor 2.0

Ce bracelet développé par la firme Affectiva¹² est un biocapteur portable, sans-fil qui mesure l'excitation émotionnelle via la conductance de la peau. Il s'agit de l'activité électrodermale (EDA) qui augmente pendant les états d'excitation, d'anxiété ou d'angoisse et diminue pendant les états d'ennui ou de relaxation. Il mesure également la température corporelle et l'activité. Il est relié par Bluetooth à l'ordinateur sur lequel est installé le logiciel Q-Live Software provenant de la même firme ; qui permet l'enregistrement et la visualisation

¹² <http://www.affectiva.com/>

des données provenant du capteur en temps réel. Nous le faisons porter au participant en premier car il requiert plus de temps pour sa calibration. La figure 15 présente le bracelet utilisé.



Figure 15. Le bracelet Affectiva Q-sensor 2.0

4.2.6 L'électrocardiogramme fourni par BIOPAC

Pour enregistrer le rythme cardiaque des participants, nous avons utilisé le système Biopac MP150 combiné au logiciel AcqKnowledge qui proviennent de la firme Biopac¹³. Il fonctionne sans-fil et le logiciel fait l'analyse automatique de l'électrocardiogramme afin d'en extraire le rythme cardiaque. Nous avons placé 3 électrodes sur le haut du corps du participant à des positions précises décrites dans le guide d'utilisation. Les électrodes sont reliées à un capteur sans fil qui communique avec le système Biopac MP150. Pour assurer un bon fonctionnement des électrodes nous avons appliqué sur eux un gel non irritant, hypoallergénique qui joue le rôle de conducteur. La figure 16 extraite du site web de la compagnie, montre le fonctionnement de ce système.

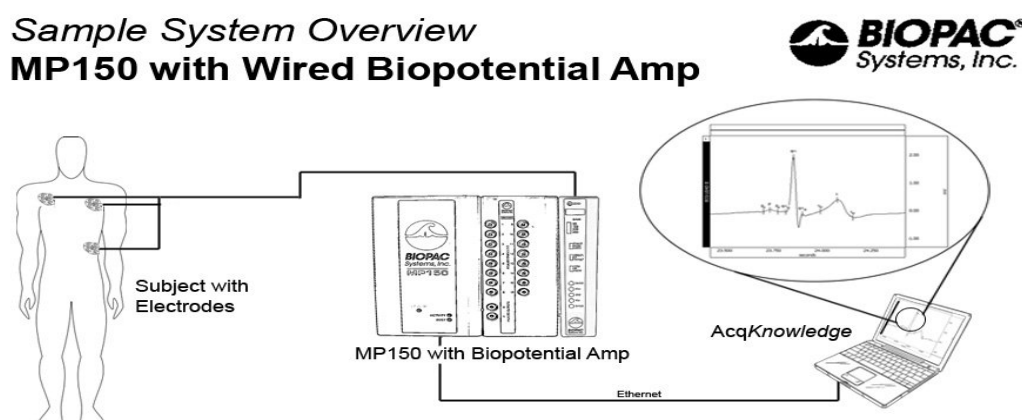


Figure 16. Le système Biopac MP150

¹³ <https://www.biopac.com>

Dans le cadre de ce mémoire, nous utilisons comme équipements de mesures le module de reconnaissance des expressions faciales combiné à l'oculomètre. La figure 17 illustre une représentation du schéma expérimental mis en place en explicitant les positions et les données collectées des capteurs placés sur le corps (à droite) et ceux intégrés dans l'ordinateur expérimental (à gauche).

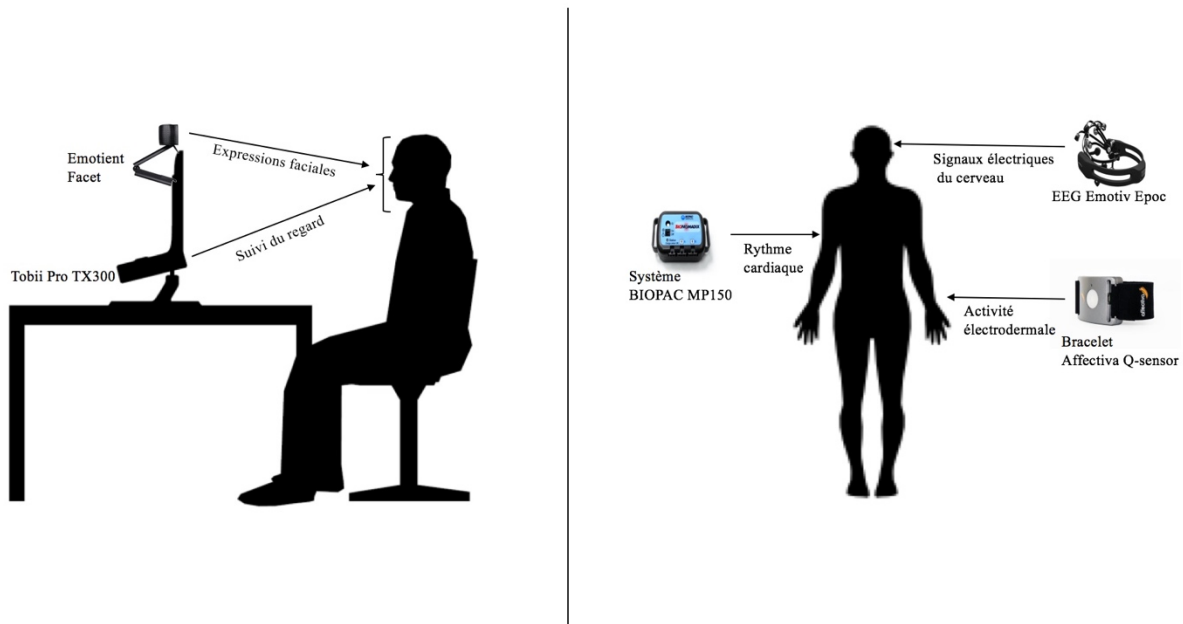


Figure 17. Dispositif expérimental

CHAPITRE 5 “ÉMOGRAPH” : la méthode et l’application

L’analyse des données recueillies a conduit au développement d’une méthode d’extraction des déclencheurs d’émotions fortes chez les participants. L’outil que nous proposons intègre cette méthode en permettant une analyse émotionnelle ainsi qu’une illustration graphique des transitions émotionnelles du joueur. Outre ces objectifs cette application permet de prédire les émotions d’un nouveau participant par rapport aux objets du jeu déjà identifiés.

Allant de la préparation des données à leur utilisation par l’application, ce chapitre décrit toutes les étapes de conception de la méthode avant de revenir sur l’application du nom d’Émograph.

5.1 La méthode

L’utilisation des données brutes extraites des senseurs nécessite une étape préalable à l’utilisation des données collectées. Les données traitées sont enregistrées sous forme de tables dans la base de données que va manipuler l’application. Le traitement et l’analyse des données ont été faits avec le langage Python notamment avec l’utilisation des bibliothèques NumPy¹⁴ et Pandas¹⁵. NumPy est la bibliothèque fondamentale pour les calculs scientifiques faits avec le langage Python, Pandas est gratuite (open source), fournit des performances élevées et une manipulation facile des structures de données et d’analyse pour le langage Python.

5.1.1 L’analyse du suivi du regard

Une fois l’expérience achevée, nous avons annoté de manière précise les séquences de jeu des participants. C’est à dire définir les AOI propres à leur parcours dans le jeu. Les AOI se réfèrent aux étapes du jeu décrites dans la section 4.1.1. Le fait que les joueurs sont tous différents, implique que nous avons une annotation par joueur. Nous avons utilisé le module

¹⁴ <http://www.numpy.org/>

¹⁵ <http://pandas.pydata.org/>

d'analyse du logiciel iMotions pour rejouer et annoter chronologiquement la vidéo de jeu des participants. La création des AOI permet au logiciel d'organiser et de catégoriser les données du regard en fonction de celles-ci. Le logiciel fait l'analyse et produit des métriques concernant les zones d'intérêts de chaque participant. Les métriques principales qui nous intéressent sont le "hit time", le "time spent". Ces métriques indiquent respectivement le moment auquel un joueur s'intéresse à un élément et le temps passé sur cet élément. La figure 18 illustre l'étape d'annotation de la vidéo de jeu.

Les données collectées ont été exportées sous forme d'un classeur Microsoft Excel nommé "Moving_AOI" contenant les données communes de tous les participants dont un extrait est présenté à la figure 19.

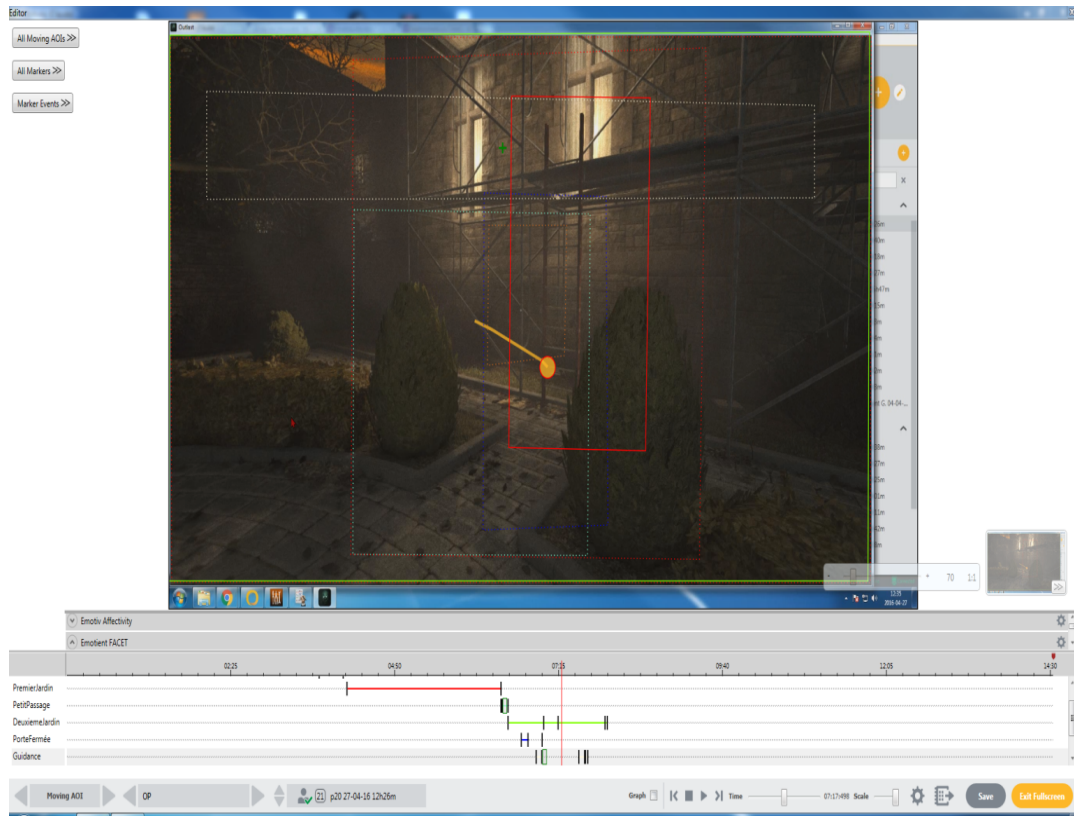


Figure 18. Annotation de la vidéo d'un participant

A	B	C	D	E	F	G	H	I
Study name	Stimulus Name	Respondent	Gender	Age	Group	AOI-Name	Hit time-G (r	Time spent-G (Tim
Outlast	OP	p20 27-04-16	Male	21	Default	Grille entree	223673	14633
Outlast	OP	p20 27-04-16	Male	21	Default	Premier jardin	248304	132592
Outlast	OP	p20 27-04-16	Male	21	Default	Petit passage	386537	2617
Outlast	OP	p20 27-04-16	Male	21	Default	Deuxieme jardin	390720	86305
Outlast	OP	p20 27-04-16	Male	21	Default	Porte fermee (deuxieme jardin)	402452	5848
Outlast	OP	p20 27-04-16	Male	21	Default	Bureau entree	204509	18836
Outlast	OP	p20 27-04-16	Male	21	Default	Cuisine	628308	9050
Outlast	OP	p20 27-04-16	Male	21	Default	Echelle	435515	4789
Outlast	OP	p20 27-04-16	Male	21	Default	Porte fermee (premier jardin)	296898	5847
Outlast	OP	p20 27-04-16	Male	21	Default	Camions	317063	6206
Outlast	OP	p20 27-04-16	Male	21	Default	Salon obscur	479093	36977
Outlast	OP	p20 27-04-16	Male	21	Default	Salon 2 (Tv)	521637	15875
Outlast	OP	p20 27-04-16	Male	21	Default	Bureau 1	554800	14353
Outlast	OP	p20 27-04-16	Male	21	Default	Sang couloir	547368	3802
Outlast	OP	p20 27-04-16	Male	21	Default	Bureau 2	587014	22961
Outlast	OP	p20 27-04-16	Male	21	Default	Sang conduite aeration	630074	4683
Outlast	OP	p20 27-04-16	Male	21	Default	Conduite aeration	634925	39755
Outlast	OP	p20 27-04-16	Male	21	Default	Monstre	660787	5831

Figure 19. Extrait du fichier "Moving_AOI"

5.1.2 L'analyse des expressions faciales

Le logiciel Emotient FACET génère des valeurs numériques pour chaque émotion de base. Ces valeurs sont présentées sur une échelle de -5 à 5 et représentent les probabilités de la présence d'une émotion. Plus une valeur générée est grande, plus il est probable qu'un expert humain la catégorise convenablement. Par exemple, une valeur de joie égale à 4, signifie un gros sourire et n'importe quel expert humain verrait la même chose. Le tableau 4 propose une explication de ce procédé en prenant la joie comme exemple.

Tableau 4. Interprétation des valeurs de FACET

Valeur	Émotion	Interprétation
-2	Joie	L'expression observée est 100 fois plus probable d'être catégorisée par un expert humain comme <i>n'étant pas</i> joyeuse.
-1	Joie	L'expression observée est 10 fois plus probable d'être catégorisée par un expert humain comme <i>n'étant pas</i> joyeuse.
0	Joie	L'expression observée est à chances égales d'être catégorisée par un expert humain comme <i>étant</i> joyeuse ou <i>n'étant pas</i> joyeuse.

1	Joie	L'expression observée est 10 fois plus probable d'être catégorisée par un expert humain comme <i>étant</i> joyeuse.
2	Joie	L'expression observée est 100 fois plus probable d'être catégorisée par un expert humain comme <i>étant</i> joyeuse.

Pendant l'enregistrement, le système de classification FACET tente de classer chaque trame de vidéo en temps réel. Cependant, selon les ressources disponibles pendant l'enregistrement, le système peut sauter des trames pour rester en phase avec les données entrantes. Nous avons procédé au re-calcul des trames ignorées avant de les exportées sous forme d'un classeur excel par participants. La figure 20 présente un extrait de ces données exportées.

StudyName	ExportDate	Name	Age	Gender	StimulusName	SlideType	EventSource	Timestamp	Joy Evidence	Anger Evidence	Surprise Evidence	Fear Evidence	Contempt Evidence	Disgust Evidence	Sadness Evidence
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	UI Emotient	47	0.8953308	-1.962669	-4.282578	-2.56378	-0.1849268	0.7475102	-2.438709	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	76								
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	196	0.8495329	-1.914344	-4.280642	-2.564664	-0.2859448	0.7703474	-2.462924	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	334	0.8002739	-1.926826	-4.185625	-2.508569	-0.06005862	0.7886484	-2.425766	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	336	1.157699	-1.866476	-4.498157	-2.482956	-0.2265256	0.597037	-2.176896	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	437	1.240291	-1.699465	-4.835064	-2.322008	-0.2236127	0.3647034	-1.62414	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	438	1.259502	-1.789097	-4.757075	-2.480948	-0.2544931	0.4712051	-1.810258	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	439	1.355551	-1.808659	-4.523743	-2.4846	-0.1375579	0.3423504	-1.956994	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	440	1.290067	-1.944465	-4.64291	-2.465719	-0.1913258	0.3649714	-2.038733	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	541	1.110662	-1.600445	-4.323603	-2.375071	-0.01169333	0.4076437	-1.764108	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	542	1.357183	-1.877982	-4.556815	-2.474208	-0.1294932	0.3581229	-1.879849	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	543	1.352213	-1.843435	-4.418516	-2.396691	-0.09495881	0.3584543	-1.832171	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	543	1.325399	-1.826708	-4.439493	-2.409326	-0.1105102	0.3547049	-1.84586	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	544	1.068539	-1.57943	-4.398458	-2.404926	0.0308438	0.4617773	-1.739216	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	545	1.205869	-1.697825	-4.235656	-2.280968	0.0001482886	0.3459498	-1.714126	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	545	1.101133	-1.673405	-4.150218	-2.186069	0.0407601	0.3659613	-1.717184	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	645	1.077265	-1.599882	-4.136751	-2.160985	0.06745407	0.3963466	-1.646581	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	647	1.251344	-1.807953	-4.022912	-1.95821	-0.0612567	0.2922418	-1.829288	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	648	1.131204	-1.464095	-3.501614	-1.529269	-0.01673711	0.2649359	-1.708815	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	669	1.102944	-1.475643	-3.289928	-1.367073	-0.0858583	0.2042049	-1.74138	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	700	0.9626099	-1.272491	-2.94457	-1.232967	0.09842514	0.2474115	-1.763078	
Outlast	20160927	p4 06-04-16.1	22	MALE	Emotient_BaselineTestImage	Emotient FA	732	1.030363	-1.304503	-2.725477	-1.065448	0.02613096	0.1234038	-1.814084	

Figure 20. Extrait du fichier exporté de FACET

La transformation des données brutes

Les données de FACET (figure 20) telles que présentées sont brutes, pour posséder des valeurs des émotions qui reflètent au plus près les expressions réelles d'un participant il faut procéder à une étape de correction des données. La réalisation de cette étape engage l'utilisation de valeurs de références. Elles sont évaluées pendant la période de calibration. Voici un exemple

pour mieux comprendre l'importance de cette étape. Supposons qu'un participant ait une expression de base plutôt joyeuse. Si ce paramètre n'est pas pris en compte et que les valeurs brutes sont utilisées, la quantité de sourires et d'autres expressions joyeuses dans une période choisie sera biaisée. Le résultat de cette étape permet de voir les changements d'expressions émotionnelles des participants par rapport à leur propre expression neutre. La correction des valeurs consiste à 2 étapes pour chaque émotion. La première est de calculer la médiane de chaque émotion sur toute la période de calibration. La seconde est de la soustraire aux valeurs des émotions brutes de la période d'expérience (iMotions, 2016). Dans notre étude, la période de calibration a consisté à l'utilisation de stimuli neutre, un écran gris pendant une période de 6 secondes. Les participants étaient assis dans une position confortable et regardaient fixement la caméra avec l'expression neutre qui leur ai propre. La figure 21 ci-dessous, illustre nos propos. Elle démontre l'impact de la correction des données sur l'étape 'Échelle' par rapport au degré de colère reconnu. Concernant cette image, la valeur de référence de colère était équivalente à 0.6920737.

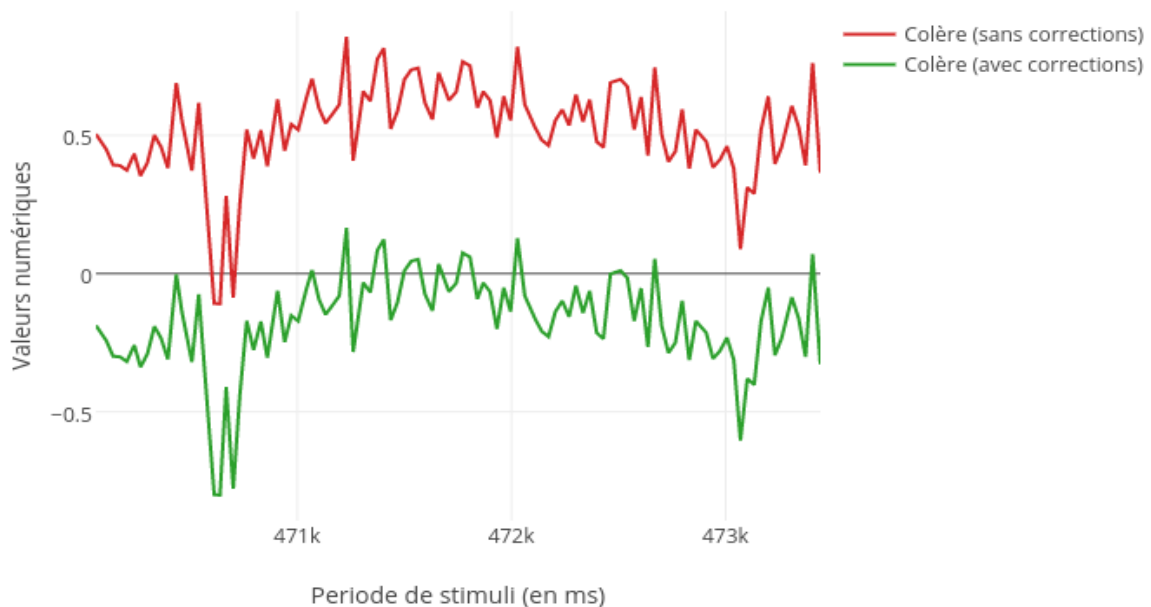


Figure 21. Effet de la correction par la valeur de référence

Un autre aspect important dont il faut tenir compte est la sélection des valeurs numériques. Nous avons procédé au seuillage. Il s'agit d'une technique qui consiste à définir un critère afin d'écartier certaines données pour mettre en avant l'information pertinente. Notre intérêt se porte vers les émotions les plus fortes/évidentes alors nous avons utilisé un seuillage basé sur l'amplitude. Cela signifie qu'une amplitude minimale est spécifiée pendant laquelle une expression faciale doit être active pour être considérée dans l'analyse. Cette technique permet de dissimuler les occurrences de faible amplitude telles qu'un léger sourire ou un froissement des sourcils subtil. Toutefois, l'idée d'utiliser cette technique soulève une nouvelle problématique. Est-ce qu'une valeur limite, fixe et commune serait un moyen efficace? En partant du même principe que toutes les personnes sont différentes, il ne serait pas commode de fixer un seuil unique pour tous. En effet, nous avons utilisés des seuils relatifs à chacun des participants. Ces seuils individuels sont calculés à partir de la distribution des données des participants. Nous avons défini le critère d'isolation au 3eme quartile. C'est cette valeur qui sépare les 75% des données les plus petites des 25% plus grandes. Avec cette technique, nous pouvons identifier individuellement les périodes pendant lesquelles les participants ont montré leurs expressions faciales les plus véhémentes. Dans la suite de ce mémoire nous verrons une description de la méthode mise en place pour manipuler les données traitées. La figure 22, montre pour la même étape dans le jeu que la figure 21, la position du seuil de colère relatif déterminé sur l'ensemble des données de jeu du participant.



Figure 22. Seuil de colère relatif

5.1.3 La recherche de l'émotion dominante

Cette étape consiste à faire la correspondance entre les données exportées. C'est-à-dire, pour chaque participant, trouver l'effet émotionnel des éléments regardés dans le jeu. Le classeur Excel extrait de l'étape d'annotation, contient les informations communes des participants. Il faut dans un premier temps séparer les données de manière individuelle. Dans un second temps nous avons appliqué la méthode suivante pour chaque participant.

- Extraire les AOI indépendamment.
- Pour chaque AOI, la segmenter en intervalles d'une fenêtre de temps de 500 millisecondes (recommandation du service de support de iMotions).
- Calculer la médiane de chacune des émotions pendant chaque fenêtre de temps.
- Si la médiane dépasse le seuil relatif, alors attribuée la valeur 1 (mise pour représenter le 'vrai') sinon 0. À la place des valeurs numériques, ces valeurs binaires seront utilisées.
- Pour chaque AOI, faire la somme des scores binaires reportés.
- Sélectionner l'émotion dont le score est le plus élevé.
- La reconnaissance des émotions par les expressions faciales est basée sur des emplacements précis du visage, alors il est très probable que plusieurs émotions utilisent les mêmes emplacements. Par exemple froncer les sourcils est un signe à la fois de colère et de confusion. De ce fait, il n'est pas anormal de rencontrer plusieurs '1' dans une même fenêtre. Pour pallier à cette situation, nous attribuons le score '1' à l'émotion dont la médiane est la plus élevée sur cette fenêtre de temps.
- Toutefois, cette sélection à elle seule ne nous épargne pas le fait d'avoir des comptes égaux d'émotions pour une AOI. Dans cette situation, nous choisissons comme émotion dominante, celle dont la somme des médianes calculées pendant la segmentation sera la plus élevée.

À la suite de cette étape, nous obtenons pour chaque participant un classeur Excel qui contient, la liste des AOI qu'il a regardé, l'émotion dominante correspondante pour chaque AOI,

la valeur numérique du point le plus élevé de cette émotion ainsi que l'intervalle de temps pendant lequel le joueur l'a exprimé. Le fichier obtenu, nous permet d'obtenir aisément une représentation graphique du parcours émotionnel d'un joueur. Sur la figure 23, ci-dessous les points médians sont déterminés sur chaque fenêtre de temps et nous observons que pour cette étape nommée "Échelle" le score binaire de colère est 0. Nous présenterons un exemple des classeurs obtenus dans le prochain chapitre, celui des résultats.

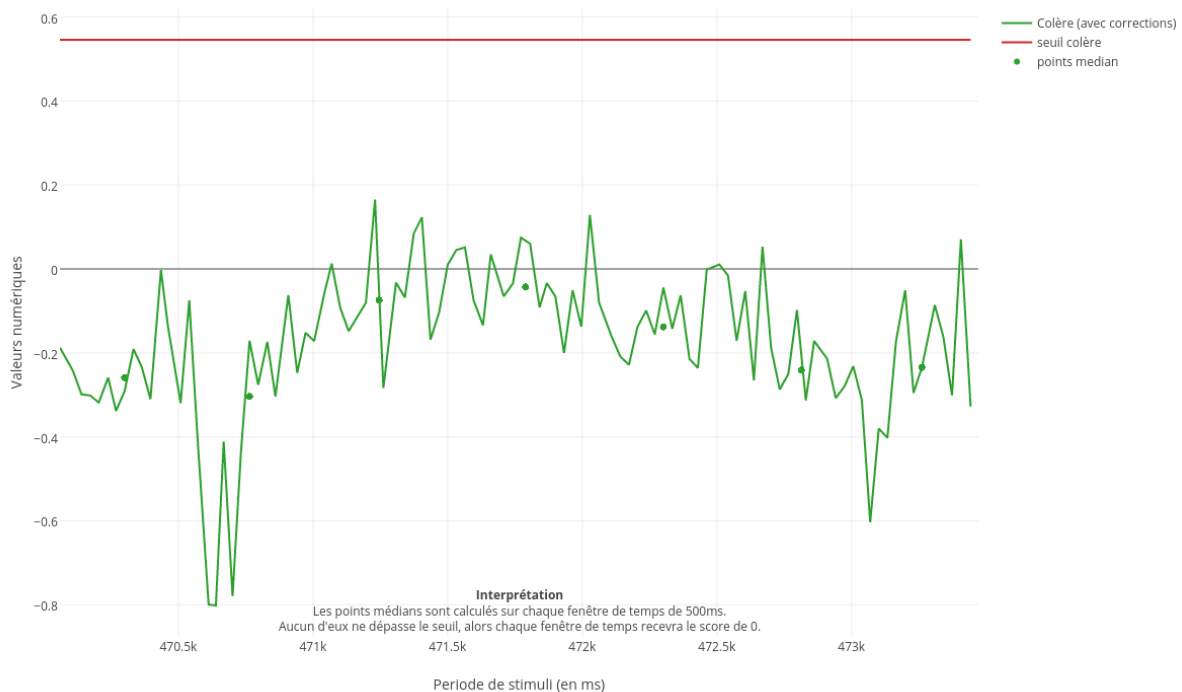


Figure 23. Exemple de recherche d'émotions fortes (Colère)

La construction des fichiers des zones d'intérêt

Rappelons que même si le parcours du joueur est fixe, il existe cependant des zones qui n'ont pas été regardées par tous les participants, d'où la nécessité de les identifier. Cette phase consiste à la construction d'un fichier unique par AOI. Ce fichier contient les informations sociodémographiques, les traits de personnalité, la catégorie des participants qui l'ont regardée et celles relatives aux émotions dominantes des participants extraites à l'étape précédente. Un exemple de ce classeur sera présenté également dans le chapitre suivant.

La construction d'un fichier global

Dans cette étape, il s'agit de constituer un fichier général des AOI, tel que sa caractéristique d'identification soit le couple (AOI, participant) par exemple, (Bibliothèque, P01), (Bibliothèque, P02) etc. Ce fichier est un regroupement des 2 fichiers précédemment extraits en 1 seul. Il contient alors toutes les informations des participants et des AOI.

La base de données

Celle-ci contient tous les fichiers créés pendant le traitement des données sous forme de tables. Nous utilisons une base de données MySQL¹⁶ pour stocker les informations. Elle joue le rôle d'intermédiaire entre le traitement des données et l'application.

5.2 L'application web dynamique : Émograph

L'application web dynamique nommée Émograph a été développée dans le but de fournir aux utilisateurs des services interactifs ainsi qu'un accès facile à diverses ressources d'informations quant aux émotions. Précisons que les utilisateurs concernés par notre application sont principalement les concepteurs de jeux vidéo, de publicités ou de pièces de théâtres, etc.

Pour le développement de l'application nous avons utilisé la plate-forme de développement web WampServer¹⁷. Il s'agit d'une plateforme qui fonctionne sous Windows et est destinée à la création des applications web dynamiques à l'aide du serveur Apache2, du langage de scripts généraliste et open source PHP5 (PHP Hypertext Preprocessor) et d'une base de données MySQL. Le module de gestion de base de données PHPMysqlAdmin y est également inclus. Les requêtes de manipulation des données sont écrites avec le langage SQL (Structured Query Language, en français langage de requête structurée).

La présence et l'utilisation de la base de données permet de faire fonctionner l'application indépendamment du stimuli d'analyse. Dans le sens où les intérêts futurs peuvent être dirigés vers l'analyse de stimuli comme des pièces de théâtres, de publicités ou d'autres

¹⁶ <http://www.mysql.com>

¹⁷ <http://www.wampserver.com>

jeux vidéo, il suffit que la base de données soit constituée de tables créées par la méthode proposée dans la section 5.1. Émograph, l'application, est composée de 2 modules. Le module d'analyse émotionnelle et le module de visualisation. La figure 24 présente l'architecture globale de base d'Émograph. Elle présente le processus d'utilisation de la méthode et de l'application.

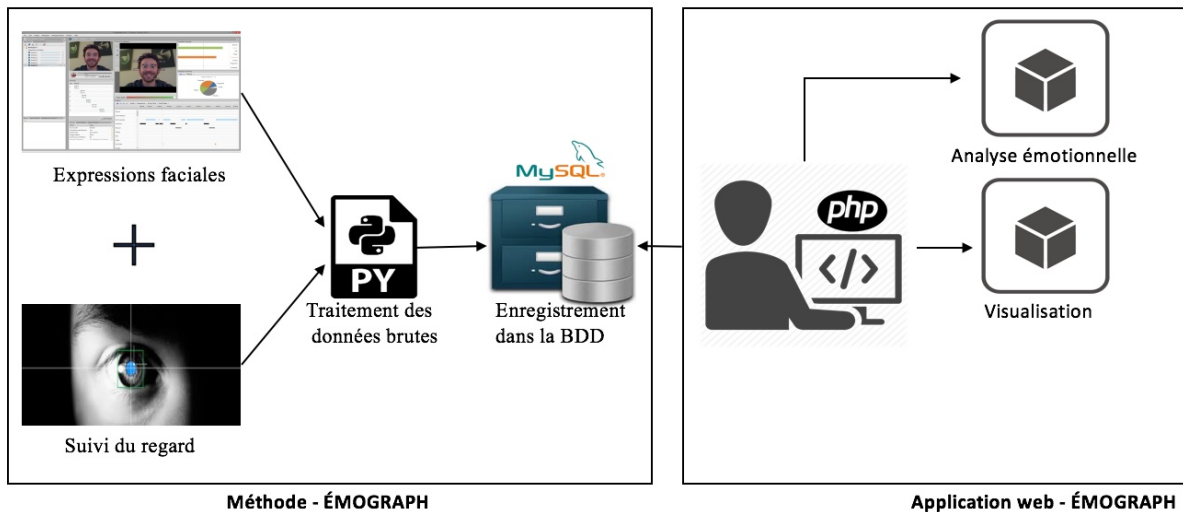


Figure 24. Architecture globale de base d'Émograph

Le module d'analyse émotionnelle

Ce module utilise les tables relatives aux participants contenant leurs informations collectées par le procédé décrit dans la chapitre 4 et aux AOI. En plus d'offrir aux utilisateurs la possibilité de connaître les émotions dominantes des personnes, il leur permet d'approfondir leur recherche en s'intéressant aux autres émotions en dehors de celle dominante. Il propose une recherche par participants ou par AOI.

En ce qui concerne les participants, l'utilisateur choisi le participant ainsi que l'émotion dont il souhaite voir les informations et le module lui affiche les informations désirées dans une fenêtre. Ces informations sont telles que, le nom de l'AOI, sa description, ses actions/challenges et le temps précis d'expression de l'émotion choisie.

Par rapport à la recherche par AOIs l'utilisateur choisi le nom de l'AOI voulue et l'application lui affiche les émotions dominantes reconnues, le pic maximal enregistré ainsi que

l'émotion correspondante et le temps d'apparition de celui-ci. Les captures d'écran illustrant le fonctionnement de ce module seront présentées dans la section 6.2.1 du chapitre suivant.

Le module de visualisation

Ce module procède à l'affichage sous forme de graphe des effets émotionnels des scènes du jeu sur les participants. Ce procédé a été vu dans la section 3.4. Le module propose de choisir le participant dont on souhaite voir le graphe. De plus il permet de visualiser l'image précise de jeu responsable de l'émotion dominante. Un exemple du module sera présenté dans la section 6.2.2 du chapitre suivant.

CHAPITRE 6 Les résultats

Ce chapitre fait état des résultats provenant de l'application de notre méthode d'évaluation émotionnelle de la séquence de jeu d'Outlast. Nous retrouvons dans ce chapitre l'analyse émotionnelle faite par la méthode sur les données traitées, sa représentation graphique et des statistiques de jeu. De plus nous continuerons en montrant les captures d'écran de l'application créée, Émograph.

6.1 La méthode

L'analyse émotionnelle

La mise en application de la méthode décrite dans la section 5.1 a donné des résultats que nous présentons dans cette section. La figure 25 montre le classeur du participant P21 obtenu. Un classeur identique est produit pour chaque participant. Sur cet exemple il s'agit d'une femme, de niveau amateur dans les jeux vidéo.

Area Of Interest	Émotion dominante	Valeur du point médian	Probabilité	Intervalle du point médian	Pic de l'émotion dominante	Apparition du pic
Grille entree	anger	2,49098444	0,996781781	0:02:59.248000	2,7644091	0:02:51.777000
Premier jardin	anger	2,59793206	0,997482478	0:04:58.985000	2,7852449	0:03:35.985000
Porte fermee (premier jardin)	anger	1,8406873	0,985773766	0:04:03.376000	2,009673	0:04:00.384000
Petit passage	surprise	1,8360396	0,985622904	0:06:33.641000	2,0368406	0:06:27.649000
Deuxieme jardin	surprise	2,2435746	0,994325157	0:07:49.076000	2,8181016	0:06:40.080000
Echelle	joy	1,11826675	0,929228898	0:07:12.580000	1,94129675	0:07:12.080000
Salon obscur	surprise	1,5929086	0,975103287	0:08:27.424000	2,0025576	0:08:19.924000
Salon 2 (Tv)	surprise	2,1482006	0,992941328	0:09:07.221000	2,3136126	0:08:52.721000
Sang couloir	surprise	2,0635346	0,991434955	0:09:45.513000	2,2020056	0:09:41.046000
Bureau 1	anger	2,3163624	0,995196625	0:09:48.241000	2,3163624	0:09:45.248000
Cuisine	fear	2,373933	0,995790457	0:10:48.981000	2,375858	0:10:30.481000
Sang conduite aeration	joy	2,29445975	0,99494942	0:10:36.313000	2,58997975	0:10:33.313000
Conduite aeration	surprise	2,2290766	0,994133647	0:10:57.819000	2,4041826	0:10:56.819000
Bibliotheque	anger	2,47712497	0,996677769	0:12:07.013000	2,8704365	0:11:44.028000
Cadavre suspendu	surprise	1,9114076	0,987885679	0:11:43.285000	2,5582846	0:11:42.785000
Cadavre qui parle	surprise	2,7199446	0,998097921	0:13:27.317000	3,2812686	0:13:19.840000
Monstre	joy	0,23861945	0,634006026	0:14:51.912000	0,23861945	0:14:50.912000
Personnage alternatif	sadness	1,38494255	0,960416249	0:15:18.152000	1,8665364	0:15:13.153000
Sang fin sequence	sadness	1,9187056	0,988085145	0:15:40.030000	2,1609914	0:15:39.041000

Figure 25. Émotions dominantes du participant P21

Sur la figure ci-dessus, nous voyons qu'outre la détermination de l'émotion dominante, le moment d'apparition de l'éminence de celle-ci est affiché dans la colonne "Apparition du pic". Sur la ligne de couleur orange à la figure 25, nous pouvons voir que l'émotion dominante déterminée pour l'étape "Cadavre qui parle" est la surprise et le temps d'apparition du pic qui

est à 00:13:19:8400 de la vidéo de jeu (13 minutes 19 secondes 8400 millisecondes). La figure 26 est une capture d'écran de jeu du participant P21 à l'étape "Cadavre qui parle" (la ligne de couleur orange sur la figure 25). Sur cette capture nous constatons l'expression de surprise sur le visage de la participante, provoquée par l'image de l'agent de sécurité empalé. En bas, à droite, dans le cercle rouge se trouve le temps de la vidéo, la figure 27 est un agrandissement du cercle qui permet de constater que ce temps est équivalent à celui inscrit dans le classeur présenté à la figure 25. Nous voyons aussi la trace de son regard et le point de fixation de celui-ci sur l'agent à ce moment précis.

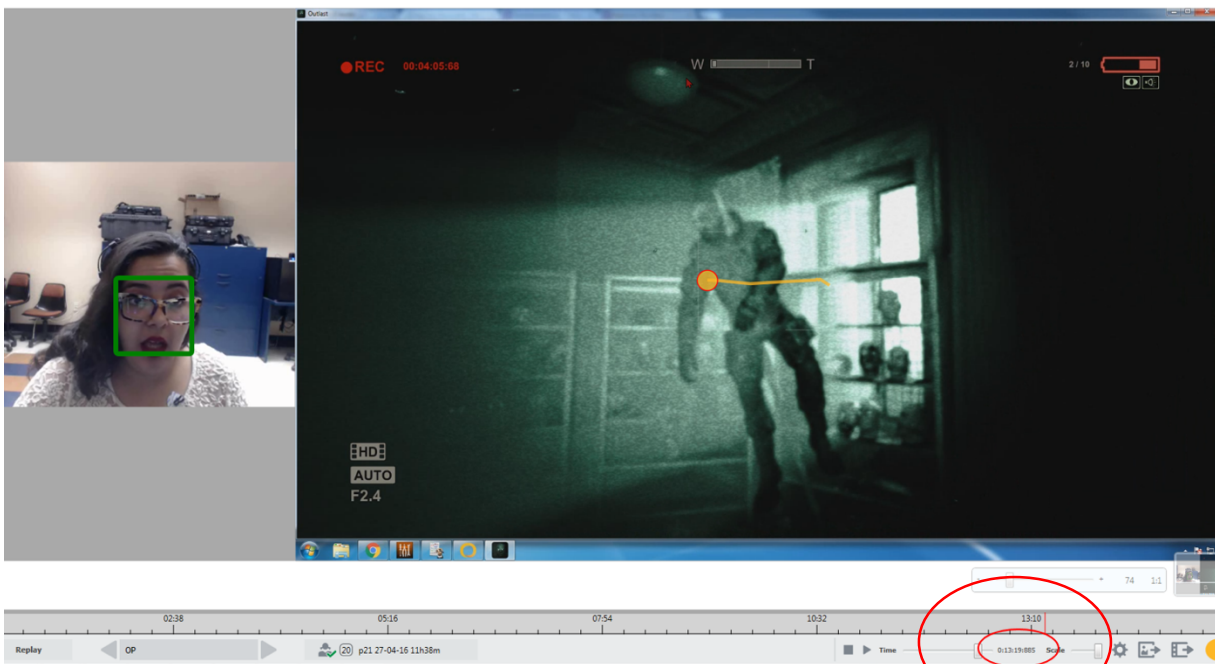


Figure 26. Capture de jeu du participant P21

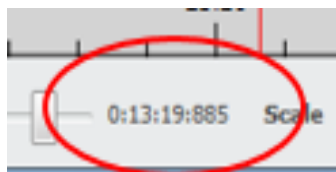


Figure 27. Temps de jeu du participant P21

Les stimuli n'engageant pas les mêmes réactions émotionnelles, nous allons montrer dans l'exemple suivant la réaction émotionnelle d'un autre participant soumis au même stimulus

(la ligne de couleur orange). Il s'agit d'un homme, de niveau intermédiaire en jeux vidéo. La figure 28 présente le classeur extrait qui lui est relatif. Comme dans l'exemple précédent, nous portons notre intérêt sur la ligne de couleur orange. L'émotion dominante déterminée est la colère et le temps d'apparition du pic de colère est à 00:15:58:204000 (15 minutes 58 secondes 204000 millisecondes) de la vidéo de jeu.

Area Of Interest	Émotion dominante	Valeur du point médian	Probabilité du point médian	Intervalle du point médian	Pic de l'émotion dominante	Apparition du pic
Bureau entree	anger	2,7304364	0,998143237	0:03:39.112000	2,8567394	0:03:32.117000
Grille entree	anger	2,6027364	0,997510106	0:04:04.904000	2,7777734	0:04:03.905000
Premier jardin	anger	3,0797484	0,999168446	0:04:23.536000	3,1334334	0:04:21.564000
Camions	anger	2,9957514	0,998991188	0:04:41.733000	3,1086484	0:04:40.733000
Porte fermee (premier jardin)	anger	2,7428294	0,998195378	0:04:53.139000	2,8596104	0:04:51.139000
Petit passage	surprise	0,922867	0,893307481	0:05:24.133000	1,508547	0:05:23.644000
Deuxieme jardin	anger	3,4914834	0,999677614	0:06:47.136000	3,4914834	0:05:33.645000
Echelle	anger	3,0985414	0,999203634	0:06:34.008000	3,2010064	0:06:14.013000
Salon obscur	anger	2,9296604	0,998825564	0:09:53.244000	3,1340474	0:09:50.268000
Salon 2 (Tv)	anger	2,8133924	0,998465593	0:10:40.364000	3,0518434	0:10:35.374000
Cuisine	disgust	1,071381	0,921791875	0:12:27.336000	1,4013874	0:12:23.836000
Sang conduite aeration	joy	0,888944	0,885630887	0:12:29.108000	0,93124	0:12:28.108000
Conduite aeration	surprise	0,263201	0,647037561	0:14:11	0,473838	0:13:05.004000
Bibliotheque	surprise	0,512356	0,764901643	0:15:48.877000	0,655848	0:15:05.900000
Cadavre suspendu	fear	0,940922	0,89720549	0:15:07.877000	0,940922	0:15:03.877000
Cadavre qui parle	anger	2,8112974	0,998458184	0:16:14.704000	3,1093924	0:15:58.204000
Porte fermee (fin)	anger	2,9088534	0,998767999	0:16:45.545000	3,0000244	0:16:43.069000
Toilettes	surprise	1,175917	0,937476485	0:17:11.177000	1,24212	0:17:07.677000
Monstre	surprise	0,01424	0,508196469	0:17:31.974000	0,088907	0:17:30.974000
Personnage alternatif	joy	1,6890541	0,979948397	0:17:51.042000	1,6890541	0:17:45.042000
Sang fin sequence	joy	1,318079	0,954130019	0:18:19.708000	1,7934464	0:18:17.708000

Figure 28. Émotions dominantes du participant P08.

La figure 29 montre la capture d'écran de jeu correspondant à cette ligne. Sur la figure nous voyons la colère exprimée sur le visage du participant et la trace de son regard sur l'agent de sécurité. De manière similaire, dans le cercle rouge en bas à droite se trouve le temps de la vidéo de jeu qui correspond effectivement au temps déterminé d'apparition du pic. La figure 30 est un agrandissement du cercle rouge afin de mieux lire le temps affiché.

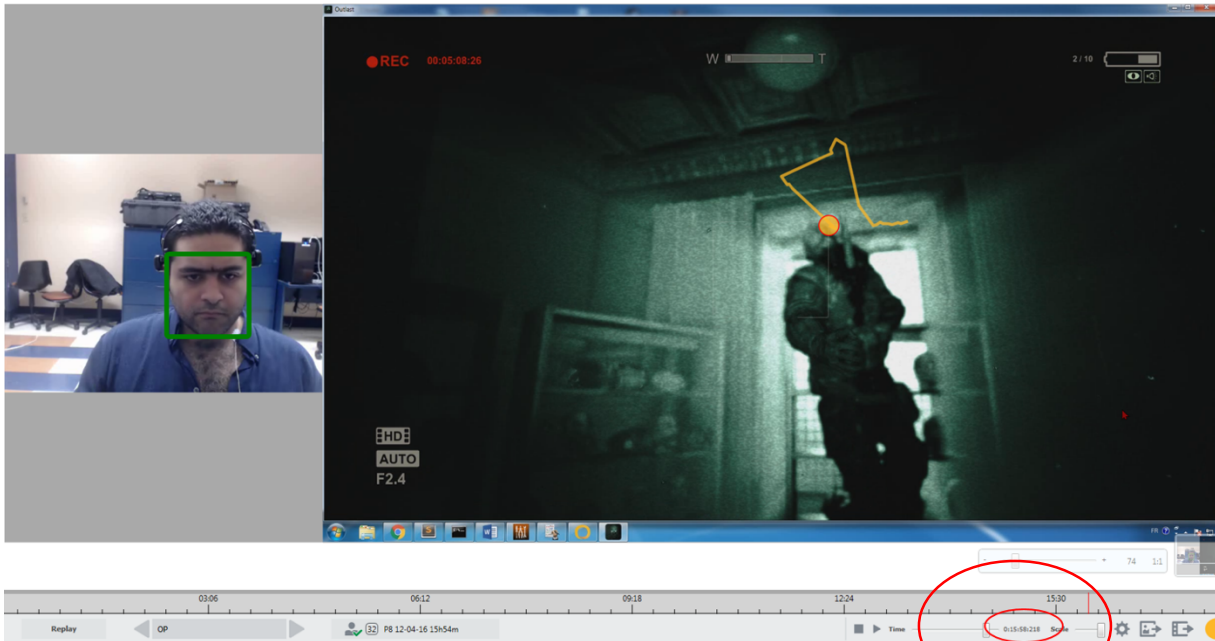


Figure 29. Capture du jeu du participant P08

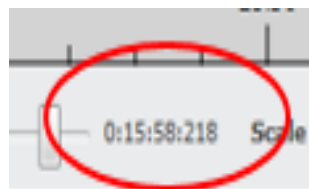





Figure 30. Temps de jeu du participant P08

Dans le tableau 5 ci-après, nous allons présenter des captures d'écrans de jeu correspondantes aux émotions exprimées par divers participants au cours des scènes du jeu. Ces captures ont été faites par l'utilisation la méthode proposée, telle que présentée avec les exemples précédents.

Tableau 5. Émotions de base reconnues chez divers participants

Émotion	Scène	Capture
Joie	Bibliothèque	 <p>The image shows a participant's face on the left, smiling broadly with a green bounding box around it. On the right is a game capture of a library scene with a green tint. A yellow arrow points to a book on a shelf. The capture includes a 'REC' indicator, a 'W T' scale, and camera settings: 'HD', 'AUTO', 'F2.4'.</p>
Colère	Salon obscur	 <p>The image shows a participant's face on the left, with a serious expression and a green bounding box around it. On the right is a game capture of a dark living room scene. A yellow arrow points to a chair. The capture includes a 'REC' indicator, a 'W T' scale, and camera settings: 'HD', 'AUTO', 'F2.4'.</p>
Surprise	Cadavre suspendu	 <p>The image shows a participant's face on the left, with a surprised expression and a green bounding box around it. On the right is a game capture of a library scene with a green tint. A yellow arrow points to a body hanging from a shelf. The capture includes a 'REC' indicator, a 'W T' scale, and camera settings: 'HD', 'AUTO', 'F2.4'.</p>

Peur	Deuxième jardin	
Dégout	Cadavre qui parle	
Tristesse	Petit passage	

Le graphe de transition émotionnelle

Ce graphe met en avant les émotions dominantes relatives aux scènes de jeu. Les valeurs numériques, qui lorsqu'elles sont présentées de manière graphique sont transformées en probabilités (entre 0 et 1), ce qui facilite leur interprétation et leur présentation. Cette transformation est faite selon la formule suivante, proposée par (iMotions, 2013) :

$$PP = \frac{1}{1 + 10e^{-LLR}}$$

Avec LLR étant la valeur numérique de l'émotion considérée.

La figure 31 présente le graphe concernant le participant P21 dont nous avons montré les données sur la figure 25. Les scènes de la séquence de jeu sont sur l'axe des abscisses tandis que les probabilités sont sur l'axe des ordonnées. Chaque point donne l'émotion dominante de la scène.

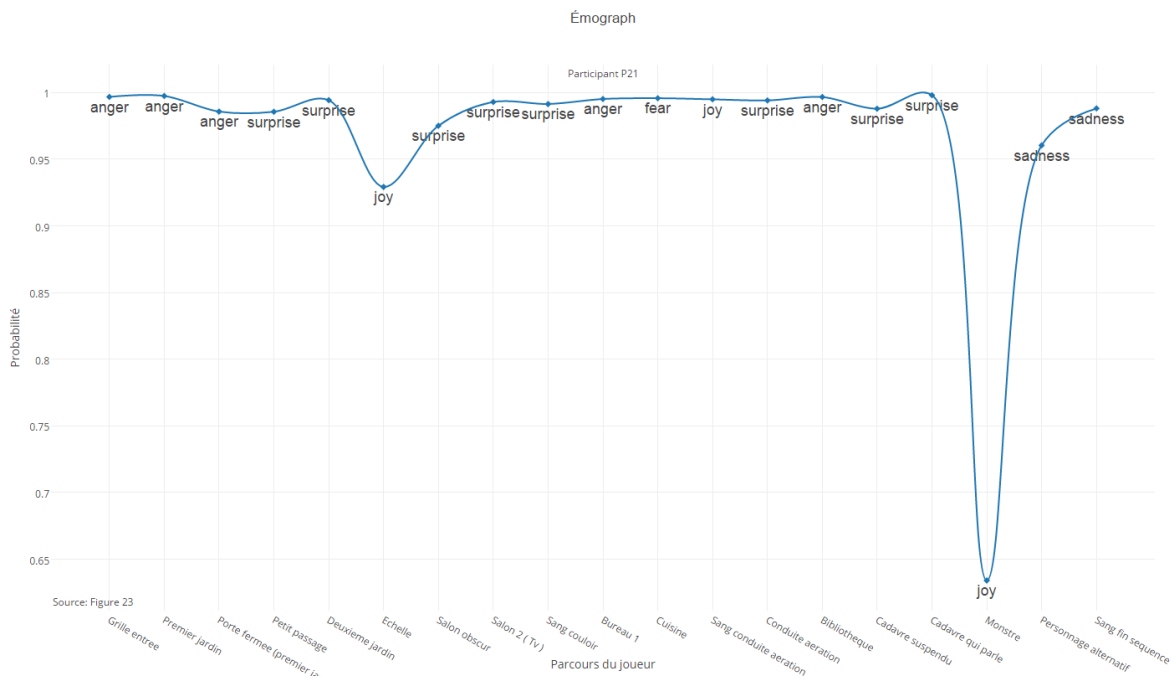


Figure 31. Graphe de transition émotionnelle du participant P21.

Statistiques

Sur toutes les scènes de la séquence nous avons déterminé 335 segments de données sur l'ensemble des participants. Nous remarquons que la colère est l'émotion la plus présente suivie de la surprise, 120 et 77 occurrences respectivement. La figure 32 présente un diagramme circulaire qui illustre les proportions des occurrences d'émotions dominantes.

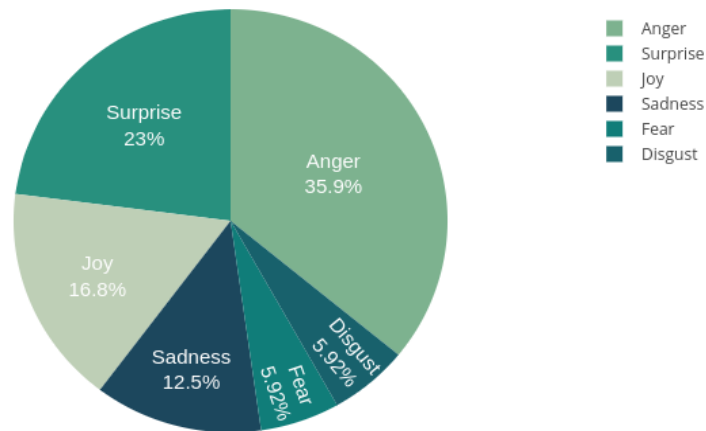


Figure 32. Statistiques globales de la séquence de jeu

L'ensemble des données du genre féminin constitue 29.7% de l'ensemble de données collectées. Chez les femmes nous avons constaté que la surprise est majoritaire suivie par la tristesse. Chez les hommes dont les données représentent 70,3% de la proportion globale, nous observons que la colère en majoritaire suivie par la surprise et la joie, les deux dernières étant très proches.

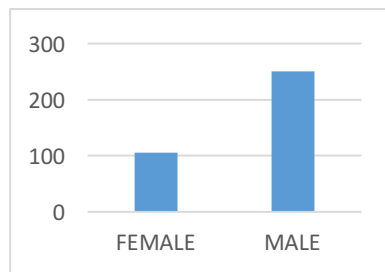


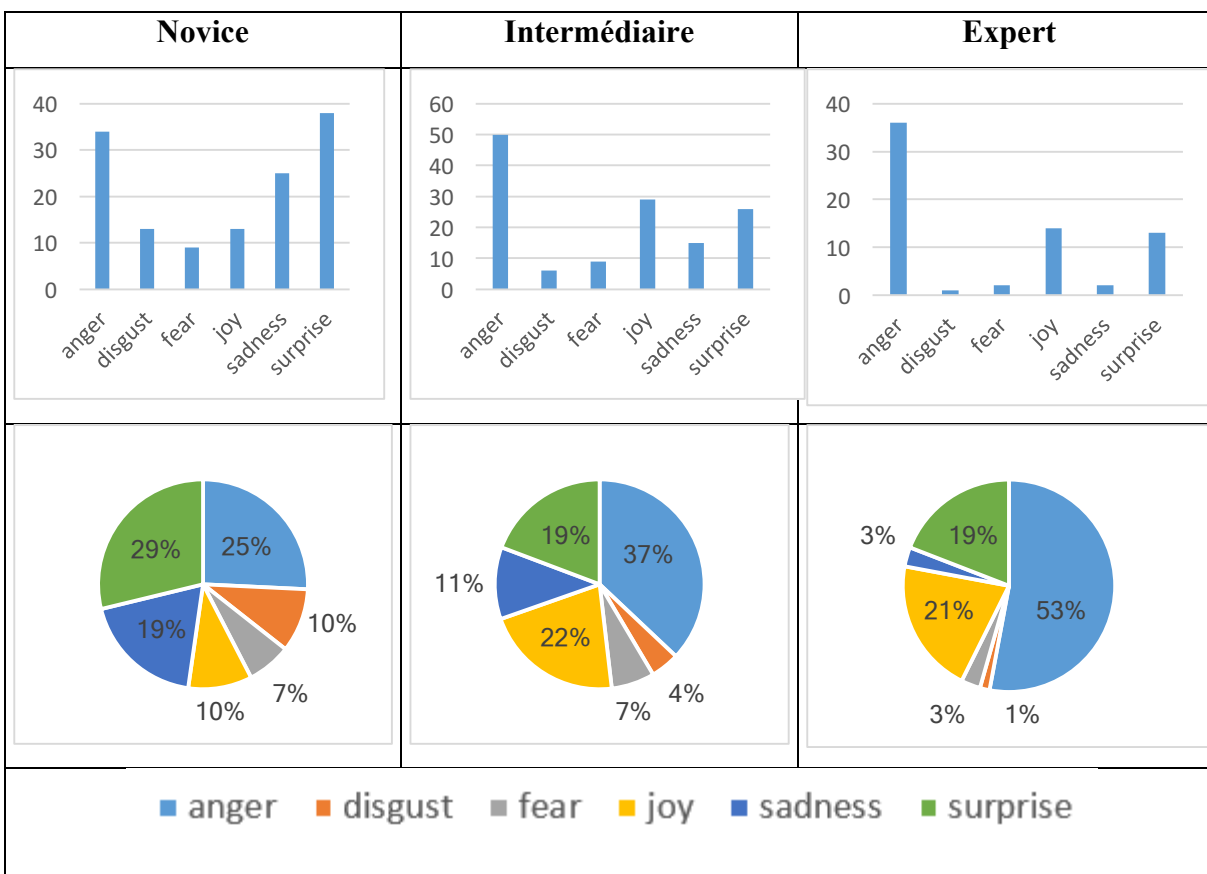
Figure 33. Répartition totale des segments selon le genre.

Par catégorie des joueurs

Nous distinguons des différences par rapport aux émotions détectées parmi les trois catégories de joueur. Ces différences s'expriment principalement chez les novices. Concernant les intermédiaires et experts, les différences ne sont pas remarquables. Pour tous les segments des novices, nous avons déterminé 29% de surprise, 25% de colère, 19% de tristesse, 10% de joie et de dégoût et 7% de peur. En ce qui concerne les joueurs de niveau intermédiaire, nous avons identifié 37% de colère, 22% de joie, 19% de surprise, 11% de tristesse, 7% de peur et 4% de dégoût. Par rapport aux joueurs experts, nous avons marqués 53% de colère, 21% de joie,

19% de surprise, 3% de tristesse et de peur et 1% de dégoût. Le tableau 6 ci-dessous, présente la comparaison entre les catégories de joueur.

Tableau 6. Occurrences d'émotions dominantes par catégories de joueurs.



Dégout/Disgust

Le dégoût a été exprimé le plus grand nombre de fois par les femmes, principalement dans la scène nommée ‘‘Personnage alternatif’’. La figure 34 montre la proportion de dégoût en fonction des zones d’intérêt.

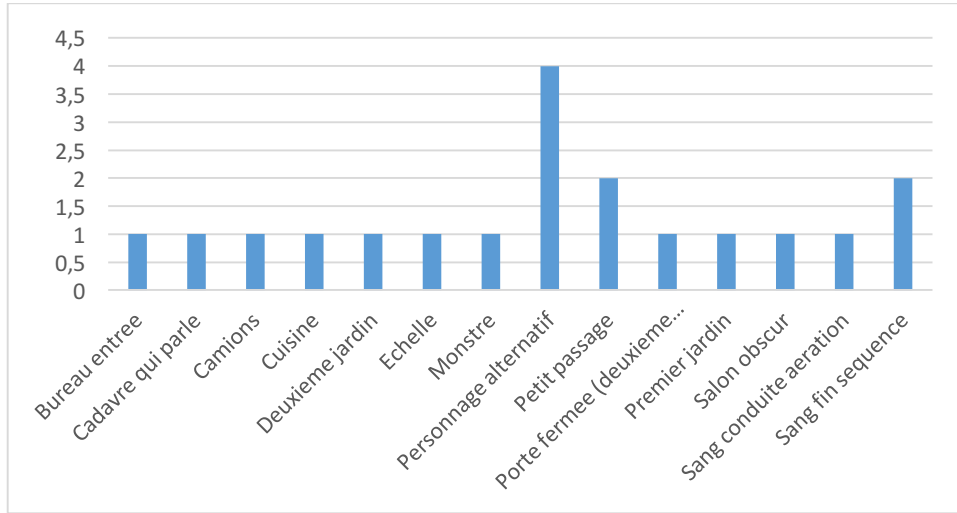


Figure 34. Proportion de dégoût par zones d'intérêt.

Peur/Fear

Les scènes, ‘Bibliothèque’, ‘Camions’, ‘Cuisine’, ‘Sang conduite aération’ sont celles pendant lesquelles la peur a été exprimée par le plus de participants. La figure 35 montre la proportion de peur en fonction des zones d'intérêt.

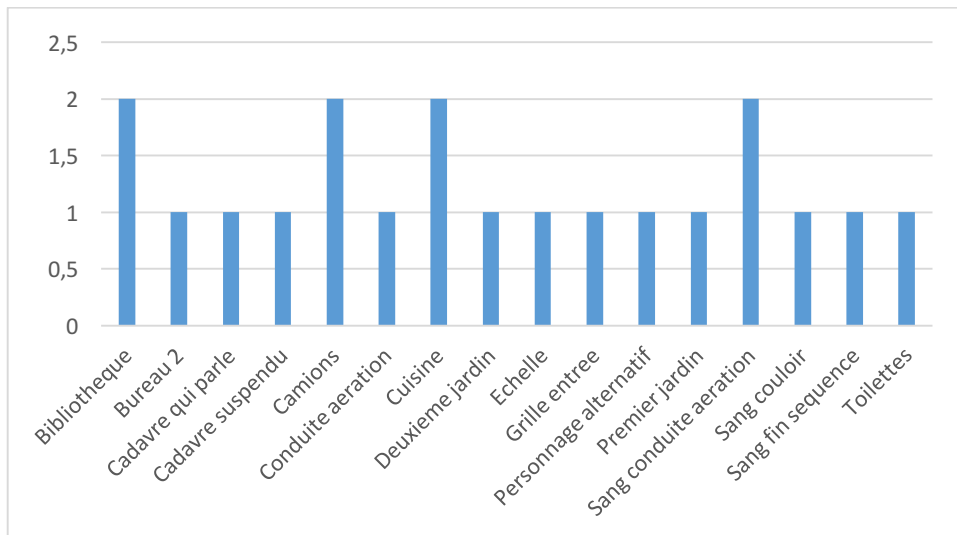


Figure 35. Proportion de peur par zones d'intérêt.

Tristesse/Sadness

Les scènes, ‘Cuisine’, ‘Grille entrée’ et ‘Sang fin séquence’, sont celles pendant lesquelles la tristesse a été exprimée par le plus participants. La figure 36 montre la proportion de tristesse en fonction des zones d'intérêt.

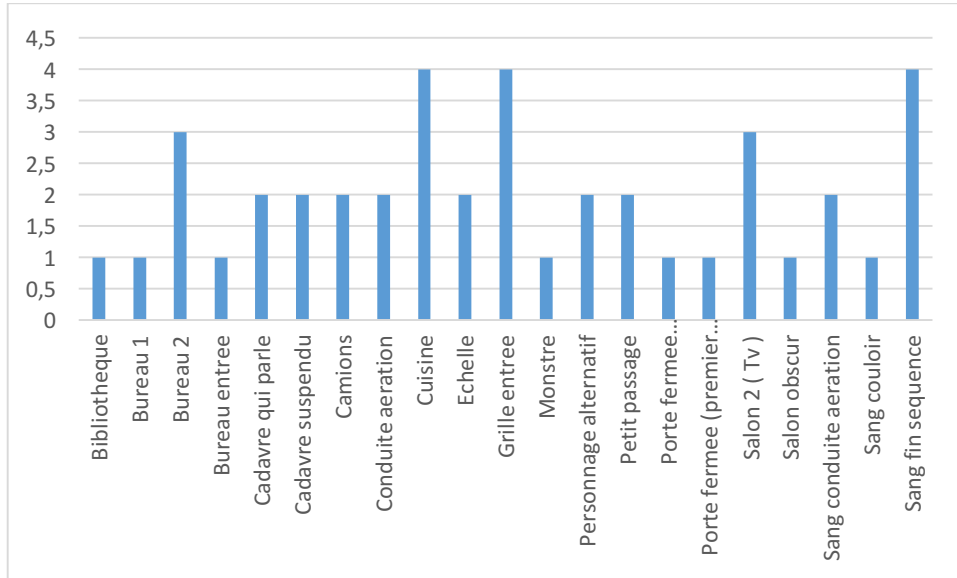


Figure 36. Proportion de tristesse par zones d'intérêt.

Joie/Joy

Les scènes, ‘‘Bibliothèque’’ et ‘‘Personnage alternatif’’, sont celles pendant lesquelles la joie a été exprimée par le plus de participants. La figure 37 montre la proportion de joie en fonction des zones d’intérêt.

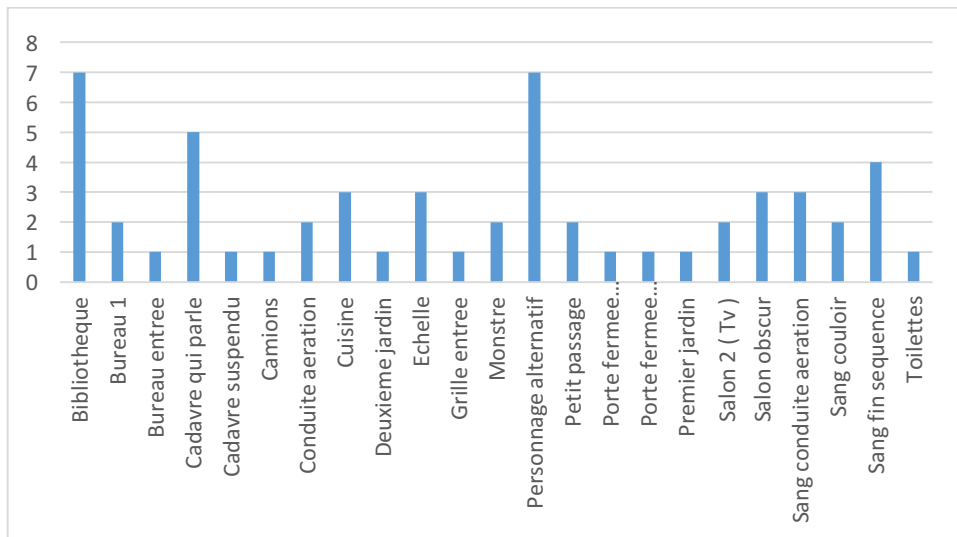


Figure 37. Proportion de joie par zones d'intérêt.

Surprise/Surprise

La scène ‘‘Deuxième jardin’’, est celle pendant laquelle la surprise a été exprimée par le plus de participants. La figure 38 montre la proportion de surprise en fonction des zones d’intérêt.

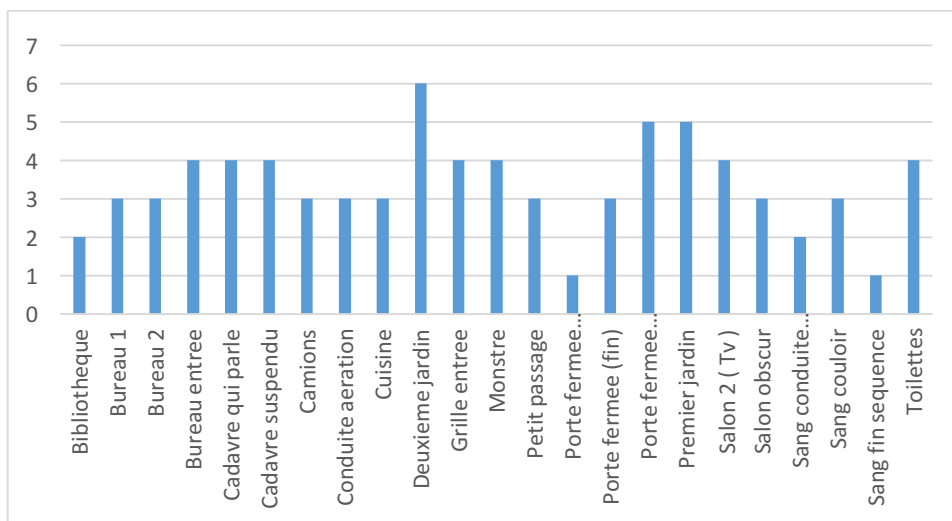


Figure 38. Proportion de Surprise par zones d'intérêt.

Colère/Anger

Les étapes ‘‘Échelle’’ et ‘‘Premier jardin’’ sont celles pendant lesquelles la colère a été exprimée par le plus de participants. La figure 39 montre la proportion de dégoût en fonction des zones d’intérêt.

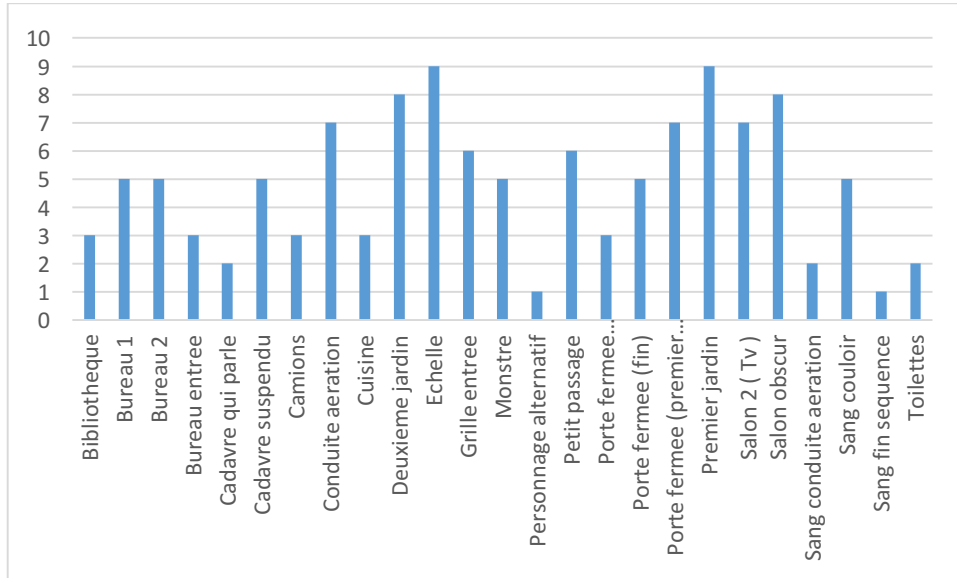
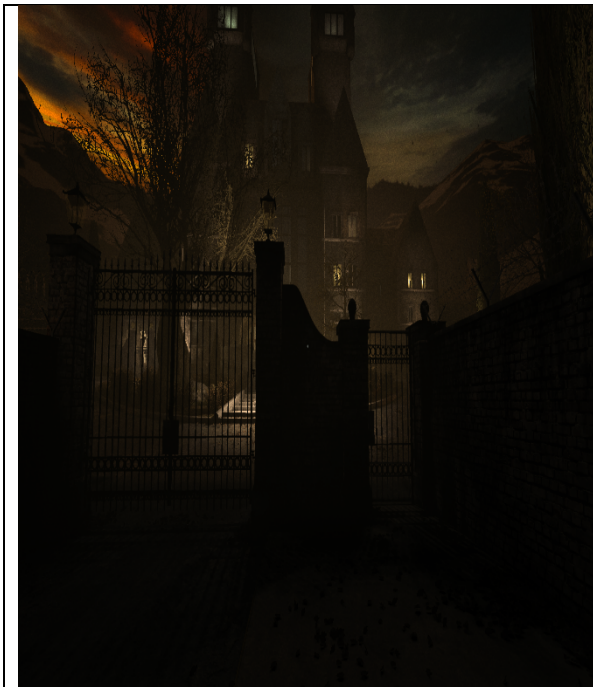


Figure 39. Proportion de Colère par zones d'intérêt.

Dans le tableau 7 ci-dessous, nous retrouvons les images prises pendant le jeu, les noms des scènes ainsi que l'émotion ou les émotions dans lesquelles elles sont impliquées.

Tableau 7. Captures de jeu.



Grille d'entrée - Tristesse



Premier jardin - Colère



Camions - Peur



Deuxième jardin - Surprise



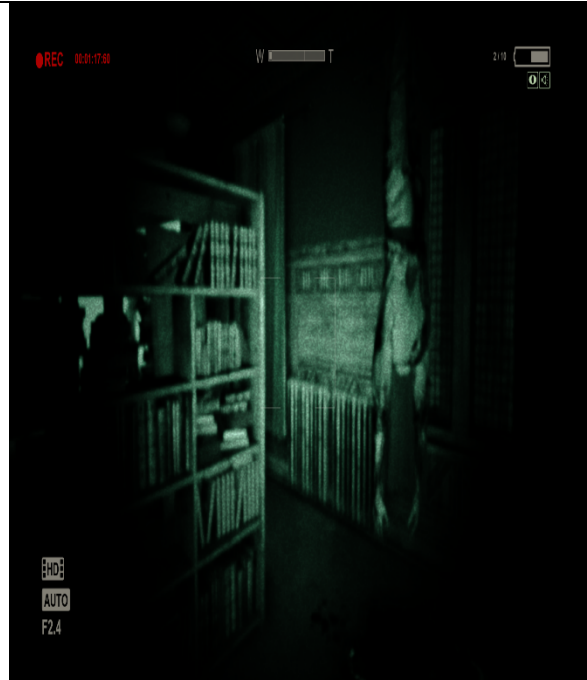
Échelle - Colère



Cuisine - Peur, Tristesse



Sang conduite aération - Peur



Bibliothèque - Peur, Joie



Je... Je vois, Dieu de Misericorde, vous m'avez envoyé un apôtre. Veillez sur votre vie, mon fils. Vous avez une mission.

Personnage alternatif - Dégout, Joie



Sang fin séquence - Tristesse

6.2 La web application Émograph

Dans cette section, nous présentons les interfaces de l'application créée et le mode de fonctionnement. La figure 40 montre la page d'accueil de l'application. Sur cette page nous pouvons distinguer les 3 modules proposés. Le module Nostradamus a été ajouté par la suite, nous reviendrons dessus dans le chapitre 7. Cette section concerne uniquement les modules d'analyse émotionnelle et de visualisation.



Figure 40. Fenêtre d'accueil d'Émograph

6.2.1 Le module d'analyse émotionnelle (Émotrigger)

Nous allons illustrer dans cette section le fonctionnement de ce module expliqué dans la section 5.2 du mémoire. Rappelons que le module propose deux types d'analyses. Celle par participants et celle par zones d'intérêt.

Analyse par participants

La figure 41 présente l'interface de l'analyse par participants dans le cas où l'utilisateur sélectionne le participant P02 et l'émotion Joy. L'utilisateur dispose de 2 listes déroulantes sur la gauche. Il sélectionne le participant et l'émotion de son choix. Le résultat de la requête de l'utilisateur s'affiche sur la droite, nous observons l'image de jeu extraite de la vidéo du

participant à partir du temps d'apparition du pic qui est affiché dans le tableau plus bas. Dans le tableau nous retrouvons les informations de la scène telles que l'intensité maximale de l'émotion, le nom, la description et les actions/défis de la scène. Sur cette figure il s'agit de la scène "Bibliothèque". Dans le cas où il ne sélectionne pas d'émotion en particulier, le module lui affiche l'émotion dominante dont le pic à l'intensité la plus élevée. De même, l'utilisateur peut ne pas choisir de participants en particulier, alors le module lui présentera l'intensité la plus élevée de l'émotion choisie ainsi que les informations présentées sur la figure 41.

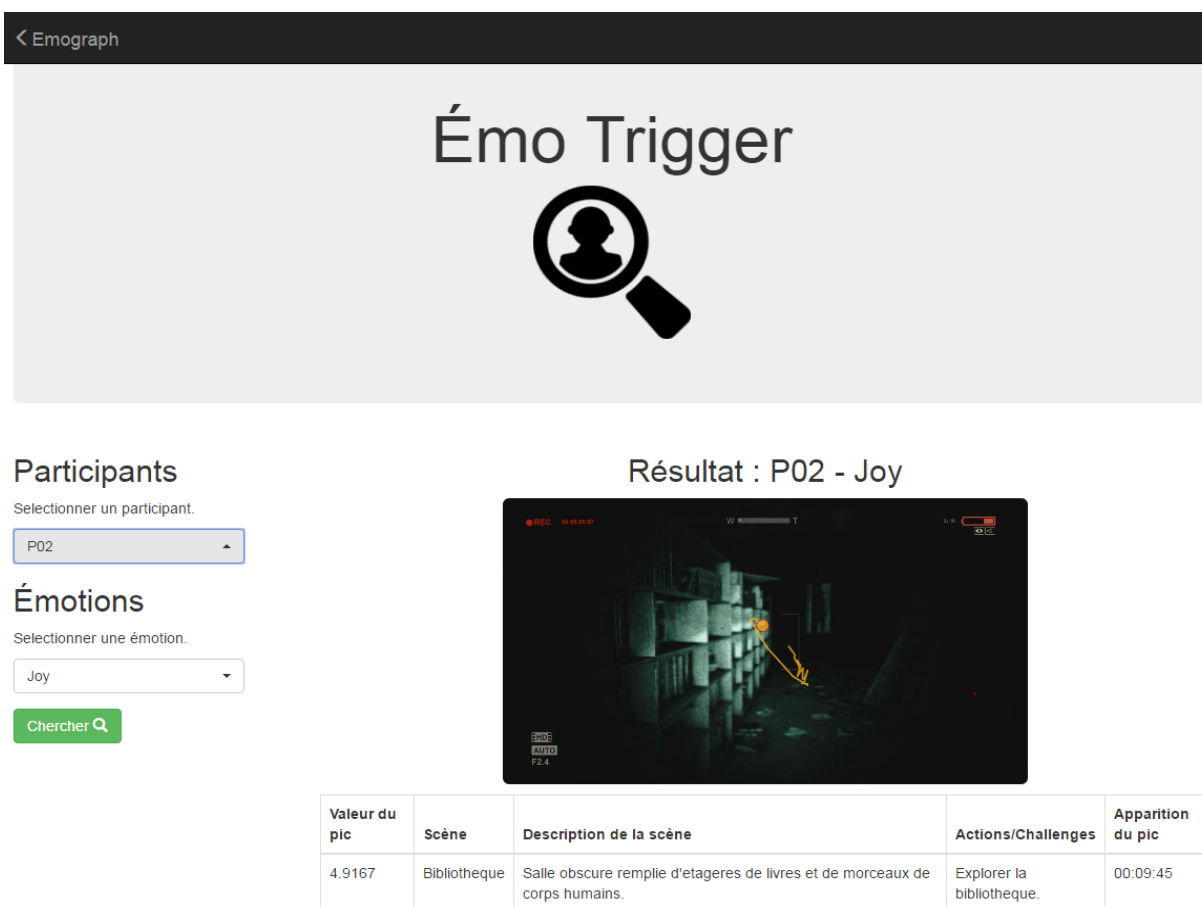


Figure 41. Interface du module d'analyse émotionnelle par participants

Analyse par zones d'intérêt

La figure 42 montre l'interface du module d'analyse émotionnelle par zones d'intérêt. Cette interface propose à l'utilisateur de choisir la zone d'intérêt dont il désire afficher les informations. Le résultat de la requête s'affiche sur la droite. Les informations affichées dans le

tableau sont une description de la zone d'intérêt, les actions ou challenges de celle-ci, les émotions qu'elle a générées, le nombre de participants qui l'ont regardée. De plus elle affiche le pic maximal exprimé ainsi que l'émotion correspondante également le participant qui l'a exprimé et le moment d'apparition de ce pic. Le module affiche également une image globale représentative de la scène concernée. Sur la figure 42 il s'agit de la scène "Bibliothèque".

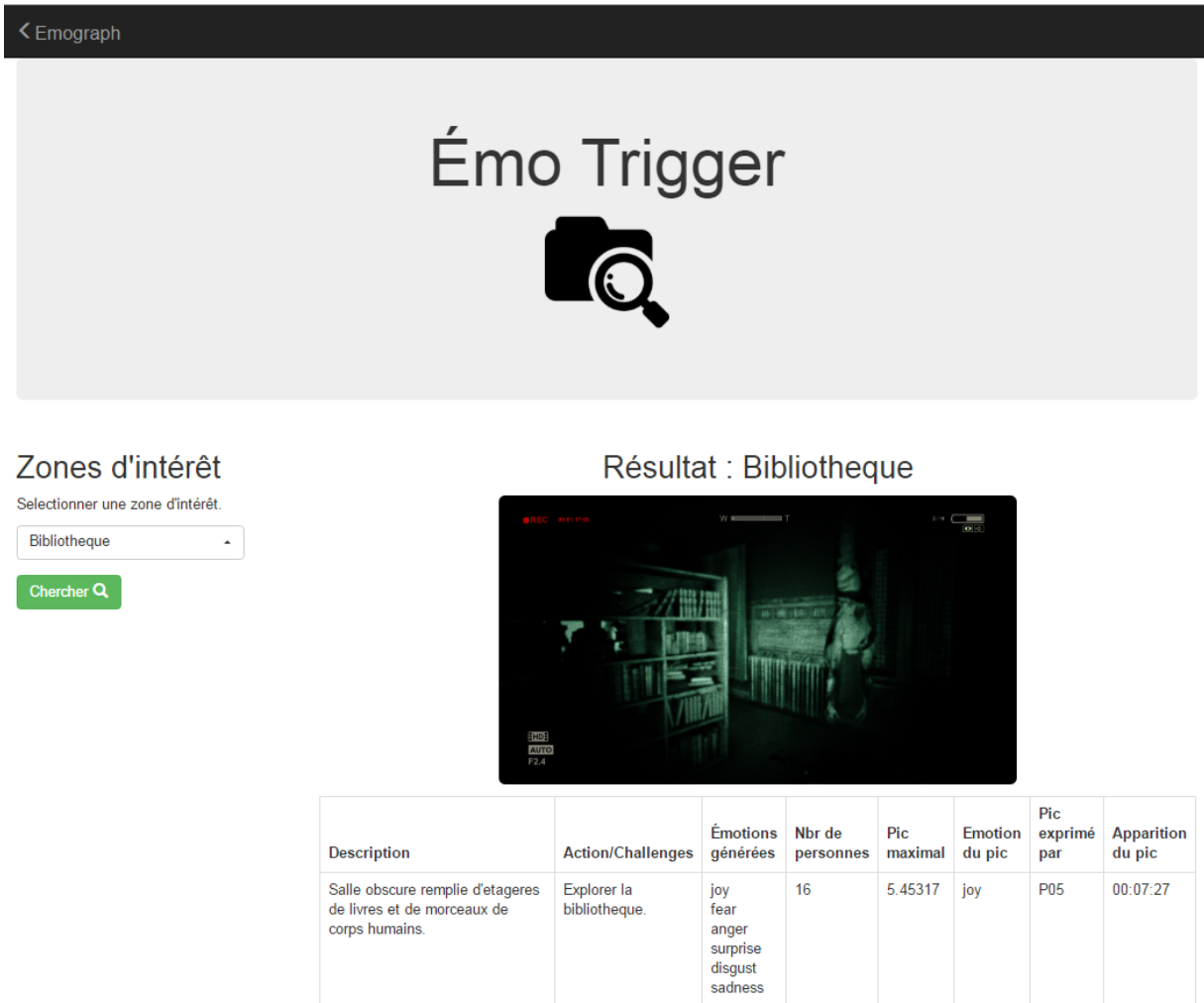


Figure 42. Interface du module d'analyse émotionnelle par zones d'intérêt

6.2.2 Le module de visualisation

La figure 43 nous fait voir l'interface du module de présentation graphique. Ce module propose à l'utilisateur de choisir parmi les participants celui dont il veut afficher le graphe. Une

fois le participant choisi, le graphe qui lui correspond s'affiche et nous pouvons lire à chaque point l'émotion dominante. Sur cette figure l'utilisateur a sélectionné le participant P02. De plus, un clic sur un point affiche l'image de jeu du pic de l'émotion concernée dans la scène. La figure 44, présente le résultat du clic de l'utilisateur sur le point "Bureau entree" à la figure 43.

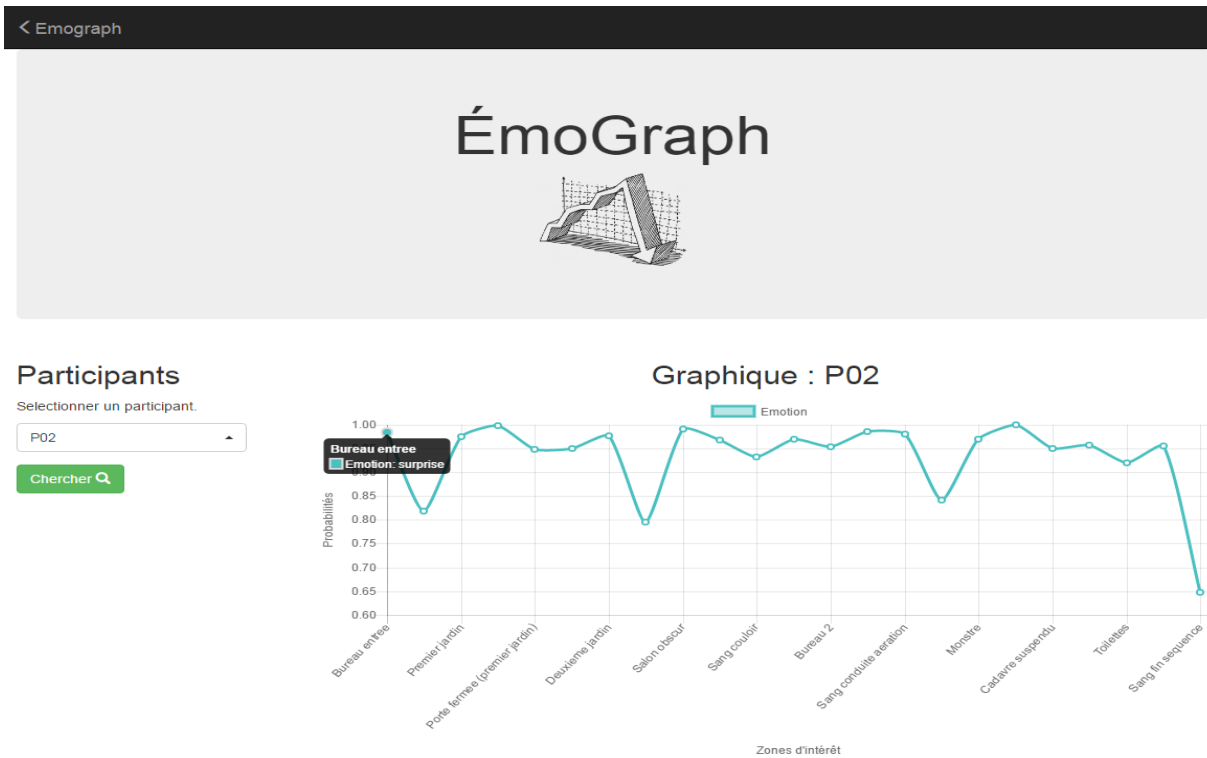


Figure 43. Interface du module de visualisation.

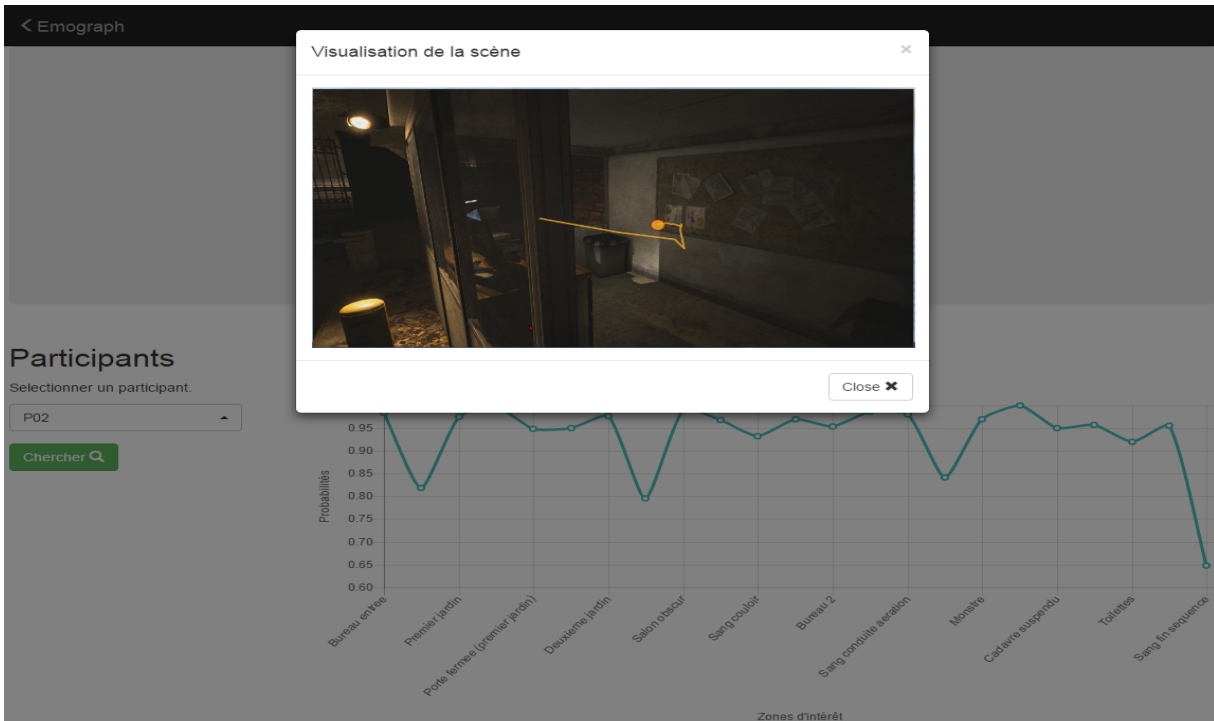


Figure 44. Image du jeu.

6.3 Les interprétations des résultats

Outlast est un tour virtuel à travers des halls de terreur, même jouer à ce jeu en plein jour rend les joueurs mal à l'aise compte tenu de son univers de jeu et de ses graphismes intensément immersifs et détaillés. Le gameplay étant relativement simple, l'arme principale d'*Outlast* se trouve dans son univers de jeu. Ce jeu crée une expérience qui fait monter le niveau d'adrénaline au détriment des expériences d'horreur et de peur en tant que telles. Dès le début du jeu, les joueurs se retrouvent dans une ambiance oppressive, plongé dans l'inconnu dépourvu de protection. Les portes qui se ferment de façon inattendue, des traces de sang dans les couloirs. Ne sachant pas exactement d'où la menace pourrait arriver, le joueur imagine le pire et reste sur ses gardes constamment. Le contexte de l'expérience et le jeu étant immersifs les joueurs ont tendance à se projeter dans l'environnement pour vivre l'expérience. Ce dernier se prend de sympathie pour l'avatar qu'il incarne et vit le moment avec lui.

Nous pouvons interpréter le fort taux de colère reconnu sur le visage des participants du fait de cet environnement oppressif dans lequel les joueurs évoluaient. En effet la vulnérabilité du joueur est exposée et contrairement aux autres jeux d'horreur, *Outlast* ne donne aucun moyen de défense aux joueurs. Cette absence de boutons d'attaque garantie que chaque ombre et chaque son insuffle une tension absolue. Cette atmosphère si pesante et imprévisible engendre un genre de tension, qui est une réponse humaine involontaire pour faire face à la situation vécue. C'est cette tension que nous avons observée chez les joueurs quand le système a reconnu la colère. Il serait intéressant dans de futures études de faire une analyse des données de conductance de la peau pour investiguer sur l'état de stress des joueurs.

De même la surprise observée vient de l'imprévisibilité et des changements soudains, brusques de l'univers de jeu. Par exemple lorsque l'agent de sécurité prend la parole dans la bibliothèque (voir figure 24), le joueur est dans un endroit dépourvu de lumière, rempli de morceaux de corps humain et de corps suspendu, cherchant à trouver la sortie de cette pièce il ne s'attend pas à ce que l'énigmatique corps suspendus dans la bibliothèque prononce des mots. Ou alors des lumières qui clignotent derrière une fenêtre dans le deuxième jardin le temps d'un battement de paupières du joueur.

Muni uniquement de sa caméra, Miles enregistre les événements et découvre progressivement l'histoire de l'asile dans les notes et documents enregistrés. Cette partie narrative joue un rôle important sur l'effet qu'a le jeu sur le joueur. En particulier dans l'expression de la tristesse sur le visage des participants. Par exemple la tristesse exprimée à la grille d'entrée (voir tableau 7) s'explique par le fait que le joueur vient de lire le document qui explique l'objectif de la présence de Miles à l'asile et décrit les activités abominables de l'hôpital qui à la base accueillait et soignait des personnes atteintes de maladies mentale.

La peur dans ce jeu en particulier dans la séquence étudiée se traduit par des sursauts de peur principalement encore appelés "jump-scare", des décors sanguinolents et l'audio du jeu qui est extrêmement efficace. Le journaliste ne parle pas mais respire. En pénétrant dans une pièce obscure, tournant dans un coin, sprintant pour échapper à une ombre qu'on croit dangereuse, chaque respiration du journaliste est obsessionnellement prononcée, avec des grosses prises d'air, des halètements accélérés et des exhalaisons épuisées en réaction aux actions et événements du jeu. La découverte des traces de pas dans une flaque de sang en dessous

de la conduite d'aération ou encore marcher dans la bibliothèque avec la vision nocturne de la caméra (voir tableau 7) altère la respiration de Miles. C'est terrifiant!

Les générateurs principaux des expressions de dégoût dans les jeux vidéo sont généralement les fluides corporels et le rejet des éléments qui ne sont pas dans les normes considérées communes par le joueur. Dans *Outlast* le sang et les parties de corps humain sont présents mais dans la séquence expérimentale, le dégoût n'a pas été exprimé fréquemment.

Les forts pics de joie reconnus viennent d'un sentiment de soulagement d'avoir échappé au danger, aussi quand le joueur se sent rassuré d'avoir traversé une partie dangereuse il a l'impression de bénéficier d'un moment de répit. Certains participants exprimaient de larges sourires tout de suite après un événement surprenant ou terrifiant. Ils proviennent également du fait que la plupart des personnes jouent aux jeux vidéo d'horreur pour profiter des multiples plaisirs de se faire peur. De ce fait lorsqu'un événement terrifiant ou surprenant se termine, les joueurs sont satisfaits et l'expriment par des sourires ou des rires dépendamment du degré d'excitation. Par exemple l'entrée dans la bibliothèque se fait juste après la surprise ou la peur créée par un cadavre qui tombe du plafond dès l'ouverture de la porte.

De plus les résultats observés de notre étude ont confirmé les différences qui existent entre les joueurs selon leurs catégories (voir le tableau 6) mais également le genre. Les novices exprimaient en majorité de la surprise (29%), de la colère (25%), de la tristesse (19%), du dégoût (10%) et de la peur (7%). Alors que les intermédiaires et experts exprimaient principalement de la colère respectivement 37% et 53%, de la joie respectivement 22% et 21%, de la surprise 19% pour les 2 catégories et de faibles pourcentages par rapport à la peur, à la tristesse et au dégoût qui sont autour de 7% maximum. Cela peut s'expliquer par leur expérience des jeux vidéo qui fait qu'ils sont moins sensibles aux éléments ou événements du jeu peu importe l'environnement alors ils ont tendances à rester concentrés sur leurs objectifs et profitent de l'expérience. D'où le fait qu'ils ressentent moins de peur, de dégoût ou de tristesse et plus de joie et de colère.

CHAPITRE 7 La prédiction émotionnelle

Nous avons expliqué dans l'introduction de ce mémoire que les émotions sont étudiées également dans le cadre des recherches en systèmes computationnels intelligents pour simuler principalement le raisonnement humain chez les robots ou agents intelligents. Selon plusieurs chercheurs, les émotions jouent un rôle dans le processus de raisonnement. Il existe d'autres travaux qui se sont focalisés essentiellement sur la simulation de la génération de l'émotion au sein des systèmes computationnels intelligents produisant ce qu'on appelle les systèmes computationnels de l'émotion ou émotionnels. Ces systèmes s'inspirent des théories de l'émotion, principalement des théories d'évaluation cognitive. Dans leur compte rendu sur les modèles computationnels de l'émotion, (Slater, Moreton, Buckley, & Bridges, 2008) affirment que la plupart des publications dans le domaine se réfèrent principalement à, ou font partie de, trois grands projets de recherches dans ce domaine, à savoir, Cognition and Affect Project (CogAff) (Sloman & Poli, 1995), Tok Project (Bates, Loyall, & Reilly, 1992), et Émile Project (Gratch, 2000).

Le CogAff a été réalisé à l'université de Birmingham. Son but global est de modéliser le comportement d'un agent pour les jeux. Une des caractéristiques de l'architecture développée est la capacité de simuler l'émotion. Les auteurs de ce travail ont suggéré une implémentation à l'aide des techniques d'intelligence artificielle comme la machine à états finis et les réseaux de neurones. Toutefois, l'aspect émotionnel dans l'architecture n'est pas basé sur une des théories de l'émotion.

Le projet Tok a été élaboré à l'université de Carnegie Mellon. C'est une architecture d'un agent intégré à un environnement virtuel, appelé l'environnement Oz. Tok contient un module comportemental orienté objectif appelé HAP (Loyall, Bryan, & Bates, 1991) et un autre module qui s'occupe des émotions appelé Em (Reilly & Bates, 1992). Ce dernier est essentiellement basé sur le modèle cognitif de l'émotion OCC qui porte le nom de ses auteurs (Ortony, Clore, & Collins, 1988). L'objectif de l'agent est spécifié dans le module HAP. Il peut avoir une émotion selon le niveau de réalisation ou non-réalisation de cet objectif. Dans cette architecture le processus de génération de l'émotion est dynamique, l'émotion change lorsque

les désirs ou les objectifs changent dans le temps. Cette architecture n'a pas été adaptée à un environnement multi-agent.

Le projet Émile est aussi une architecture agent qui se base sur l'architecture du module Em (Reilly & Bates, 1992) et sur celle du "The Affective Reasoner" par (C. D. Elliott, 1992). L'agent est capable non seulement d'évaluer une situation par rapport à ses besoins à l'aide du modèle OCC, mais aussi d'essayer de s'adapter à elle en planifiant et en changeant d'objectifs (Gratch & Marsella, 2003). Ces objectifs peuvent aussi être influencés par la présence et par l'action d'autres agents. Ainsi, Émile permet une simulation dans un contexte où cohabitent plusieurs agents. Cette architecture a été étendue et renommée EMA (Emotion and Adaptation) (Gratch & Marsella, 2004).

Il existe d'autres modèles computationnels de l'émotion intégrés à des architectures d'agents intelligents comme Poepke and Robot3¹⁸, SAFIRA¹⁹, etc. Les chercheurs ont rapporté que la majorité de ces travaux, comme dans les projets Tok et Émile, sont basés sur le modèle OCC. Ce dernier est décrit de manière à faciliter sa conversion dans des logiciels agents. Le modèle OCC a été surtout utilisé par les modèles computationnels pour simuler la génération de l'émotion de l'agent dans un environnement multi-agent. Un modèle computationnel basé sur l'OCC pourrait aussi servir à la compréhension du mécanisme qui soutient la génération de l'émotion d'un individu dans un contexte d'interaction homme-machine (Trabelsi & Frasson, 2010).

Dans ce mémoire, nous nous inspirons du modèle OCC dans l'optique de doter la machine d'une capacité de compréhension de la situation émotionnelle dans laquelle un participant se retrouve à chaque étape du jeu. Cette approche permettrait alors à la machine de prédire l'émotion d'un joueur. Ainsi à partir des émotions dominantes identifiées par la méthode Émograph, nous faisons une modélisation des étapes du jeu suivant les variables du modèle OCC. Dans les prochaines sections de ce chapitre, nous présenterons le modèle OCC, nous reviendrons sommairement sur l'apprentissage machine particulièrement celui supervisé et nous

¹⁸ <http://www.peopleandrobots.org/>

¹⁹ <http://www.sics.se/safira>

montrons la méthode que nous proposons pour prédire les émotions d'un joueur et son intégration dans l'application Émograph.

7.1 Le modèle OCC

Pour prédire l'émotion d'un individu nous avons exprimé la nécessité de traduire les informations émotionnelles sur la scène de jeu de manière compréhensible par la machine. Pour capturer ce processus nous avons alors utilisé un formalisme relativement simple et intuitif, permettant de décrire une scène en fonction de la perception et de l'évaluation cognitive faite par les joueurs. Le modèle d'évaluation cognitive de l'émotion d'Ortony, Clore et Collins (Ortony et al., 1988) est intéressant dans la mesure où il a été construit pour être facilement adapté aux travaux d'intelligence artificielle.

7.1.1 Les types de situations

Selon le modèle, les émotions sont provoquées chez une personne quand la situation influence ses objectifs, ses normes et valeurs ou fait appel à son goût. Dépendamment de l'évaluation faite de la personne, une situation génératrice d'émotions est basée sur 3 catégories :

- Un événement ayant des conséquences sur ses besoins et objectifs.
- Une action d'agent qui répond aux valeurs et normes ou qui en dévie.
- L'aspect d'un objet qui est en accord avec ses goûts.

Un événement qui a des conséquences sur l'individu lui-même ou sur autrui engendre en général un sentiment de plaisir ou de déplaisir. Il y a 3 classes de conséquences d'événements. Ces catégories sont, les conséquences sur autrui, les conséquences sur l'avenir personnel et les projets futurs, et les conséquences sur son bien-être actuel.

Un agent peut être une personne (soi-même ou une autre), une entité abstraite ou une institution (école, hôpital, entreprise...) qui peut agir en accord ou pas avec ses normes et valeurs et à laquelle on peut s'identifier ou non. L'action d'un agent crée une approbation ou une désapprobation sans qu'elle ait pour autant des conséquences sur soi.

Un objet engendre une appréciation ou un dégoût. Il peut être concret comme des habits, des livres, ou il peut être abstrait comme une idée ou un concept.

Cependant, il existe une quatrième catégorie implicite de situations. Elle correspond à une combinaison de deux catégories, l'action d'un agent et l'évènement. L'action d'un agent (approuvée ou désapprouvée) donne lieu à un évènement pouvant affecter les objectifs d'un individu (plaisir ou déplaisir).

La classification préliminaire précédente a permis à Ortony, Clore et Collins de définir 4 grandes catégories de réactions émotionnelles associées à ces 4 types de situations. La catégorie relative à l'évènement est elle-même divisée en trois sous catégories selon les conséquences de l'évènement (sur autrui, sur les projets, sur le bien-être). La figure 25 montre l'arbre d'illustration de la structure cognitive des réactions émotionnelles selon le type de réactions.

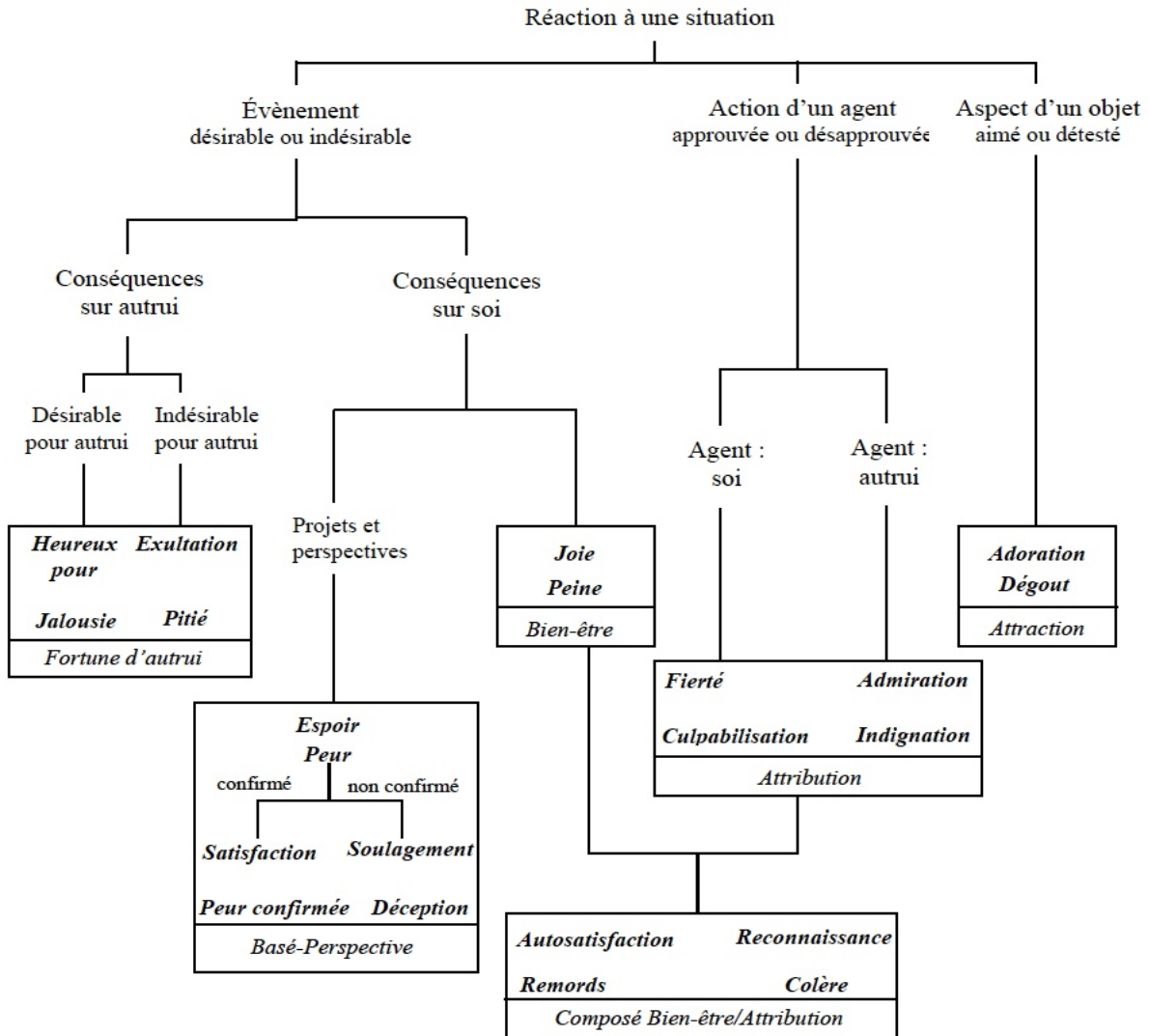


Figure 45. Modèle Ortony, Clore et Collins (Ortony et al., 1988)

La structure générale du modèle présentée sur la figure 45, donne, selon les types de situations, à six groupes d'émotions qui sont : Fortune d'autrui, Basé-Perspective, Bien-être, Attribution, Composé Bien-être/Attribution et Attraction. Ces groupes sont composés de paires d'émotions opposées, au total il existe 11 paires soit 22 émotions englobées par le modèle dont les 6 émotions de base sur lesquelles portent notre recherche.

7.1.2 Le lien entre les types de situations et les émotions

Le modèle fait une description semi-formelle des règles entre chaque paire d'émotions et les types de situations. Le tableau 8 montre les situations correspondantes à chacune des paires

d'émotions considérées. Les auteurs ont ainsi caractérisé de manière semi-formelle différentes situations et à associer à chacune l'émotion correspondante.

Tableau 8. Relations entre les émotions et les types de situations.

Émotion	Situation	Émotion	Situation
Heureux pour autrui	Un événement désirable par autrui et souhaité par soi (pour autrui).	Pitié	Événement indésirable par autrui et non souhaité par soi (pour autrui).
Exultation	Un événement indésirable par autrui et souhaité par soi.	Jalousie	Événement désirable par autrui et non souhaité par soi.
Espoir	La perspective d'un événement désirable.	Peur	La perspective d'un événement indésirable.
Satisfaction	La confirmation de la perspective d'un événement désirable.	Peur confirmée	La confirmation de la perspective d'un événement indésirable.
Joie	Un événement désirable	Peine (Tristesse)	Un événement indésirable
Fierté	Une action louable de soi n'ayant pas de conséquences sur soi.	Culpabilisation	Une action répréhensible de soi n'ayant pas de conséquences sur soi.
Admiration	Une action louable d'autrui n'ayant pas de conséquences sur soi.	Indignation	Une action répréhensible d'autrui n'ayant pas de conséquences sur soi.

Reconnaissance	Une action louable d'autrui ayant de bonnes conséquences sur soi.	Colère	Une action répréhensible d'autrui ayant de mauvaises conséquences sur soi.
Autosatisfaction	Une action louable de soi ayant de bonnes conséquences sur soi.	Remords	Une action répréhensible de soi ayant de mauvaises conséquences sur soi.
Adoration	Un objet attractif.	Dégoût	Un objet répulsif.

7.1.3 Les variables d'évaluation

Les évaluations des situations se rapportent à des facteurs ou variables dites cognitives. Ces variables caractérisent l'interprétation d'une situation par une personne. La réponse émotionnelle est différente du fait d'une situation désirable ou non, attendue ou inattendue, etc. Le modèle OCC dispose de ses propres variables descriptives, elles nous serviront à prédire la réponse émotionnelle des participants. Ces variables sont divisées en 3 catégories :

- Les variables globales, communes à toutes les situations, incluant la surprise et le sens de la réalité.
- Les variables centrales, chacune d'elle est associée à un type de situation.
- Les variables locales, elles aussi propres à un type de situation. Cependant, elles diffèrent dans le cas d'un événement selon que ce dernier a des conséquences sur autrui ou sur les projets de l'individu.

Le tableau 9 suivant présente les variables d'évaluation du modèle OCC et leur signification.

Tableau 9. Les variables d'évaluation du modèle OCC

	Variable d'évaluation	Signification
Variables globales	Surprise	Une situation attendue ou inattendue.
	Sens de la réalité	Une situation réelle et vécue ou irréelle et imaginaire.
Variables centrales	Désirabilité	L'occurrence d'un événement est désirable ou indésirable.
	Approbation	L'action d'un agent est louable ou condamnable.
	Attirance	Un objet (concret, idée, etc.) est attirant ou répugnant.
Variables locales	Désirabilité par autrui	L'occurrence d'un événement est désirable ou indésirable par autrui.
	Estime pour autrui	Autrui est apprécié ou méprisé.
	Mérite d'autrui	Autrui mérite ou ne mérite pas ce qui s'est passé.
	Vraisemblance	L'événement anticipé va ou ne va pas se produire.
	Réalisation	L'événement anticipé s'est ou ne s'est pas finalement réalisé.
	Effort	Les efforts fournis pour que l'événement anticipé soit évité ou soit réalisé.

	Agent	Réalise l'action, soi ou autrui.
	Puissance du lien	Lien avec l'agent qui réalise l'action
	Déviation	L'action de l'agent sort de la norme (positivement ou négativement)
	Disposition	Disposé ou non disposé à être attiré par l'objet.
	Familiarité	L'objet est familier ou étranger.

Rappelons que les variables globales sont incluses dans toutes les situations, alors que les variables centrales et locales sont particulières à une situation donnée pour en restituer son contenu informationnel. Le tableau 10 ci-dessous, montre les variables centrales et locales impliquées à chaque type de situation.

Tableau 10. Variables centrales et locales impliquées à chaque type de situation.

Situations Variables	Événement ayant des conséquences sur			Action d'un agent	Action d'un agent + événement	Aspect d'un objet
	Autrui	Plans	Bien-être			
Désirabilité	✓	✓	✓		✓	
Approbation				✓	✓	
Attirance						✓
Désirabilité par autrui	✓					
Estime pour autrui	✓					

Mérite d'autrui	✓					
Vraisemblance		✓				
Réalisation		✓				
Effort		✓				
Agent				✓	✓	
Puissance du lien				✓	✓	
Déviation				✓	✓	
Disposition						✓
Familiarité						✓

Ce modèle permet l'identification des variables impliquées et permet de représenter une scène du jeu par des variables précises. Il s'agit d'une description informelle d'une scène qui a engendré une émotion. La prochaine section de ce mémoire propose une formalisation des scènes à partir de cette description, c'est à dire comment faire une représentation par une combinaison de valeurs numériques.

7.1.4 La description formelle des scènes

La section 7.1.3 nous indique qu'une scène est décrite à partir de 3 types de variables. Ainsi nous avons défini pour chaque variable les valeurs qu'elles peuvent prendre. Nous avons défini ces échelles en nous basant sur l'interprétation des valeurs numériques produites du système d'analyse des expressions faciales. Les variables globales (surprise, sens de la réalité) prennent des valeurs booléennes par exemple, une situation est soit attendue ou inattendue. Les variables centrales et locales prennent des valeurs sur une échelle entre -1 et 1. Il existe des variables comme "Disposition" qui ne peuvent prendre que 3 valeurs, soit l'on est

extrêmement, moyennement ou pas du tout disposé à être attiré par l'objet. Cette échelle est présentée dans le tableau 11. Le tableau 12, montre les valeurs attribuées aux variables.

Tableau 11. Échelle des valeurs prises par les variables centrales et locales.

Valeur	Interprétation
-1	Intensité extrêmement négative de la variable (une <i>non-désirabilité</i> de l'occurrence d'un événement).
-0.5	Intensité moyennement négative de la variable (une <i>non-désirabilité moyenne</i> de l'occurrence d'un événement).
0	La variable n'est pas incluse dans la description de la situation donnée.
0.5	Intensité moyennement positive de la variable (une <i>désirabilité moyenne</i> de l'occurrence d'un événement).
1	Intensité extrêmement positive de la variable (une <i>désirabilité</i> de l'occurrence d'un événement).

Tableau 12. Les variables globales, centrales et locales et les valeurs associées.

	Variable d'évaluation	Valeurs
Variables globales	Surprise	1 (Vrai) ou 0 (Faux).
	Sens de la réalité	1 (Vrai) ou 0 (Faux).
Variables centrales	Désirabilité	1, 0.5, 0, -0.5, -1.
	Approbation	1, 0.5, 0, -0.5, -1.
	Attirance	1, 0.5, 0, -0.5, -1.

Variables locales	Désirabilité par autrui	1, 0.5, 0, -0.5, -1.
	Estime pour autrui	1, 0.5, 0, -0.5, -1.
	Mérite d'autrui	1, 0.5, 0, -0.5, -1.
	Vraisemblance	1, 0.5, 0, -0.5, -1.
	Réalisation	1, 0.5, 0, -0.5, -1.
	Effort	1, 0.5, 0, -0.5, -1.
	Agent	1 (Soi) ou 0 (autrui).
	Puissance du lien	1, 0.5, 0.
	Déviation	1, 0.5, 0, -0.5, -1.
	Disposition	1, 0.5, 0.
	Familiarité	1, 0.5, 0, -0.5, -1.

Ainsi, une scène peut être décrite formellement par un vecteur de 16 valeurs numériques qui lui est propre. Dans la section suivante nous verrons comment cette transformation est utilisable par un ordinateur pour prédire l'émotion d'un individu.

7.2 L'apprentissage machine

Dans cette section, nous présentons les principes de base de l'apprentissage machine. Nous nous focaliserons sur l'apprentissage supervisé, les algorithmes des *k-plus proches voisins* et des *forêts d'arbres de décision*. Ces algorithmes sont ceux utilisés dans notre étude.

Chez les êtres humains, l'apprentissage est un processus qui vise à améliorer ses performances au fur et à mesure de l'exercice d'une activité. Pour apprendre il faut s'entraîner,

obtenir une expérience cumulative de l'activité. L'expérience permet aux humains d'être plus efficaces dans les nouvelles tâches de la vie courante.

L'apprentissage machine est né de la volonté des chercheurs en intelligence artificielle de doter la machine de la capacité d'apprendre et de s'améliorer automatiquement dans l'exécution d'une tâche donnée (Mitchell, 1997) (Vapnik, 2000) (Bishop, 2006). Ces algorithmes suscitent de l'intérêt dans plusieurs disciplines et sont appliqués dans des domaines très variés tels que la neuroscience, les statistiques, la biologie etc..... Ils sont généralement appliqués à de grandes bases de données contenant plusieurs cas ou exemples du phénomène à étudier. Leur but consiste à identifier des patrons récurrents et des régularités implicites. Il faut préciser qu'apprendre ne signifie pas mémoriser, c'est être capable de généraliser à de nouveaux cas. L'apprentissage est au cœur des succès modernes de l'intelligence artificielle, par exemple en février 2011, Watson²⁰, le système d'IBM bat des champions humains au jeu télévisé *Jeopardy!*²¹ en se fondant sur l'apprentissage automatique à partir de données textuelles. Il n'y a pas que les jeux qui en bénéficient, les applications comme Google search, Facebook, Netflix, Siri etc... qui fonctionnent avec la reconnaissance de la parole, de l'écriture, etc... en profitent également. Les performances des algorithmes varient selon le domaine d'application et le type de problème à résoudre. À performances égales, on préfère l'algorithme le plus simple à implémenter, le plus facile à interpréter, le moins coûteux en termes de complexité spatiale et temporelle, parmi d'autres critères de choix. L'apprentissage s'attaque principalement à trois catégories de problèmes : l'apprentissage supervisé, non supervisé et par renforcement.

En ce qui concerne l'apprentissage supervisé, on dispose d'observations sous forme de couples d'entrées et de sorties. L'apprentissage consiste alors à chercher un lien entre les sorties et les entrées. Lorsque la cible est une étiquette de classes (indiquant à quelle classe l'entrée appartient, parmi plusieurs), il s'agit d'un problème de classification. Lorsque la cible est une (ou plusieurs) valeur réelle à prédire, alors il s'agit d'un problème de régression.

Par rapport à l'apprentissage non-supervisé, celle-ci s'applique à des observations qui n'ont pas de sorties explicites. Sa tâche consiste à décrire l'association entre les entrées et

²⁰ <http://www-03.ibm.com/innovation/ca/fr/watson/>

²¹ <https://www.jeopardy.com>

l'organisation des observations. Il existe 3 familles dans l'apprentissage non-supervisé. La catégorisation (l'algorithme des k-moyennes (MacQueen, 1967)), la réduction de dimensionnalité (l'analyse en composantes principales (Jolliffe, 2002)) et l'estimation de densité de probabilité (le modèle de mélange de densités gaussiennes (McLachlan & Basford, 1988)).

Dans l'apprentissage par renforcement (Sutton & Barto, 1998), il s'agit d'un agent intelligent artificiel qui doit apprendre à décider quelles actions effectuer dans un environnement changeant afin de maximiser une récompense totale. L'algorithme Q-Learning (Watkins & Dayan, 1992) est un exemple.

Dans notre étude, notre intérêt se porte sur l'apprentissage supervisé, particulièrement au problème de classification car notre tâche consiste à prédire les émotions d'un individu étant donné les descriptions formelles des scènes dont nous disposons.

7.2.1 L'apprentissage supervisé

En apprentissage supervisé, on suppose que les variables cibles sont reliées aux variables d'entrées par une fonction $f : \mathbf{x} \rightarrow \mathbf{y}$, sauf que f est inconnue. La tâche d'apprentissage consiste alors à chercher un algorithme qui produit une sortie qui est la bonne prédiction de la cible. Cet algorithme trouve la fonction f ou du moins propose une estimation viable. La plupart des algorithmes d'apprentissage nécessitent une représentation vectorielle des exemples, il s'agit d'un vecteur de valeurs numériques de taille d . Une entrée x est en général représentée par un vecteur de dimension d telle que $x \in \mathbb{R}^d$, avec d étant la dimensionnalité de l'entrée. Les dimensions prennent une place importante dans le problème d'apprentissage car elles déterminent s'ils sont applicables ou non en raison de leur complexité algorithmique. Dans la notation en général on utilise n pour le nombre d'exemples, d pour la dimension de l'entrée, c'est à dire le nombre de traits caractéristiques de l'entrée et m pour la dimension de la cible à prédire. Un ensemble de donnée sera souvent organisé dans une matrice $n \times (d + m)$. L'apprentissage supervisé est constitué essentiellement de 2 tâches.

La classification est utilisée lorsque la variable cible y admet un nombre fini de valeurs. Alors la fonction trouvée est alors appelée classifieur. Si la cible admet 2 valeurs possibles, la classification est dite binaire, dans le cas contraire (plus de 2 valeurs) elle est dite multiclasse.

La régression est utilisée lorsque la variable cible y est continue, dans ce cas la tâche consiste à prédire la valeur estimée de y (le y ‘‘moyen’’ qui correspond à x).

7.2.2 Les phases de l’apprentissage

L’entraînement

Cette phase consiste à trouver la fonction f qui en premier lieu fonctionne bien sur l’ensemble d’entraînement, puis en second lieu qui est utilisable sur de nouveaux points (de test) qui n’ont jamais été vus dans l’ensemble d’entraînement. L’objectif dans cette phase est de minimiser le risque empirique, c’est à dire entraîner le modèle (adapter ses paramètres) pour qu’il fasse le minimum d’erreurs sur l’ensemble d’entraînement. On fait référence à la fonction de perte, c’est elle qui mesure la qualité d’une prédiction. Elle peut prendre la forme d’une erreur quadratique dans le cas de la régression ou d’un pourcentage d’exemples mal classés dans le cas d’une classification. Un algorithme d’apprentissage correspond souvent à la combinaison des éléments suivants :

- La spécification d’une famille de fonction F telle que $f \in F$.
- Une manière d’évaluer la qualité de la fonction par l’utilisation de la fonction de perte L .
- Une méthode d’optimisation afin de minimiser la fonction de perte.

Il s’agit alors d’entraîner un modèle de manière à ce qu’il fasse un minimum d’erreurs sur l’ensemble d’entraînement. Toutefois l’erreur sur l’ensemble d’entraînement est un estimateur biaisé de l’erreur de généralisation, nous verrons comment traiter cette donnée dans le paragraphe suivant. Apprendre c’est choisir parmi un ensemble de fonctions. La capacité correspond à la taille de l’ensemble F de fonctions considérées. Le choix de l’ensemble de fonction et des hyperparamètres permettent de contrôler la capacité du modèle. Dans le cas où le modèle choisi à une capacité trop élevée on tombe dans le cadre d’un sur-apprentissage. En cas de capacité trop faible, on parle de sous-apprentissage. L’art de l’apprentissage consiste à

trouver un compromis entre le sous-apprentissage et le sur-apprentissage, c'est à dire que notre modèle doit avoir une capacité optimale. La figure 46 présente des courbes d'apprentissage, les points sont les exemples, et la courbe est une représentation graphique de la fonction f . Cette figure est une illustration du concept de capacité évoqué. Sur cette figure, la courbe de gauche représente le cas d'un sous-apprentissage, celle de droite, le cas d'un sur-apprentissage et celle du milieu représente le cas optimal.

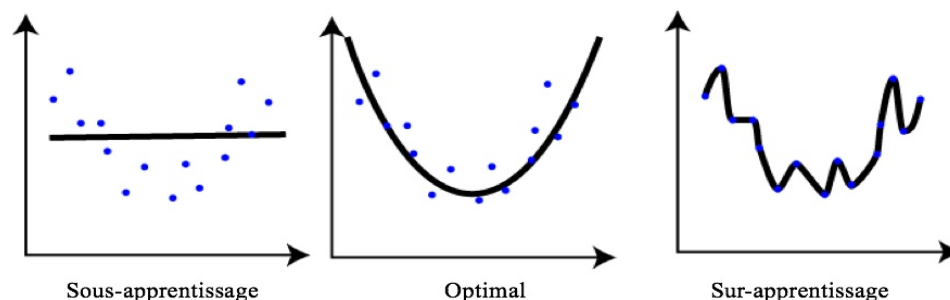


Figure 46. L'art de l'apprentissage.

L'entraînement/Validation

Rappelons que l'objectif est de bien généraliser sur de nouveaux exemples or les paramètres du modèle sont spécialisés pour minimiser l'erreur sur les exemples d'entraînement. Ainsi l'erreur d'entraînement n'est pas un bon estimé de l'erreur de généralisation. De ce fait nous devons passer par une étape de sélection de modèle pour choisir lequel sera utilisé sur un exemple de test et ainsi estimer d'une bonne manière l'erreur de généralisation de l'algorithme. Cette étape se fait de la manière suivante, on sépare nos données en 3, un ensemble d'entraînement de taille n , un ensemble de validation de taille n' et un ensemble de test de taille m . On suppose que l'entraînement donne suite à plusieurs alternatives de modèles identifiées par des paramètres additionnels appelé *hyperparamètres*, un exemple d'hyperparamètre est le k de voisins dans l'algorithme des *k-plus proches voisins*. Pour chaque valeur d'hyperparamètres considérée :

- On entraîne le modèle pour minimiser l'erreur sur l'ensemble d'entraînement.
- On évalue l'erreur sur l'ensemble de validation.

- On utilise finalement la valeur des hyperparamètres donnant l'erreur la plus basse sur l'ensemble de validation.

On évalue la performance de généralisation en mesurant les erreurs sur l'ensemble de test qui n'a jamais été regardé pendant l'entraînement / validation. Il est important de préciser que l'évaluation de performance finale s'effectue sur des données qui n'ont servi ni pour l'entraînement, ni pour la sélection de modèle.

Si l'on ne dispose pas d'un nombre important de données on peut utiliser la technique de *validation croisée en K blocs* (*K-fold cross validation* en anglais) qui est fondée sur les techniques d'échantillonnage. L'idée est de répéter plusieurs fois la procédure entraînement/test (respectivement validation) en divisant différemment l'ensemble de données. On divise l'ensemble de données en K échantillons, puis on sélectionne un des K échantillons comme ensemble de validation et les $(K-1)$ autres comme ensemble d'entraînement. Après avoir évalué la performance de généralisation, l'opération se répète ainsi K fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. Dans le cas où $K=n$, cette technique est appelée *leave-one-out cross validation* (LOOCV), la figure 47 présente ce procédé. Dans ce mémoire nous utiliserons cette technique comme méthode de validation.

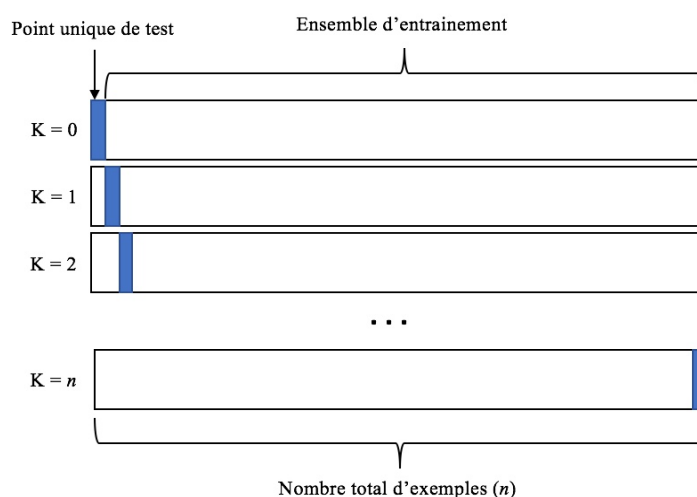


Figure 47. Procédé de la technique *Leave-one-out cross validation*.

7.2.3 Les algorithmes utilisés

Les k-plus proches voisins pondérés

La technique des *k plus proches voisins* (*kPPV*) (Devijver & Kittler, 1982) est une méthode à base de voisinage. L'idée consiste à faire voter les voisins de l'exemple de test. C'est une méthode de type non-paramétriques. Le *kPPV* est l'un des algorithmes les plus simples à implémenter. Soit n observations telles que $(\mathbf{x}_i, \mathbf{y}_i)$ avec $i = 1 \dots n$, et une nouvelle observation (\mathbf{x}, \mathbf{y}) dont la cible \mathbf{y} est à prédire. L'algorithme fonctionne comme tel :

- Calcule les distances entre \mathbf{x} et les \mathbf{x}_i pour $i = 1 \dots n$.
- Identifie les voisins les plus proches de \mathbf{x} , $(\mathbf{x}_{(1)}, \mathbf{y}_{(1)}), \dots, (\mathbf{x}_{(k)}, \mathbf{y}_{(k)})$.
- Classe \mathbf{x} en lui associant la classe majoritaire parmi ses voisins.

Formellement, pour la classification multiclasse (avec $Y_i \in 1 \dots m$) et $V(\mathbf{x})$ étant l'ensemble des k plus proches voisins de \mathbf{x} dans l'ensemble d'apprentissage. On a :

$$f(\mathbf{x}) = \operatorname{argmax} \left(\frac{1}{\sum_{i=1}^n w_i} \sum_{i=1}^n w_i \operatorname{onehot}_m(Y_i) \right) \text{ Avec : } w_i = I_{\{\mathbf{x}_i \in V(\mathbf{x})\}}.$$

Onehot est un vecteur de m composantes dont le bit de la composante dans la parenthèse est mis à 1 et le reste à 0. I la fonction d'identification qui renvoie 1 si la condition est vraie, 0 sinon. Le nombre k est un hyperparamètre qui peut être spécifié ou bien choisi après un test de différentes valeurs sur un ensemble de validation. Cet algorithme, basé sur une approche à base de voisinage, peut être un avantage par rapport aux autres algorithmes paramétriques lorsque la fonction à apprendre se révèle très complexe (Mitchell, 1997). Il y a plusieurs variantes de l'algorithme selon, par exemple, la mesure de distance (Euclidienne, Minkowski, etc.) et la pondération des classes.

Nous avons présenté la version classique de l'algorithme. Celle-ci a ensuite été améliorée en donnant à chacun des k exemples des poids dans la classification dépendamment de leur distance par rapport à l'exemple \mathbf{x} . Une mesure communément utilisée est la distance Euclidienne standard ou L2. Ainsi, plus une observation est proche de l'exemple à classer, plus elle a du poids, et plus elle a de l'influence dans la classification. Cette variante de l'algorithme est appelée le *k Plus Proches Voisins pondéré* (*kPPV*) ou *pondéré-distance* (*Distance-Weighted*

en anglais). Une des méthodes serait de donner comme poids à un exemple l'inverse de sa distance au carré (Mitchell, 1997), c'est celle qui est utilisée dans notre cadre. Dans ce cas, la classification se fait comme suit :

- Si la distance entre l'exemple à classer et un exemple d'entraînement est égale à 0, la classe prédite est celle de ce dernier.
- S'il existe plusieurs exemples identiques à l'exemple à classer, on lui affecte la classe majoritaire parmi eux.
- S'il n'existe pas d'exemples identiques, appliquer le *kPPV pondéré*.

Selon Mitchell (1997), le *kPPV pondéré* peut s'avérer très utile pour la résolution de plusieurs problèmes pratiques, car il est robuste contre le bruit des données et efficace lorsqu'il utilise de grandes quantités de données.

Notre choix s'est porté sur cet algorithme car outre sa simplicité relative d'implémentation, ce type d'algorithme, non-paramétriques ne nécessitent pas un entraînement permanent du modèle de classification contrairement à ceux paramétriques qui dont la phase d'entraînement nécessite l'utilisation de tout l'ensemble d'entraînement. De plus les algorithmes paramétriques ne possèdent pas d'alternatives pour un apprentissage séquentiel. Une autre raison est que le *kPPV* permet de limiter à *k* le nombre d'exemples à prendre en compte et il s'agit d'un des algorithmes les plus populaire et performants pour résoudre différents types de problèmes.

Les forêts d'arbres de décisions

L'objectif d'un arbre de décision est de réaliser une classification en utilisant une séquence de questions fermées, les questions sont organisées dans un arbre. La structure d'un arbre de décision est composé de nœuds et de branches (Breiman, Friedman, Stone, & Olshen, 1984). Un nœud correspond à une variable d'entrée et chaque branche qui en sort constitue une valeur possible de cette variable (si la variable est discrète) ou un intervalle (si la variable est continue). Les nœuds feuilles de l'arbre représentent généralement les valeurs dominantes (plus fréquentes) de la variable cible. L'arbre est interprété comme un ensemble de règles ou chemins, allant de la racine aux feuilles. Les autres chemins s'interprètent d'une manière similaire. Tous les chemins sont créés de manière à différencier au mieux les exemples d'entraînement entre les

nœuds feuilles (valeurs de variable cible). Parmi les avantages des arbres de décisions, nous pouvons citer son efficacité à l'entraînement et à la classification ainsi que son interprétabilité relativement simple. Un autre avantage majeur est la construction automatique de tels arbres par des algorithmes d'apprentissage supervisé. Ces algorithmes sélectionnent automatiquement les variables discriminantes à partir de données non-structurées et potentiellement volumineuses. Ils peuvent ainsi permettre d'extraire des règles logiques de cause à effet qui n'apparaissent pas initialement dans les données brutes.

Nous avons présenté une méthode de classification par l'utilisation d'un arbre de décision, une extension de cette méthode en apprentissage automatique est la construction de *forêts d'arbres de décisionnels* (*random forest classifier* en anglais) dont l'algorithme a été développé par (Breiman, 2001). Le terme *forêts aléatoires de décision* (*random decision forest*) vient à l'origine de (Ho, 1995). Cette méthode combine l'idée de "bagging" et de sélection aléatoire de variables afin de construire une collection d'arbres de décisions avec un contrôle sur la variation. Il s'agit d'un apprentissage effectué sur plusieurs arbres de décisions et la classification est faite en faisant voter l'ensemble des arbres en choisissant la réponse majoritaire. Chaque arbre est construit en utilisant l'algorithme suivant :

- Soient N le nombre d'exemples d'entraînement et M le nombre variables dans le classifieur.
- Le nombre m de variables d'entrées à utiliser pour déterminer la décision à chaque nœud de l'arbre est donné ; m doit être beaucoup plus petit que M .
- Choisir un ensemble d'entraînement pour cet arbre. L'ensemble est construit en choisissant les exemples N fois avec remplacement provenant de tous les N exemples d'entraînements disponibles. Utiliser le reste des exemples pour estimer l'erreur de classification de l'arbre en prédisant leurs classes.
- Pour chaque nœud de l'arbre, choisir aléatoirement m variables sur lesquelles baser la décision à chaque nœud. Calculer le meilleur découpage basé sur ces m variables dans l'ensemble d'entraînement.
- Chaque arbre est entièrement construit et non élagué.

Pour la prédiction :

- Un nouvel exemple est envoyé à travers l'arbre en partant de la racine. On lui assigne la classe de l'exemple d'entraînement du nœud terminal dans lequel il finit son chemin.
- Cette procédure est répétée sur tous les arbres de la collection et la majorité des votes de tous les arbres est enregistrée comme prédiction de la forêt d'arbres de décisions.

Notre choix s'est porté sur cet algorithme à causes de ses nombreux avantages. Nous pouvons citer entre autres, le fort taux de précision enregistré sur plusieurs jeux de données par exemple cette étude de (Shanthi & Ramani, 2012) qui rapportent un score de 95% de précision en utilisant une classification spécifique des positions assises dans une voiture pour trouver les schémas d'accidents dans les données globales des accidents de la route. De plus, il peut supporter un grand nombre de dimensions sans avoir à les traiter au préalable aussi il génère un estimateur interne non biaisé de l'erreur de généralisation durant le processus de construction de la forêt. Une autre raison qui n'est pas moindre est la résistance particulière de cet algorithme au sur-apprentissage.

7.3 Les approches proposées

Des sections 7.1 et 7.2 nous avons montré comment décrire de manière formelle une scène de jeu et présenté la notion d'apprentissage machine et introduit le lien avec notre étude. Dans cette section nous présenterons notre approche qui a conduit à la création d'un nouveau module nommé Nostradamus pour l'application. Ce module permet aux concepteurs ou chercheurs d'évaluer si l'émotion d'un individu sera en accord avec celle qu'il souhaite produire dans une nouvelle scène. Outre cette fonctionnalité, il permet d'anticiper l'émotion d'un individu qui regarderait une scène qui a déjà été regardée par d'autres individus. La figure 48 présente l'évolution de l'architecture présentée sur la figure 24.

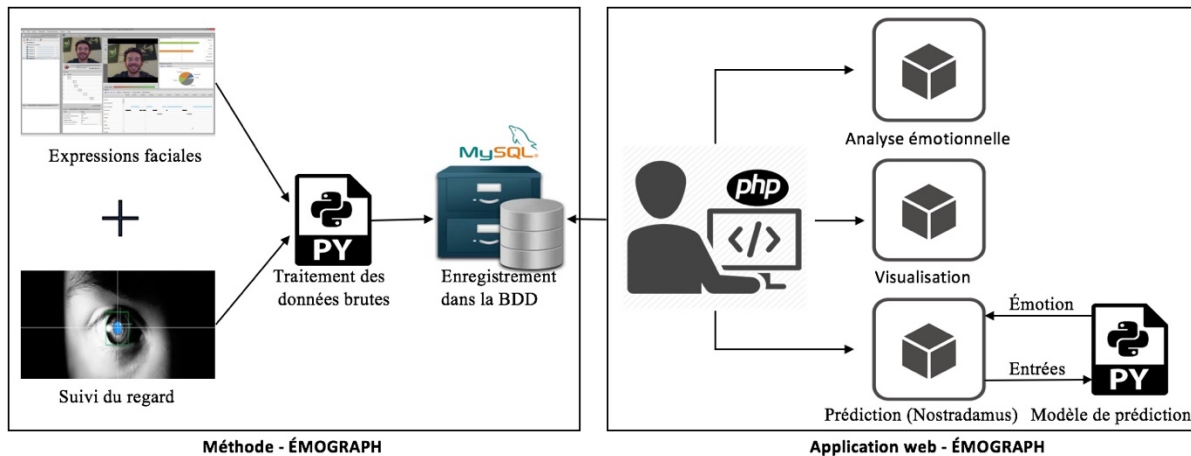


Figure 48. Architecture finale du système.

Dans le cadre d'un apprentissage supervisé (voir section 7.2), un ensemble d'entraînement contient plusieurs exemples du phénomène étudié. Chaque exemple est de la forme d'un couple de variables d'entrées et de sorties. Dans notre contexte, les descriptions formelles des scènes constituent des variables d'entrées de l'ensemble d'entraînement, il s'agit du vecteur de valeur numériques obtenu du modèle OCC, et les émotions des participants sont les variables de sorties. Cependant, les vecteurs de variables d'entrées ne se contiennent pas uniquement aux descriptions numériques des scènes.

L'importances des données sociodémographiques et des traits de personnalité

De nombreux psychologues s'accordent sur le fait que les réponses émotionnelles humaines dépendent de plusieurs facteurs. Des études comme celles de (Hess, 2001), (Bradley, Codispoti, Sabatinelli, & Lang, 2001) ont montré l'effet du genre de la personne sur les réponses émotionnelles des individus dans des situations données. À son tour, (Scherer, 2004) met en avant l'influence de l'âge sur les émotions. De même dans cette étude (Scherer, 1997) expose les effets du facteur culturel dans l'apparition des émotions. Également l'origine ethnique des personnes a une influence sur l'interprétation de l'information ainsi sur l'expression d'émotions comme l'expose cette étude de (Hess, Blairy, & Kleck, 2000). Ainsi les facteurs économiques, les croyances culturelles, les races, l'origine ethnique des individus jouent un rôle dans la différenciation des émotions chez les individus. Toutefois, le facteur le plus étudié comme cause de différenciation de réponse émotionnelle reste la personnalité comme le montre cette étude

(Scherer, Wranik, Sangsue, Tran, & Scherer, 2004). Des psychologues comme (Arnold, 1960), (Larsen & Ketelaar, 1991) affirment dans leurs études que les individus diffèrent selon leur prédisposition à vivre l'émotion. Par exemple les individus ayant des scores élevés du trait névrotique auraient tendance à être plus réactifs aux situations stressantes et à s'énerver et s'inquiéter plus qu'un autre (Schwebel & Suls, 1999).

Ainsi les informations sociodémographiques et les traits de personnalité sont nécessaires pour faire une bonne différenciation des émotions. Les traits de personnalité du *big five* représentent le modèle de caractérisation de la personnalité le mieux accepté et le plus commun utilisé dans la littérature en psychologie. Alors qu'il semble exister une multitude de variables de la personnalité, 5 se démarquent car elles donnent de larges explications sur les réponses d'une personne aux questions sur sa personnalité :

- L'extraversion (extroversion en anglais), les individus ayant un score élevé ont tendances à être ouvert et sociable.
- La stabilité émotionnelle plus connu sous le nom de névrotisme (neuroticism en anglais), correspond à une prédisposition au ressenti d'émotions négatives.
- L'agréabilité (agreeableness en anglais), caractérise les individus sympathiques et optimistes.
- La conscience (conscientiousness en anglais), ce trait caractérise les personnes minutieuse et appliquée.
- L'intellect/l'imagination (intellect/imagination en anglais), ce facteur est souvent appelé ouverture aux nouvelles expériences (openness to experience), les individus avec un faible score ont tendances à être traditionnels et conventionnels.

Ces dimensions sont mesurables par un test de personnalité sous forme de questionnaire, que nous avons présenté dans la section 4.1.2 du mémoire. Une fois que le participant remplit le questionnaire, une note entre 0 et 100 est attribuée à chaque trait de personnalité, ainsi la personnalité d'un individu est décrite par un vecteur de 5 dimensions. De ce fait, les variables d'entrées de l'ensemble d'entraînement deviennent alors une combinaison des informations sociodémographiques et des traits de personnalité du participant.

7.3.1 La construction des ensembles d'entraînements

Rappelons que la problématique de prédiction n'était une des problématiques de base de notre recherche. Elle est venue après que nous ayons mené les expériences, de ce fait notre protocole expérimental n'a pas pris en compte le facteur de description formelle des scènes. Dans le cadre de l'apprentissage supervisé (voir la section 7.2) un ensemble d'entraînement contient plusieurs exemples du phénomène étudié. Les valeurs des variables d'entrée correspondent à des caractéristiques décrivant un cas particulier du phénomène, et celles de sortie contiennent les réponses associées à ce cas. Dans notre contexte, l'ensemble d'entraînement est constitué d'un côté des descriptions des scènes, des informations sociodémographiques et des traits de personnalité du participant et d'un autre côté des émotions suscitées par chaque scène regardée.

Description numérique des scènes du côté du joueur

Pour pallier au fait de ne pas disposer des descriptions des scènes, nous avons contacté les participants et nous sommes placés dans la peau des joueurs ainsi grâce aux vidéos de jeu, aux informations des joueurs, et aux émotions qui ont été extraites par la méthode Émograph (voir chapitre 5) nous avons fait la description formelle la plus précise de chaque zone d'intérêt regardée par les joueurs. De cette étape nous obtenons un fichier individuel par participants. D'un autre côté, nous avons constitué également un fichier unique par zone d'intérêt qui contient les descriptions formelles des scènes de tous les participants qui l'ont regardée. L'ensemble d'entraînement global est alors constitué de l'association de toutes les descriptions des scènes faites par les participants. Les informations sociodémographiques et de personnalités sont incluses dans tous les ensembles créés.

Description numérique d'une nouvelle scène du côté du concepteur

De l'autre côté du miroir, nous nous sommes placés dans la peau d'un concepteur de jeux vidéo (ou de stimuli visuel) et avons défini une description générale de nouvelles scènes en fonction des émotions impliquées dans notre étude, la joie, la colère, la surprise, la peur, le dégoût et la tristesse. Donc dépendamment du choix du concepteur, ces descriptions seront les exemples de test de notre module.

Description numérique générale des scènes

Cette étape consiste à la création d'un fichier qui contient une description générale par scène. C'est à dire pour chacune des variables descriptives, on va lui attribuer la valeur qui a été rencontrée le plus souvent dans les descriptions des participants. Par exemple, dans la scène "Bibliothèque" sur 17 participants qui l'ont regardée 15 ont octroyé la valeur 1 à la variable "Désirabilité" alors dans le fichier de description générale cette scène aura la valeur 1 pour la désirabilité.

Dans notre contexte d'étude, notre objectif est de prédire les émotions d'un individu selon 2 aspects. Le premier par rapport à un objectif émotionnel d'une nouvelle scène défini par le concepteur que nous avons appelé *Objectif 1*. Le second, en face de scènes existantes, c'est-à-dire déjà regardées par d'autres individus, qui est nommé *Objectif 2*. Nous verrons l'utilisation des fichiers créés à travers les approches que nous présenterons dans les sections suivantes. Pour chacune des approches, les modèles disposent comme entrées de vecteurs de variables de 25 dimensions, à savoir les 16 variables du modèle OCC, 4 variables sociodémographiques (le genre, l'âge, l'appartenance ethnique et la catégorie du joueur) et les 5 variables associées aux traits de personnalité, chaque exemple est de la forme $\mathbf{x}_i (x_{i1}, x_{i2}, \dots, x_{i25})$. Les approches ont été développées avec le langage Python et la librairie scikit-learn²², c'est une librairie gratuite (open source) qui fournit des outils efficaces pour les problèmes d'apprentissage machine. Elle est construite sur les librairies NumPy²³, SciPy²⁴ et Matplotlib²⁵.

7.3.2 L'approche par analyse individuelle

Dans cette approche nous basons la prédiction sur l'historique des réactions émotionnelles des individus face aux scènes regardées dans le jeu. C'est à dire que les ensembles d'entraînement sont individuels, c'est à dire qu'il est propre à chaque participant ou scène dépendamment de l'objectif. Cette approche de prédiction par similarité que nous retenons est

²² <http://scikit-learn.org/stable/>

²³ <http://www.numpy.org/>

²⁴ <https://www.scipy.org>

²⁵ <http://matplotlib.org>

très répandue dans plusieurs domaines, en marketing, en physique, en biologie etc. C'est une méthode de prédiction des événements futurs ou inconnus sur la base de vécus similaires ou comparables. En marketing, par exemple, on utilise ce type d'analyse pour la prévision et la reconnaissance des préférences des consommateurs et leurs réactions émotionnelles pour des produits (objets) nouveaux. C'est dans cette approche que nous découvrons l'utilité des fichiers individuels présentés dans la section 7.3.1.

Le nombre n d'exemples est dépendant du nombre de zones d'intérêt regardées par le participant pour l'*objectif 1* et du nombre de participants qui ont regardé la scène pour l'*objectif 2*. Alors pour chaque objectif $D_n = (n \times 25)$. Le nombre d'exemples maximum était de 25 soit le nombre de scènes et de 16 soit le nombre maximal de personnes qui ont regardés une scène. La sortie est l'émotion prédite par le modèle, le nombre de classes maximal est de 6 qui est le nombre d'émotions exprimées par les individus. Chez certains participants ce nombre est plus petit car les participants ne les expriment pas toutes.

Nous avons entraîné et validé nos modèles par la méthode *leave-one-out Cross validation* compte tenu sur nombre réduits d'exemples disponibles sur les données du participant ou de la scène, avant de lui soumettre un exemple de test. Nous présenterons les résultats qui découlent de cette approche dans la section 7.4 de ce chapitre.

7.3.3 L'approche par l'analyse générale

Avec cette approche, nous appuyons notre prédiction sur un ensemble unique d'entraînement qui contient toutes les réponses émotionnelles de tous les participants dans toutes les scènes. Dans ce cadre, nous avons 335 exemples, $D_n = (335 \times 25)$. Nous avons entraîné et validé nos modèles par la méthode de *validation croisée en k blocs (k-fold Cross validation* en anglais). Nous avons trouvé le nombre optimal de blocs avec une recherche par quadrillage. Les résultats de cette approche seront présentés dans la section 7.4 de ce chapitre.

7.4 Les résultats et leurs interprétations

Dans cette section, nous présentons et comparons les résultats en termes de taux de précision enregistrés sur les étapes d'entraînement/validation des modèles prédictifs correspondant aux approches présentées dans la section 7.3. Également nous expliquons le

fonctionnement du module développé à cet effet et montrons des captures d'écran du module intégré dans l'application. Nos résultats sont présentés en termes de précision qui représente le nombre de prédictions correctes divisé par le nombre total de prédictions faites. Formellement elle représente le ratio $VP / (VP + FP)$ avec VP représentant les exemples correctement prédits et FP les exemples considérés comme étant corrects mais qui ne le sont pas en réalité.

7.4.1 La prédiction

7.4.1.1 Approche par analyse individuelle

Pour l'algorithme *kPPV pondéré*, la méthode de validation LOOCV permet pour chaque participant de choisir le nombre optimal de voisins (k) en testant la performance sur plusieurs valeurs de k . Nous faisons varier k entre 2 et 7 car au-delà de ce nombre de voisins les scores de précisions sont les plus bas.

Pour l'algorithme des *forêts d'arbres de décisions*, nous avons trouvé avec la même méthode de validation les paramètres optimaux pour la prédiction du modèle. Les paramètres concernés sont le nombre d'arbres dans la forêt (le nombre d'estimateurs), le nombre de dimensions à considérer pendant le choix de la meilleure division et le nombre minimum d'exemples requis pour être à la feuille d'un nœud, une feuille avec un petit nombre d'exemples rend le modèle plus enclin à capturer le bruit dans les données d'entraînement. Dans tous les cas, le meilleur score était obtenu avec un nombre maximal de dimensions égal à la racine carrée du nombre total de dimensions et le nombre minimum d'exemples était de 1. Seul le nombre d'arbres dans la forêt variait.

Objectif 1 : individu vs nouvelle scène

La performance de notre modèle est évaluée individuellement, c'est à dire que la précision varie d'un participant à l'autre. Les scores de précision vont de 42% à 90%. La précision moyenne sur tous les participants est de 85% avec l'algorithme des *kPPV pondéré*.

La précision va de 75% minimalement à 95% au maximum. La précision moyenne est de 90%, en ce qui concerne l'algorithme des *forêts d'arbres de décisions*.

Le nombre maximal de classes pour l'émotion est de 6, et le nombre minimal de 4.

Objectif 2 : scène vs nouvel individu

Dans cet objectif, la précision varie d'une scène à l'autre avec des scores allant de 27% à 59% de précision. La précision moyenne sur toutes les scènes est de 30% avec l'algorithme des *kPPV pondéré*.

La précision va de 40% minimalement à 80% au maximum. La précision moyenne est de 64%, en ce qui concerne l'algorithme des *forêts d'arbres de décisions*.

Le nombre de classes pour l'émotion est de 6.

7.4.1.2 Approche par analyse générale

Dans cette approche nous avons fait varier le nombre de blocs pour la validation entre 2 et 10 en ce qui concerne le *kPPV pondéré* et entre 5 et 10 pour les *forêts d'arbres de décisions* dans le but de faire l'optimisation des paramètres (ou hyperparamètre) des algorithmes.

Pour l'algorithme *kPPV pondéré*, le nombre de bloc qui a donné le taux de précision le plus élevé est de $K^*=2$ avec le nombre de voisins optimal (k) de, $k^*=4$. Le taux de précision est de 84%, également nous avons fait varier k entre 2 et 7.

Pour l'algorithme des *forêts d'arbres de décisions*, nous avons trouvé un nombre optimal de blocs de $K^*=5$, un nombre optimal d'arbres équivalent à 40, le nombre maximum de dimensions équivalent à la racine carrée du nombre total de dimensions, une profondeur d'arbre maximale, c'est à dire que les nœuds sont étendus jusqu'à ce que toutes les feuilles soient pures ou jusqu'à ce que toutes les feuilles contiennent moins 2 échantillons et un nombre d'exemples minimum de 1. La précision enregistrée est de 98%.

Le tableau 13 fait le récapitulatif des résultats obtenus par nos algorithmes dans les approches.

Tableau 13. Récapitulatif des résultats de l'apprentissage machine.

	Approche individuelle		Approche générale
Méthode de validation	LOOCV		K blocs
Objectifs	<i>Objectif 1</i>	<i>Objectif 2</i>	<i>Général</i>
Algorithmes			
<i>k Plus Proches Voisins pondéré</i>	85%	30%	83% avec $K^* = 2$ et $k^*=4$
<i>Forêts d'arbres de décision</i>	90%	64%	98% avec $K^*=5$

La figure 49 ci-dessous, montre le rapport de classification correspondant au score de 98% enregistré et présenté dans le tableau 13. Sur la figure nous observons le nombre d'exemples dans la colonne "support" et les métriques d'évaluations, la précision, le rappel et le f-score. Rappelons que le rappel peut se traduire formellement par le ratio $VP / (VP + FN)$ avec FN qui représente le nombre de prédictions qui auraient dû être correctement classées. Cette métrique exprime intuitivement la capacité d'un classifieur à trouver tous les exemples positifs. Tandis que le f-score est la moyenne harmonique pondérée de la précision et du rappel. Il s'exprime par la formule : $fscore = \frac{2.Precision.Rappel}{Precision + Rappel}$, dans notre contexte la précision et le rappel sont de pondération équivalente. Le tableau 14 présente quant à lui la matrice de confusion correspondante.

	precision	recall	f1-score	support
Joy	1.00	0.98	0.99	56
Anger	0.96	1.00	0.98	120
Surprise	1.00	1.00	1.00	77
Fear	1.00	0.85	0.92	20
Disgust	1.00	1.00	1.00	20
Sadness	0.98	0.95	0.96	42
avg / total	0.98	0.98	0.98	335

Figure 49. Rapport de classification

Tableau 14. Matrice de confusion

	Joy	Anger	Surprise	Fear	Disgust	Sadness
Joy	55	1	0	0	0	0
Anger	0	120	0	0	0	0
Surprise	0	0	77	0	0	0
Fear	0	2	0	17	0	1
Disgust	0	0	0	0	20	0
Sadness	0	2	0	0	0	40

De ces résultats, l'algorithme des *forêts d'arbres de décision* dont la performance dépasse toutes les autres est celui qui a été intégré à notre module selon l'approche générale, dont nous présentons le principe de fonctionnement dans la section 7.4.2.

7.4.2 Le module de prédiction (Nostradamus)

La prédiction de l'émotion a conduit à l'ajout d'un module supplémentaire à l'application Émograph. Il s'agit du module que nous avons nommé Nostradamus en référence à l'illustre Michel de Nostredame dit Nostradamus qui était un apothicaire, médecin bien qu'il fût expulsé de la faculté de médecine de Montpellier, français de l'époque de la renaissance

surtout très connu pour ses prédictions sur la marche du monde. Cette section décrit dans un premier lieu le principe général de fonctionnement du module et en présente ensuite des captures d'écran.

7.4.2.1 Principe général de fonctionnement

Objectif 1 : Individu vs Nouvelle scène

Dans le contexte de cet objectif, le module fonctionne de la manière suivante :

- L'utilisateur sélectionne un participant dont il souhaite prédire l'émotion.
- L'utilisateur sélectionne l'émotion qu'il souhaite produire dans la nouvelle scène.
- Le module crée l'exemple de test en combinant la description numérique de la nouvelle scène, cette description correspond à l'émotion désirée par l'utilisateur (le concepteur) (voir la section 7.3.1) et les données sociodémographiques et de personnalité du participant sélectionné. Toutes les données sont retrouvées dans la base de données (voir le chapitre 5).
- En arrière-plan, le module fait appel au script de prédiction. Ce script applique l'algorithme des *forêts d'arbres de décision* par l'approche d'analyse générale (voir les sections 7.3.3 et 7.4.1.2).
- Une fois cette étape complétée, le module soumet l'exemple de test au modèle et affiche le résultat de la prédiction et le score de précision calculé à l'étape d'entraînement/validation. Ce score sert d'indicateur de fiabilité de la prédiction pour l'utilisateur.

Objectif 2 : Scène vs Nouvel individu

Pour cet objectif, le module fonctionne de la manière suivante :

- L'utilisateur sélectionne la scène dont pour laquelle il souhaite prédire l'émotion du nouvel individu.
- L'utilisateur remplit un formulaire qui contient les champs pour toutes les informations du nouvel individu (sociodémographiques et traits de personnalité).

- Le module crée l'exemple de test en combinant la description numérique générale de la scène (voir la section 7.3.1), les données sociodémographiques et de personnalité du nouveau participant que l'utilisateur a saisi.
- En arrière-plan, le module fait appel au script de prédiction. Ce script applique l'algorithme des *forêts d'arbres de décision* par l'approche d'analyse générale (voir les sections 7.3.3 et 7.4.1.2).
- Une fois cette étape complétée, le module soumet l'exemple de test au modèle et affiche le résultat de la prédiction et le score de précision calculé à l'étape d'entraînement/validation. Ce score sert d'indicateur de fiabilité de la prédiction pour l'utilisateur.

7.4.2.2 Captures d'écran du module

Prédiction de l'émotion d'un individu par rapport à une nouvelle scène

La figure 50, présente l'interface du module Nostradamus concernant la prédiction de l'émotion d'un individu en face d'une nouvelle scène. Ce module propose aux concepteurs de nouveaux stimuli de sélectionner en premier lieu le participant dont il veut prédire l'émotion et en second lieu l'objectif émotionnel de la nouvelle scène. En retour, le module lui affiche dans un tableau les informations personnelles du participant, l'émotion qu'il désire produire dans la nouvelle scène et l'émotion prédite par le module, avec l'indicateur de fiabilité qui est en fait le taux de précision enregistré pendant l'entraînement/validation.

< Emograph

Nostradamus



Participants

Selectionner un participant.

P21

Objectif emotionnel de la nouvelle scène

Selectionner une émotion.

Surprise

Prédire

Résultat : P21 - Emotion désirée : Surprise

Participant	Émotion désirée	Émotion prédite
AGE : 20 GENRE : FEMALE CATEGORIE : 1 PERSONALITE Extroversion : 96 Emotional stability : 84 Agreeableness : 89 Conscientiousness : 12 Intellect/Imagination : 88 ÉMOTIONS (Nombre d'exemples) anger: 5 surprise: 8 joy: 3 fear: 1 sadness: 2	Surprise	Surprise Taux de précision : 98.77 %

Note : Ce module permet au concepteur de prévoir si l'émotion de l'individu choisi sera en concordance avec l'objectif émotionnel de la future scène.

Figure 50. Interface du module de prédiction pour l'objectif 1.

Prédiction des émotions d'un nouvel individu par rapport à une scène déjà regardée

La figure 51 présente l'interface du module Nostradamus concernant la prédiction de l'émotion d'un nouvel individu en face d'une scène qui a déjà été regardée par d'autres participants. Dans ce cadre, le module propose au concepteur de stimuli en premier lieu de sélectionner la scène dans laquelle ils souhaitent prédire l'émotion du nouveau participant et dans un second lieu de transmettre les informations de celui-ci en remplissant le formulaire affiché à cet effet. Le module affiche les résultats de la requête du concepteur dans un tableau qui contient un rappel des données du participant, la description, l'(es) action(s) et une image de la scène sélectionnée et l'émotion prédite du nouveau participant avec l'indicateur de fiabilité qui est le taux de précision du modèle pendant l'entraînement/validation.

< Emograph

Nostradamus



Zones d'intérêt
 Sélectionnez une scène.
 Bureau entree

Informations du nouveau participant

Age
20

Genre
 Homme Femme

Catégorie
 Débutant Intermediaire Expert

Extraversion
58

Nevrotisme
74

Agreabilite
78

Conscience
7

Intellect
36

Prédire

Résultat : Bureau entree

Participant	Scène	Émotion prédite
Age : 20 Genre : 1 Catégorie : 1 Extraversion : 58 Nevrotisme : 74 Agreabilite : 78 Conscience : 7 Intellect : 36	Description : Situé a la barriere d'entree de l'asile, il contient quelques dessins d'enfants et du mobilier de bureau Action(s) : Explorer le bureau situe avant les grilles d'entree de l'asile. 	Anger Taux de précision : 98.9 %

Note : Ce module permet au concepteur de prévoir l'émotion d'un nouvel individu en face de la scène choisie.

Figure 51. Interface du module de prédiction pour l'objectif 2.

Interprétation des résultats

Les bonnes performances enregistrées par nos algorithmes nous rassurent et confirment la pertinence dans le choix de la modélisation des scènes par le modèle théorique OCC présenté à la section 7.1. À cette modélisation s'ajoute l'importance des informations individuelles, nous avons évalué les performances du même algorithme, avec la même approche, en ne les introduisant pas dans le vecteur d'entrée. Nous avons observé le score chuter à 87%. Nous remarquons que dans notre cadre d'étude l'approche générale fournit de meilleurs résultats que celle individuelle, nous estimons que cela vient du fait du nombre d'exemples réduits dont cette dernière dispose. Toutefois il serait intéressant de disposer de plus d'exemples dans les 2 approches pour effectuer une meilleure comparaison. Aussi les performances enregistrées par

l'algorithme des *forêts d'arbres de décisions* dépassent celles des *kPPV pondéré* peu importe l'approche. Cela fait de cet algorithme le fer de lance sur lequel nous allons baser nos prochaines études. De plus, ces scores raisonnables donnent à l'application la capacité d'ouvrir plus de portes à de nouvelles explorations dans la conception de stimuli multimédia, dans le mesure où les concepteurs peuvent avoir une représentation de l'effet émotionnel du stimuli sur des personnes.

CHAPITRE 8 Les conclusions

Le divertissement multimédia a pris une place quasi indispensable dans le monde moderne actuel, qu'il s'agisse des jeux vidéo, des publicités, des applications mobiles etc. La recherche sur les contenus multimédia tels que les jeux vidéo, les interfaces d'applications ou de sites web, les publicités, les vidéos courtes ou longues vise à tester l'effet d'une nouvelle génération de stimuli visuels sur les consommateurs. Dans notre étude, nous avons porté notre intérêt sur les rapports entre jeux vidéo et les émotions des joueurs. Identifier les scènes dans lesquelles des réponses émotionnelles particulières sont attendues par les concepteurs est crucial.

Nous avons vu que les jeux vidéo créent plusieurs types de réactions et comportements chez les joueurs dépendamment du genre de jeu. Le jeu choisi dans notre cadre était le jeu *Outlast* qui est de genre survival-horror. Ce jeu d'horreur par ses graphiques et son gameplay possède des propriétés visuelles qui engagent des réponses émotionnelles fortes de la part des joueurs. Nous avons vu qu'il est possible de mesurer les émotions humaines par plusieurs techniques, les questionnaires d'auto-évaluation, les mesures physiologiques et notamment les expressions faciales. Les outils d'analyse d'expressions faciales sont de grande utilité dans les tentatives de compréhension des émotions humaines. C'est la technique de mesure utilisée pour notre étude. Nous avons abordé les éléments générateurs d'émotions dans les jeux vidéo et aussi explicité l'impact émotionnel de l'interface graphique sur les joueurs. De ce fait nous avons choisi une séquence de jeu et en avons fait une division en étapes en fonction des éléments graphiques et du gameplay. Pour comprendre les interactions entre les joueurs et l'interface nous avons intégré à notre étude la technique d'oculométrie. Nous voyons que les émotions humaines représentent un domaine de recherche complexe et qu'il en existe une grande variété. Dans notre étude nous avons choisi de faire notre analyse suivant l'approche catégorielle en nous penchant sur les émotions primaires telles que catégorisées par (Ekman & Friesen, 1971). L'enjeu de notre recherche était de combiner les procédés de reconnaissance des émotions humaines et l'oculométrie dans le domaine du jeu vidéo. Le but étant d'extraire d'une séquence de jeu, les émotions dominantes ainsi que de mettre en avant les éléments visuels qui les déclenchent, sous forme d'une courbe qui représente les émotions dominantes d'une étape à l'autre du jeu. Il s'agit

de trouver une méthode qui détermine les éléments principaux dont résultent les expressions faciales les plus fortes. Pour développer notre méthode nous avons conduit une expérience sur 21 individus. Nous avons éprouvé des problèmes techniques pendant l'expérience qui nous ont obligés à retirer 4 individus du fait d'une mauvaise qualité des données collectées. À partir de cette expérience, nous proposons une méthode d'analyse émotionnelle.

La méthode proposée dans ce mémoire est une contribution originale. En effet son utilisation permet de cerner avec un degré élevé de précision les émotions fortes créées chez un individu en train de jouer, comparativement aux études précédentes dans le domaine qui se limitent à des aspects plus généraux. De plus l'intégration de l'oculométrie dans le processus d'analyse fournit un autre niveau de précision quant à la conception graphique des stimuli.

En effet, l'intégration des techniques de corrections des données brutes par valeurs de références et de fixation de seuils d'amplitudes relatifs rend cette méthode adaptative aux différences entre les personnes. Le découpage temporel des scènes quant à lui amène de la précision dans l'analyse. Nous avons pu observer la qualité de la précision de notre méthode au millième de secondes près à partir des exemples présentés dans le chapitre 6 aux figures 25,26,27,28,29 et 30. Le fichier créé relatif au participant, présenté sur la figure 25 fait la correspondance avec la capture réalisée sur la figure 26. Sur cet exemple montré en figure 26, nous observons de la surprise sur le visage du participant, avec la trajectoire et le point de fixation de son regard. Nous voyons la trace de son regard sur les éléments graphiques impliqués dans la production de la surprise, notamment l'agent de sécurité. De plus, nous avons identifié avec précision à quel instant de la scène apparaît l'expression faciale la plus forte. Toutes les captures de jeu suivantes ont été réalisées à partir des informations obtenues par la méthode. Un autre avantage de cette méthode est sa versatilité. Elle peut être appliquée à tout type de stimuli multimédia, que ça soit une pièce de théâtre, des vidéos courtes ou longues. Son application sur des données provenant d'analyse d'expressions faciales et de suivi de regard entraîne la construction d'une base de données avec les fichiers résultants.

Un autre avantage de notre méthode est la généralisation qu'elle propose. Pour généraliser l'utilisation de cette méthode, nous avons développé une application qui interagit avec la base de données constituée par les résultats extraits de la méthode. L'application développée à cet effet, inclut un grand nombre de fonctionnalités pour les utilisateurs. Ces

fonctionnalités sont décrites et présentées sur les figures de la section 6.2. Le fait de la généralisation permet d'étendre l'application de cette méthode à des stimuli tels que les pièces de théâtres, les publicités etc.

Au cours de notre recherche, la problématique de la prédiction des émotions a émergé. De ce fait nous avons tenté de répondre à ce problème par l'utilisation et la formalisation d'un modèle d'évaluation cognitive, en l'occurrence le modèle OCC (Ortony et al., 1988). Le choix de ce modèle est venu du fait que la plupart des travaux sur les systèmes intelligents qui visaient à simuler le processus émotionnel humain s'en sont inspirés. Aussi ce modèle permet de répondre au problème de représentation de l'information émotionnelle des scènes de jeu, en le combinant aux émotions extraites par la méthode de détection/sélection d'émotions proposée dans ce mémoire. Pour prédire les émotions, il est nécessaire d'avoir recours aux algorithmes d'apprentissage machine, dans notre travail nous avons utilisé le *kPPV pondéré* et les *forêts d'arbres de décisions*, nous avons présentés ces algorithmes dans la section 7.2.3 de ce mémoire. Ces algorithmes nécessitent un ensemble d'apprentissage que nous avons construit par les descriptions numériques des scènes auxquelles nous avons rajoutées les informations sociodémographiques et les traits de personnalité des individus. Dans ce mémoire nous présentons 2 approches, la première qui se base sur une analyse individuelle et la seconde qui utilise l'historique des scènes regardées et des émotions exprimées par les joueurs. Au sein de ces 2 approches, nous avons identifié 2 objectifs, le premier était de prédire l'émotion d'un individu par rapport à l'objectif émotionnel d'une nouvelle scène de jeu et le second était de prédire l'émotion d'un nouvel individu par rapport à une scène de jeu qui a déjà été regardée par des individus auparavant. Les performances de l'approche générale enregistrées sont satisfaisantes, atteignent un score de 98% et dépassent celles de l'approche individuelle qui sont de 90% pour l'*objectif 1* et de 64% pour l'*objectif 2*, tous les scores sont reportés dans le tableau 13. Ce score élevé témoigne de la fiabilité sur laquelle peuvent se reposer les concepteurs en utilisant le module développé. Néanmoins, il nous faut trouver un moyen d'évaluation de notre outil en situation réelle, soit par des interviews avec les joueurs pendant lesquelles nous discutons des prédictions de l'outil ou par un questionnaire d'auto évaluation des émotions que nous comparerons avec les prédictions du module.

Travaux futurs

Dans le but d'aller en profondeur dans la compréhension des interactions émotionnelles complexes entre un jeu vidéo et un joueur, nous pourrions améliorer cette étude par le traitement de données collectées par d'autres senseurs et leur intégration à cette méthode. Cette amélioration impliquerait d'investiguer sur les niveaux de stress, d'anxiété et de tension et d'également se pencher sur les émotions plus complexes ressenties dans les jeux vidéo comme le triomphe, le désespoir, l'espoir, la détresse, la fierté afin de faire la différence parmi les émotions primaires distinguées et affermir la personnalisation de l'analyse. En restant dans la même lignée de l'intégration des senseurs, un autre axe à exploiter serait la combinaison de ceux-ci. En effet certaines personnes peuvent dissimuler leurs émotions. Ce qui veut dire que l'analyse des expressions faciales peut ne pas refléter le ressenti réel de l'individu. Dans ce sens il serait innovateur de procéder à une combinaison de l'analyse des autres capteurs physiologiques. Notamment les signaux EEG. Car l'étude de (Benlamine et al., 2016) reportée dans la section 2.3 a produit d'excellents résultats dans la reconnaissance d'expressions faciales à partir de EEG. Alors en combinant ces études il serait possible d'évaluer si un individu était en train de dissimuler ses émotions.

Aussi presque tous les senseurs utilisés dans la littérature et dans cette étude nécessitent un contact direct avec le sujet et une certaine préparation de l'équipement et du participant pour fonctionner correctement (ex. : l'application du gel conducteur sur les électrodes pour enregistrer le rythme cardiaque). Il serait alors intéressant d'intégrer certains des nouveaux senseurs embarqués qui font progressivement leur apparition sur le marché. Ceux-ci sont conçus de telles sortes qu'ils peuvent être dissimulés dans l'environnement du participant. On pense, entre autres, aux capteurs de tension placés dans les chaises comme le montre cette étude (D'Mello & Graesser, 2010), aux senseurs d'activité électrodermale sur la souris. Certains appareils, tels les caméras thermiques, permettent des mesures sans contact du rythme cardiaque via le changement de température du visage. On voit même l'apparition de fonctionnalité de mesure intégrée à même l'objet interactif. Par exemple, une équipe de chercheurs du MIT²⁶ a créé une application Android permettant de mesurer le rythme cardiaque à l'aide de la camera

²⁶ Massachusetts Institute of Technology

frontale d'un téléphone portable²⁷. Cependant les capteurs cités sont moins précis et produisent un enregistrement plus bruité.

En ce qui concerne le module de prédiction, une piste d'amélioration serait de baser la prédiction sur une combinaison des 2 approches présentées dans ce mémoire. De plus, il serait avantageux de reconduire une expérience en tenant compte de l'intérêt de prédiction comme donnée initiale, dans ce cas il faudrait compléter l'expérience présentée dans ce mémoire par une étape de description des scènes par les participants. Cette étape consisterait à réaliser des interviews avec les joueurs des après la séquence d'analyse et d'extraction des émotions par Émograph. Cette étape pouvant être effectuée dans un délai très court elle n'engendrerait que quelques minutes supplémentaires dans le protocole expérimental de base. Aussi la nouvelle expérience apporterait une plus grande quantité de données qui viendrait élargir l'ensemble d'apprentissage et ainsi améliorer la comparaison des approches car il serait intéressant d'observer le comportement de nos algorithmes avec les descriptions faites par les joueurs eux-mêmes. De plus, une nouvelle expérience serait un test viable de notre méthode et application en condition réelle d'utilisation. Une piste supplémentaire serait le test de nos méthodes et de l'application sur des séquences courtes de jeux vidéo dont nous détenons le code source. Cela pour introduire la notion d'adaptation du jeu en fonction des émotions du joueur à notre approche comme dans cette étude de (Ghali, Ouellet, & Frasson, 2016) dans laquelle l'objectif est de fournir de l'aide automatiquement à un étudiant en fonction de son état émotionnel. Posséder le code source du jeu implique que nous pourrions rajouter les caractéristiques des éléments visuels de jeu telles que la couleur, la texture, les dimensions, aux descriptions formelles des scènes provenant de l'évaluation cognitive. Cette approche permettrait de d'atteindre non seulement la conception de scènes mais également celle des éléments à l'intérieur des scènes de jeu. Des séquences plus courtes permettront d'être plus précis sur l'étude des éléments graphiques.

²⁷ www.cardiio.com

Bibliographie

Adams, E. (2009). *Fundamentals of Game Design*: New Riders Publishing.

Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology*, 78(4), 772.

Arnold, M. B. (1960). *Emotion and personality: Neurological and physiological aspects* (Vol. 2): Columbia University Press.

Bailey, K., & West, R. (2013). The effects of an action video game on visual and affective information processing. *Brain research*, 1504, 35-46.

Bates, J., Loyall, A. B., & Reilly, W. S. (1992). *An architecture for action, emotion, and social behavior*. Paper presented at the European Workshop on Modelling Autonomous Agents in a Multi-Agent World.

Benlamine, M., Chaouachi, M., Frasson, C., & Dufresne, A. (2016). Physiology-based Recognition of Facial Micro-expressions using EEG and Identification of the Relevant Sensors by Emotion
. *Proceedings of the 3rd International Conference on Physiological Computing Systems - Volume 1: PhyCS*.

Bishop, C. M. (2006). Pattern recognition. *Machine Learning*, 128.

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion*, 1(3), 300.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *Journal of experimental psychology: Learning, Memory, and Cognition*, 18(2), 379.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*: CRC press.

D'Mello, S., & Graesser, A. (2010). *Mining bodily patterns of affective experience during learning*. Paper presented at the Educational Data Mining 2010.

Darwin, C. (1872). The expression of emotion in animals and man. *London: Methuen.*(1877), *A biographical sketch of an infant. Mind*, 2, 285-294.

De la Torre, F., & Cohn, J. F. (2011). Facial expression analysis *Visual analysis of humans* (pp. 377-409): Springer.

Devijver, P. A., & Kittler, J. (1982). *Pattern recognition: A statistical approach*: Prentice hall.

Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, 6(3-4), 169-200.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial cues*: Englewood Cliffs, NJ: Prentice Hall.

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: Investigatoris Guide*: Consulting Psychologists Press.

Ekman, P., & O'Sullivan, M. (1991). Facial expression: Methods, means, and moues.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotions. *J. Cole*, 207-283.

Elliott, C., & Brzezinski, J. (1998). Autonomous agents as synthetic characters. *AI magazine*, 19(2), 13.

Elliott, C. D. (1992). The affective reasoner: A process model of emotions in a multi-agent system.

Gauducheau, N. (2009). *Mesurer les émotions de l'utilisateur: quels fondements pour une démarche d'évaluation des systèmes interactifs?* Paper presented at the Proceedings of the 21st International Conference on Association Francophone d'Interaction Homme-Machine.

Ghali, R., Ouellet, S., & Frasson, C. (2016). LewiSpace: An Exploratory Study with a Machine Learning Model in an Educational Game. *Journal of Education and Training Studies*, 4(1), 192-201.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological assessment*, 4(1), 26.

Gratch, J. (2000). *Emile: Marshalling passions in training and education*. Paper presented at the Proceedings of the fourth international conference on Autonomous agents.

Gratch, J., & Marsella, S. (2003). Fight the way you train: the role and limits of emotions in training for combat. *Brown J. World Aff.*, 10, 63.

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5(4), 269-306.

Gratch, J., Marsella, S., & Petta, P. (2009). Modeling the cognitive antecedents and consequences of emotion. *Cognitive Systems Research*, 10(1), 1-5.

Hazlett, R. L., & Benedek, J. (2007). Measuring emotional valence to understand the user's experience of software. *International Journal of Human-Computer Studies*, 65(4), 306-314.

Hess, U. (2001). The experience of emotion: situational influences on the elicitation and experience of emotions. *Emotions, qualia, and consciousness*, 386-396.

Hess, U., Blairy, S., & Kleck, R. E. (2000). The Influence of Facial Emotion Displays, Gender, and Ethnicity on Judgments of Dominance and Affiliation. *Journal of Nonverbal Behavior*, 24(4), 265-283. doi:10.1023/a:1006623213355

Ho, T. K. (1995). *Random decision forests*. Paper presented at the Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on.

iMotions. (2013). Attention Tool FACET Module Guide. Retrieved from https://imotions.com/wp-content/uploads/2013/08/130806_FACET_Guide.pdf

iMotions. (2016). Facial Expressions Analysis - POCKET GUIDE.

Iqbal, F. (2015). Understanding User Interaction in a Video Game by using Eye Tracking and Facial Expressions Analysis.

Isbister, K., Höök, K., Laaksolahti, J., & Sharp, M. (2007). The sensual evaluation instrument: Developing a trans-cultural self-report measure of affect. *International Journal of Human-Computer Studies*, 65(4), 315-328.

Jolliffe, I. (2002). *Principal component analysis*: Wiley Online Library.

Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychol Bull*, 100(1), 78-100.

Kwon, O.-W., Chan, K., Hao, J., & Lee, T.-W. (2003). *Emotion recognition by speech signals*. Paper presented at the INTERSPEECH.

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, 61(1), 132.

Lazarus, R. S. (1991). Emotion and adaptation. 1991. *Cité en*, 9.

Lazzaro, N. (2004). Why we play games: Four keys to more emotion without story.

Lee, C. M., & Narayanan, S. S. (2005). Toward detecting emotions in spoken dialogs. *IEEE transactions on speech and audio processing*, 13(2), 293-303.

Lewis, M., Haviland-Jones, J. M., & Barrett, L. F. (2010). *Handbook of emotions*: Guilford Press.

Littlewort, G., Bartlett, M. S., Fasel, I., Susskind, J., & Movellan, J. (2006). Dynamics of facial expression extracted automatically from video. *Image and Vision Computing*, 24(6), 615-625.

Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). *The computer expression recognition toolbox (CERT)*. Paper presented at the Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on.

Loyall, B., Bryan, A., & Bates, L. J. (1991). Hap a reactive, adaptive architecture for agents.

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. Paper presented at the Proceedings of the fifth Berkeley symposium on mathematical statistics and probability.

Mandryk, R. L., Atkins, M. S., & Inkpen, K. M. (2006). *A continuous and objective evaluation of emotional experience with interactive play environments*. Paper presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems.

Matsumoto, D. (1992). More evidence for the universality of a contempt expression. *Motivation and Emotion*, 16(4), 363-368.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion*, 23(2), 209-237.

McLachlan, G. J., & Basford, K. E. (1988). *Mixture models*: New York: Marcel Dekker.

Mirza-Babaei, P., & McAllister, G. (2011). *Using physiological measures in conjunction with other usability approaches for better understanding of the player's gameplay experiences*. Paper presented at the Game Research Methods Seminar.

Mitchell, T. M. (1997). *Machine learning*. 1997. Burr Ridge, IL: McGraw Hill, 45, 37.

Newzoo. (2014). TOP 100 COUNTRIES REPRESENT 99.8% OF \$81.5BN GLOBAL GAMES MARKET. Retrieved from <https://newzoo.com/insights/articles/top-100-countries-represent-99-6-81-5bn-global-games-market/>

Nielsen, J., & Pernice, K. (2010). *Eyetracking web usability*: New Riders.

Ortony, A., Clore, A., & Collins, G. (1988). *The cognitive structure of emotions*: Cambridge University Press.

Rajae-Joordens, R. J. (2008). Measuring experiences in gaming and TV applications *Probing Experience* (pp. 77-90): Springer.

Reilly, W. S., & Bates, J. (1992). *Building emotional agents*.

Rollings, A., & Morris, D. (2003). *Game architecture and design: a new edition*.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178. doi:10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, 110(1), 145.

- Russell, S., Graetz, M., & Witaenem, W. (1962). Spacewar. *Computer software*.
- Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological bulletin*, 99(2), 143.
- Scherer, K. R. (1997). The role of culture in emotion-antecedent appraisal. *Journal of Personality and Social Psychology*, 73(5), 902.
- Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, research*, 92, 120.
- Scherer, K. R. (2004). HUMAINE deliverable D3c: Preliminary plans for exemplars: Theory.
- Scherer, K. R., Wranik, T., Sangsue, J., Tran, V., & Scherer, U. (2004). Emotions in everyday life: probability of occurrence, risk factors, appraisal and reaction patterns. *Social Science Information*, 43(4), 499-570.
- Schwebel, D. C., & Suls, J. (1999). Cardiovascular reactivity and neuroticism: Results from a laboratory and controlled ambulatory stress protocol. *Journal of Personality*, 67(1), 67-92.
- Shaker, N., Yannakakis, G. N., & Togelius, J. (2010). *Towards Automatic Personalized Content Generation for Platform Games*. Paper presented at the AIIDE.
- Shanthi, S., & Ramani, R. G. (2012). *Classification of Seating Position Specific Patterns in Road Traffic Accident Data through Data Mining Techniques*. Paper presented at the Second International Conference on Computer Applications, ICCA.
- Slater, S. I., Moreton, R., Buckley, K., & Bridges, A. (2008). A Review of Agent Emotion Architectures. *Eludamos. Journal for Computer Game Culture*, 2(2), 203-214.
- Sloman, A., & Poli, R. (1995). *SIM_AGENT: A toolkit for exploring agent designs*. Paper presented at the International Workshop on Agent Theories, Architectures, and Languages.

SMI. (2011). Case Study Eye Tracking: Games User Research. Retrieved from http://www.smivision.com/fileadmin/user_upload/downloads/case_studies/cs_smi_games.pdf

Stemmler, G. (2004). Physiological processes during emotion. *The regulation of emotion*, 33-70.

SuperData. (2015). SuperData/Google event: Gamers are dead. Long live gaming. Retrieved from <https://www.superdataresearch.com/gamers-are-dead/>

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1): MIT press Cambridge.

Tan, E. S. (1996). Emotion and the Structure of Narrative film: Film as an Emotion Machine, prijev. B. Fasting: Mahwah, NJ: Erlbaum.

Tobii. (2015). User experience with eye tracking for Guerrilla Games' Killzone 3. Retrieved from http://acuity-ets.com/downloads/Tobii_CustomerCase_Valsplat_GuerrillaGames.pdf

Trabelsi, A., & Frasson, C. (2010). *The Emotional Machine: A Machine Learning Approach to Online Prediction of User's Emotion and Intensity*. Paper presented at the 2010 10th IEEE International Conference on Advanced Learning Technologies.

Vapnik, V. N. (2000). The nature of statistical learning theory, ser. Statistics for engineering and information science. *New York: Springer, 21*, 1003-1008.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4), 279-292.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin*, 98(2), 219.

Wundt, W. M. (1907). *Outlines of psychology*: W. Engelmann.

Annexe A Formulaire de consentement

Titre de la recherche : Étude des réactions émotionnelles et cérébrales dans un environnement d'apprentissage. Brain based Tutoring Systems. NSERC

Chercheurs : Asma Ben Kheder, Ramla Ghali, Sahbi Benlamine, étudiants PhD, DIRO, Samira Bouslimi, Mahmoud Sfaia,, Sebastien Ouellet, étudiants MSc, et autre étudiants à venir DIRO, Directeur du Laboratoire : Claude Frasson, Professeur titulaire, DIRO, Université de Montreal

RENSEIGNEMENTS AUX PARTICIPANTS

Objectifs de la recherche.

Ce projet de recherche vise à étudier les réactions émotionnelles et cérébrales dans un environnement d'apprentissage. Nous allons ainsi enregistrer l'activité cérébrale du participant au fur et à mesure qu'il est en train d'expérimenter notre système d'apprentissage.

Participation à la recherche

Votre participation à cette recherche consiste tout d'abord à passer un test de niveau de connaissance initial. Ensuite vous expérimentez un système d'apprentissage en remplissant un questionnaire de motivation après chaque étape accomplie. Enfin vous répondrez à un test final des connaissances. Ceci implique d'enregistrer votre activité cérébrale (EEG) et votre suivi visuel (oculométrie de suivi du regard). L'expérience se déroulera dans un local du département d'informatique de l'Université de Montréal et durera ~1h. Vous bénéficierez d'un montant de 20\$ après votre participation à cette expérience.

Confidentialité

Les renseignements que vous nous donnerez demeureront confidentiels. Chaque participant de l'étude se verra attribuer un numéro d'identification. Aucune information permettant de vous identifier d'une façon ou d'une autre ne sera publiée. Ces données seront conservées durant sept ans, conformément à la politique habituelle de l'Université de Montréal. Après ce délai, elles seront totalement détruites.

Avantages et inconvénients

L'utilisation des capteurs physiologiques est indolore et sans risque. En participant à cette recherche, vous pourrez contribuer à l'avancement des connaissances sur l'apprentissage par les jeux et ses potentiels. Vous bénéficierez d'un montant de 20\$ après votre participation à l'expérience.

Droit de retrait

Votre participation est entièrement volontaire. Vous êtes libre de vous retirer en tout temps par avis verbal, sans préjudice et sans devoir justifier votre décision. Si vous décidez de vous retirer de la recherche, vous pouvez communiquer avec l'assistant de recherche, au numéro de téléphone indiqué à la page suivante du formulaire. Si vous décidez de vous retirer de la recherche, veuillez communiquer avec les chercheurs indiqués À votre demande, tous les renseignements qui vous concernent pourront aussi être détruits. Cependant, après le déclenchement du processus de publication, il sera impossible de détruire les analyses et les résultats portant sur vos données.

CONSENTEMENT

Je déclare avoir pris connaissance des informations ci-dessus, avoir obtenu les réponses à mes questions sur ma participation à la recherche et comprendre le but, la nature, les avantages, les risques et les inconvénients de cette recherche.

Après réflexion et un délai raisonnable, je consens librement à prendre part à cette recherche. Je sais que je peux me retirer en tout temps sans préjudice et sans devoir justifier ma décision.

Signature :

Date : / /2016

Nom : Prénom :

Environnement expérimental : E1 : E2 : E3 : E4 :

Je déclare avoir expliqué le but, la nature, les avantages, les risques et les inconvénients de l'étude et avoir répondu au meilleur de ma connaissance aux questions posées.

Signature du chercheur Date : (ou de son / /2016
représentant)

Nom : Prénom :

Pour toute question relative à la recherche, ou pour vous retirer de la recherche, vous pouvez communiquer avec René Lanciné Doumbouya à l'adresse courriel suivante : rene.lancine.doumbouya@umontreal.ca ou Sahbi Benlamine à l'adresse courriel suivante : benlaminem@iro.umontreal.ca

Toute plainte relative à votre participation à cette recherche peut être adressée à l'ombudsman de l'Université de Montréal, au numéro de téléphone (514) 343-2100 ou à l'adresse courriel ombudsman@umontreal.ca. (L'ombudsman accepte les appels à frais virés).

Un exemplaire du formulaire de consentement signé doit être donné au participant.

Jeux vidéo cités

Format : Titre, Développeur, Année, Publication

Ace Attorney, Capcom, 2001, Capcom

Call of Duty Modern Warfare 2, Infinity Ward, 2009, Activision

Doom, id Software, 1993, GT Interactive

Fahrenheit, Quantic Dream Aspyr, 2005, Atari, Inc

FIFA 17, EA Canada, 2016, EA Sports

Fight Night Champion, EA Canada, 2011, Electronic Arts

Final Fantasy IV, Square, 1991, Square

Final Fantasy Tactics, Square, 1998, Sony Computer Entertainment

F1 2013, Codemasters, 2013, Codemasters

Grand Theft Auto, DMA Design, 1997, BMG Interactive

Grand Theft Auto III, DMA Design, 2001, Rockstar Games

Grand Theft Auto III : Vice City, Rockstar North, 2002, Rockstar Games

Grand Theft Auto : San Andreas, Rockstar North, 2004, Rockstar Games

League of Legends, Riot Games, 2009, Riot Games

The Longest Journey, Funcom, 2000, Empire Interactive, Tri Synergy

The Lord of the Rings Online, Turbine, Inc., 2007, Turbine, Inc.

Metal Gear, Konami, 1987, Ultra Games

Metroid Prime, Nintendo, 2002, Nintendo

Naruto Shippūden Ultimate Ninja Storm 4, CyberConnect2, 2016, Bandai Namco Games

Outlast, Red Barrels, 2013, Red Barrels

Uncharted: Drake's Fortune, Naughty Dog, 2007, Sony

Uncharted 2 : Among Thieves, Naughty Dog, 2009, Sony

Uncharted 3: Drake's Deception , Naughty Dog, 2011, Sony

SimCity, Maxis, 2013, Electronic Arts

SimLife, Maxis Software, 1992, Maxis Software

Spacewar, Stewart Russel (and others), 1962

Spy Hunter, Bally Midway, 1983, Bally Midway

Street fighter, Capcom, 1987, Capcom

Tetris, Spectrum HoloByte, 1984, Spectrum HoloByte

Tom Clancy's End War, Ubisoft Shanghai, 2009, Ubisoft

