



HAL
open science

Modélisation et évaluation de la fidélité d'un clone virtuel

Mehdi Boukhris

► **To cite this version:**

Mehdi Boukhris. Modélisation et évaluation de la fidélité d'un clone virtuel. Synthèse d'image et réalité virtuelle [cs.GR]. Université Paris Saclay (COMUE), 2015. Français. NNT : 2015SACLS177 . tel-01424127

HAL Id: tel-01424127

<https://theses.hal.science/tel-01424127>

Submitted on 2 Jan 2017

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

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

NNT : 2015SACLS177

THESE DE DOCTORAT
DE L'UNIVERSITE PARIS-SACLAY,
Préparée à l'université Paris-Sud

ÉCOLE DOCTORALE N° 580
STIC: Sciences et technologies de l'information et de la communication

Spécialité de doctorat : Informatique

Par

Mr BOUKHRIS Mehdi

Titre de la thèse
Modélisation et Evaluation de la Fidélité d'un Clone Virtuel

Thèse présentée et soutenue à Orsay, le 4/12/2015 :

Composition du Jury :

Mr, Jacquemin, Christian	Professeur, Université Paris Sud,
Mme, Lisetti, Christine	Professeur associé, Florida University,
Mme, André, Elisabeth	Professeur, Universität Augsburg,
Mme, Pesty, Sylvie	Professeur, Université de Grenoble 2,
Mr, Martin, Jean-Claude	Professeur, Université Paris Sud,

Président du jury
Rapporteur
Rapporteur
Examinatrice
Directeur de thèse



Résumé

L'identification des visages est primordiale lors de nos interactions sociales. Ainsi, notre comportement change suite à l'identification de la personne avec laquelle nous interagissons. De plus, les travaux en psychologie et en neurosciences ont observé que le traitement cognitif face à un visage familier diffère de celui que nous avons face à un visage inconnu. Les dernières techniques de rendu 3D et les dernières avancées des scans 3D ont permis la création de visages virtuels photo-réalistes modélisant des personnes réelles existantes. La tendance actuelle pour modéliser des humains virtuels est de se baser sur des techniques d'acquisition de données réelles (issues de scans et de sessions de capture de mouvement). Les recherches et applications en humains virtuels ont connu un intérêt croissant pour ces clones virtuels (des agents ayant un aspect familier ou du moins reconnaissable). Les clones virtuels sont donc de plus en plus répandus dans des interfaces homme-machine et dans l'industrie audio-visuelle.

L'étude de la perception et de l'interaction avec des clones virtuels est donc nécessaire pour appréhender la conception et l'évaluation de cette technologie. En effet, très peu d'études se sont penchées sur l'évaluation de la fidélité de ces clones virtuels. L'objectif de cette thèse consiste à explorer cet axe de recherche en examinant le processus de perception de la fidélité d'un visage virtuel, clone d'une personne réelle (que l'on connaît ou non).

Nos travaux répondent à plusieurs questions de recherche : Quels sont les éléments qui nous permettent d'évaluer la ressemblance du clone virtuel avec son référent ? Parmi les multiples possibilités de techniques de rendu, d'animation et d'acquisition de données qu'offre l'informatique graphique, quelle est la meilleure combinaison pour assurer le plus haut degré de fidélité perçue ?

L'apparence visuelle n'est cependant qu'une des composantes qui interviennent dans la reconnaissance de personnes familières. Les autres composantes comprennent ainsi l'expressivité mais aussi le traitement des connaissances que nous avons sur cette personne (par exemple sa manière particulière d'évaluer une situation émotionnelle et de l'exprimer via son visage).

Nos contributions apportent des éléments de réponse à ces questions à plusieurs niveaux. Nous avons défini un cadre conceptuel identifiant les principaux concepts pertinents pour l'étude de la fidélité d'un visage virtuel. Nous avons aussi étudié l'aspect visuel de la fidélité à travers l'exploration de différentes techniques de rendu. Nous avons étudié dans une autre étape l'impact de la familiarité dans le jugement de la fidélité. Finalement, nous avons proposé les premières briques d'un modèle informatique individuel basé sur une approche cognitive des émotions qui permettrait de guider l'animation expressive du clone virtuel. Ces travaux de thèse ouvrent des perspectives pour la conception et l'amélioration de clones virtuels, mais aussi plus généralement des interfaces homme-machine basées sur des agents expressifs.

Abstract

Face identification plays a crucial role in our daily social interactions. Indeed, our behavior changes according to the identification of the person with whom we interact. Moreover, several studies in Psychology and Neurosciences have observed that our cognitive processing of familiar faces is different from the cognitive processing of unfamiliar faces.

Creating photorealistic and animated human-like face of a real person is now possible thanks to recent advances in Computer Graphics and 3D scan systems. Recent rendering techniques are challenging our ability to distinguish between computer generated faces and real human faces. Besides, the current trend to model virtual humans is to involve real data collected using scans and motion capture systems.

Research and applications in virtual humans have experienced a growing interest in so-called virtual clones (agents with a familiar aspect or at least recognizable). Virtual clones are therefore increasingly used in human-machine interfaces and in the audiovisual industry.

Studies about the perception and interaction with virtual clones are therefore required to better understand how we should design and evaluate this kind of technology. Indeed, very few studies have tried to evaluate virtual clones' fidelity with respect to the original human (hereafter called "the referent").

The main goal of this thesis is to explore this line of research. Our work rises several research questions : What are the features of the virtual clone that enable us to evaluate the resemblance between a virtual clone and its referent ? Among several possibilities of rendering, animation and data acquisition techniques offered by Computer Graphics, what is the best combination of techniques to ensure the highest level of perceived fidelity ?

However, visual appearance is not the only component that is involved in recognizing familiar people. The other components include facial expressiveness but also the possible knowledge that we have about the referent (e.g. his particular way of assessing an emotional situation and expressing it through his face).

Our contributions provide answers to these questions at several levels. We define a conceptual framework identifying the key concepts which are relevant for the study of the fidelity of a virtual face. We explore different rendering techniques. We describe an experimental study about the impact of familiarity in the judgment of fidelity. Finally, we propose a preliminary individual computational model based on a cognitive approach of emotions that could drive the animation of the virtual clone.

This work opens avenues for the design and improvement of virtual clones, and more generally for the human-machine interfaces based on expressive virtual agents.

Acknowledgements

Un travail de thèse est à mon sens assimilable à une escalade d'un sommet qui semble bien lointain. Pas à pas, tu franchis les différents obstacles de cette montagne, accompagnée parfois par le doute et la frustration, mais heureusement aussi par l'encouragement, de la confiance et l'aide des gens qui t'entourent qui te tire vers le haut et t'aide à te surpasser. Quand tu te retrouves, finalement à destination au sommet du pic savourant la beauté du paysage, tu te rends compte que c'est grâce à ce travail d'équipe que tu as la chance de jouir d'un moment aussi rare.

C'est pour cela que par ces quelques mots je tenais à adresser mes sincères remerciements à toutes les personnes qui de près ou de loin ont cru en moi et qui dans les moments les plus durs m'ont rappelé mes premières motivations et ma curiosité pour le domaine de la recherche.

Mon premier remerciement s'adresse à l'ensemble des membres qui m'ont fait l'honneur de siéger au sein de mon jury. Merci aux rapporteurs, Pr Elisabeth André et Pr Christine Lisetti pour leurs retours constructifs et leurs commentaires positifs et encourageants.

Je voulais également adresser ma reconnaissance à mon directeur de thèse Jean-Claude Martin qui m'a patiemment supervisé. Merci pour toutes les discussions scientifiques que nous avons partagées.

Merci à tous les membres de mon équipe, pour tous ces échanges sur diverses problématiques de recherche. Mais aussi pour tous nos échanges non professionnels qui étaient un peu moins sérieux, mais tout aussi intéressants. J'adresse également des remerciements bien particuliers à mes camarades de bureau pour leur soutien pour tous les débats qu'on a pu avoir et pour des moments de détente mémorables.

Merci à tous mes amis qui soient à distance dans ma chère Tunisie ou aux personnes parisiennes dont j'ai eu la chance de faire la connaissance plus récemment. Une pensée particulière à certaines connaissances de la maison de la Tunisie auprès desquelles j'ai vécu des moments inoubliables.

Pour terminer, je voulais particulièrement remercier tous les membres de ma famille pour leur soutien quotidien malgré la distance qui nous séparait. Un grand merci à mes parents pour être tout simplement là, un grand merci à ma sœur Imen pour son aide incessante et finalement un grand merci à ma grand-mère pour sa bonté infinie.

Table des matières

1	Introduction générale	1
1.1	Contexte général	2
1.2	Motivation	7
1.3	Problématiques et objectifs scientifiques	7
1.4	Démarche incrémentale	9
1.5	Organisation du mémoire	11
I	Paradigme conceptuel pour l'étude de la fidélité	13
2	Vers une définition de la fidélité pour des clones virtuels	15
2.1	Pourquoi s'intéresser à la fidélité?	16
2.2	Les différentes définitions de la fidélité	17
2.3	Les différentes composantes de la fidélité d'un visage	18
2.4	Différences interindividuelles lors de la perception et de l'identification d'un visage	23
2.5	Limites de l'existant & conclusions	24
3	Cadre Conceptuel pour l'étude de la fidélité	25
3.1	Introduction	26
3.2	Principales dimensions de la fidélité	26
3.3	Qu'est ce qu'on évalue?	27
3.4	Les entités impliquées lors du jugement de la fidélité	30
3.5	Quelles sont les variables qui influencent le jugement de la fidélité?	32
3.6	Synthèse	33
II	Capture, rendu et animation d'un clone virtuel	35
4	État de l'art : Conception d'humains virtuels et techniques de numérisation	37
4.1	Introduction	38

4.2	Création du maillage	38
4.3	Rendu	45
4.4	Animation	49
4.5	Limites de la modélisation d'humains virtuels	54
5	Modélisation, rendu et animation d'un clone virtuel	57
5.1	Introduction	58
5.2	Capture des données statiques	58
5.3	Capture des données dynamiques	61
5.4	Synthèse	64
III	Évaluation de la fidélité visuelle	67
6	État de l'art : Fidélité visuelle	69
6.1	Introduction	70
6.2	L'interaction entre l'apparence et l'expressivité lors de l'identification . . .	70
6.3	Le rôle de la familiarité	77
6.4	Des métriques pour l'évaluation de visages virtuels	78
6.5	Synthèse	78
7	Étude perceptive pour l'étude de la fidélité visuelle	81
7.1	Introduction	82
7.2	Mise en place des niveaux de réalisme d'apparence	82
7.3	Méthode	83
7.4	Résultats	87
7.5	Discussion et synthèse	92
7.6	Conclusion	93
IV	Vers une fidélité comportementale	95
8	État de l'art : Fidélité cognitive	97
8.1	Introduction	98
8.2	Les différentes définitions d'une identité psychologique	100
8.3	La structure affective de l'identité	102
8.4	Approches, théories et modèles des émotions en psychologie	106
8.5	Modèle computationnel des émotions	111
8.6	Synthèse	114
9	Vers une fidélité comportementale : réflexions sur une architecture d'évaluation émotionnelle personnalisée	117
9.1	Introduction	118
9.2	Choix du cadre théorique des émotions	118

9.3	Justification du choix du modèle CPM	119
9.4	Architecture d'un modèle computationnel d'émotion personnalisé	121
9.5	Cas d'étude : Protocole d'élicitation GEMEP	124
9.6	Conclusion et perspectives	126
10	Conclusions et perspectives	129
10.1	Introduction	130
10.2	Démarche méthodologique	131
10.3	Contributions	131
10.4	Limites de nos travaux	134
10.5	Perspectives	135
	Références	139

Table des figures

1.1	Agent assistant médical (Bickmore et al, 2014)	3
1.2	Agent signeur (Heloir et al. 2008)	3
1.3	Exemple de rendu photo-réaliste de visage ¹	4
1.4	La doublure de l’actrice ”Ellen Page” dans Beyond two souls (Quantic dreams, sortie juin 2012)	4
1.5	Capture d’écran du film ”La fabuleuse histoire de Benjamin Button”	4
1.6	Emily project(Alexander, 2010)	5
1.7	Clone de Kevin Spacy dans Call of duty 4 (sorti novembre 2014)	5
1.8	Clone virtuel de Barack Obama (Metallo et al, 2015)	6
1.9	Plan du manuscrit	11
2.1	Exemple de clones virtuels dans le jeu vidéo (sortie 2015)	16
2.2	Évaluation holistique du visage (Young et al,1987)	19
2.3	L’importance de la pigmentation et de la forme du visage lors de la recon- naissance (Russel et al,2007)	20
2.4	Les éléments les plus importants lors de l’évaluation de visages (Ramamur- thy et al,2012)	20
2.5	Les sourcils, éléments les plus importants pour l’identification (Sadro et al., 2003)	21
2.6	L’importance de l’expressivité pour l’identification (knappmeyer et al.,2003)	22
2.7	Les régions les plus importantes pour l’évaluation de l’expressivité (Berisha et al.,2010)	22
2.8	L’effet des conditions d’illumination sur l’identification des visages (Bra- jet,1998)	22
2.9	L’effet de l’angle de vue sur l’identification des visages (Wallis,2001)	23
3.1	Cadre conceptuel pour l’étude de la fidélité	26
3.2	Cadre conceptuel pour l’étude de la fidélité : les dimensions de l’identité .	27
3.3	Cadre conceptuel pour l’étude de la fidélité : Fidélité psychologique	30
3.4	Cadre conceptuel pour l’étude de la fidélité : les entités impliquées	32

3.5	La relation entre les différentes entités	33
3.6	Cadre conceptuel pour l'étude de la fidélité	34
4.1	Évolution de rendu pour les visages virtuels	38
4.2	Taxonomie des techniques d'acquisition de formes 3D (Amor et al., 2011) .	39
4.3	Exemple de système à base de triangulation laser ¹	40
4.4	Extrait de la base de corps humains 3D CAESAR anthropometric database (Robinette,1999)	41
4.5	Extrait de la base de visage (Liu et al., 2004)	41
4.6	Exemple de système à base de lumière structuré (Von der Pahlen et al., 2014)	42
4.7	Une set-up de photogrammétrie de qualité professionnelle (infinite Realities ³)	43
4.8	LightStage Version 5 (Von der Pahlen et al., 2014)	44
4.9	Composantes lumineuses séparées, combinées, et photo de référence (Ma et al., 2008)	45
4.10	Extrait du film Terminator Genisys (2014) ⁴	46
4.11	Trajectoire de la lumière dans une surface partiellement transparente (Ji- menez, 2012)	47
4.12	Vue d'ensemble de l'algorithme de rendu de peau SSSSS : Screen-Space Sub- Surface Scattering (Jimenez, 2009)	48
4.13	Ensemble des Shaders pour un rendu photo-réaliste des yeux (Jimenez, 2012)	49
4.14	Ensemble des muscles faciaux (Frakas & Munro, 1987)	49
4.15	Les actions units correspondant à la moitié supérieure du visage(Ekman & Friesen, 1977)	50
4.16	Norme MPEG-4	51
4.17	Exemple de paramètres de Skinning. À Gauche : Positions des points-clés. À Droite : Influence de l'un des points-clés (Rouge : très fort, Vert : Moyen, Noir : Aucune influence) (Courgeaon et al., 2011)	52
4.18	Modèle d'animation musculaire (Sifakis et al., 2005)	53
4.19	Scan de position de Blendshapes au sein du projet Digital Ira (Von der Pahlen et al., 2014)	54
4.20	Hiroshi Ishiguro et son double robotique Geminoid HI-1 (Bartneck, 2009) .	55
5.1	Déroulement du scan Kinect	60
5.2	Données brutes du scan Kinect	60
5.3	Phase de retopologie du maillage avec le logiciel Zbrush	60
5.4	Plaquage de la texture	61
5.5	Disposition des marqueurs lors de la session de capture de mouvements . .	64
5.6	Pipeline de création de clones virtuels	65
6.1	Exemple de 10 styles de rendu classés du plus abstrait au plus réaliste (Mc- donnell et Al., 1012)	72
6.2	Étude sur la modification de la forme d'un visage 3D (ring et Al., 2014) . .	73

6.3	Texture réaliste, bronze et en trait variant sur 3 niveaux de détail (Macdor- man, 2009)	74
6.4	Signature dynamique : exemple d'expression caractéristique de l'hypothèse de "l'information supplémentaire"	75
6.5	Étude de l'apport de l'expressivité pour l'identification de visage célèbre (Bennetts et al., 2013)	77
7.1	Exemples de stimuli statiques utilisés dans l'expérimentation	83
7.2	Impact de la variable (D)dynamique sur l'évaluation de fidélité	87
7.3	Effet de l'interaction entre les variables dynamique(D) et rendu(R) sur l'évaluation de fidélité	88
7.4	Évaluation de la fidélité sur la base du rendu et de la familiarité avec le réfèrent(1= non-familier, 2= modérément familier, 3= très familier)	89
7.5	La différence entre l'évaluation de la fidélité dans la condition statique et la condition dynamique par groupe de familiarité	90
7.6	Interaction du rendu et de la dynamique dans le jugement de la "Likeability" .	91
7.7	L'impact de la familiarité sur l'évaluation de la "Likeability"	91
8.1	Le Circomplexe de Plutchik	107
8.2	Modèle OCC (Ortony et al., 1990)	109
8.3	architecture dynamique du CPM (Scherer, 2005)	110
8.4	Modèles computationnels de l'émotion (Maresella, 2010)	111
8.5	Les principaux modelés computationnels des émotions se basant sur la théorie de l'appraisal (Rodriguez, 2014)	112
8.6	L'architecture EMA (Marsella, 2009)	113
8.7	Modèle cognitif pour un agent virtuel intégrant une fonction affective basée sur CPM (Marinier et al., 2005)	114
8.8	Représentation computationnelle du processus émotionnelle faciale (Grizard, 2006)	115
9.1	Architecture séquentielle du modèle CPM (Sander et al., 2005)	120
9.2	Prédiction d'apparition des actions unitaires (AUs) suite à l'évaluation de SEC de CPM pour 5 émotions de base (Scherer, 2007)	121
9.3	Ensemble des SEC retenus pour l'évaluation émotionnelle	122
9.4	Modèle computationnel pour une expressivité émotionnelle fidèle	124
9.5	Annotation manuelle d'un script du corpus GEMEP pour un même individu	125

Liste des tableaux

8.1	Les questionnaires de mesure des traits émotionnels	105
-----	---	-----

Introduction générale

Sommaire

1.1	Contexte général	2
1.1.1	Un intérêt grandissant pour les agents virtuels	2
1.1.2	Clones virtuels	4
1.1.3	Attrait des visages familiers	6
1.2	Motivation	7
1.3	Problématiques et objectifs scientifiques	7
1.3.1	Qu'est-ce que la fidélité d'un clone virtuel?	8
1.3.2	Comment concevoir un clone virtuel fidèle?	8
1.3.3	Comment concevoir un clone virtuel interactif et fidèle?	8
1.3.4	Comment évaluer l'impact de la perception et de l'interaction avec un clone virtuel	9
1.4	Démarche incrémentale	9
1.4.1	Paradigme conceptuel pour l'étude de la fidélité	10
1.4.2	Fidélité visuelle	10
1.4.3	Fidélité comportementale	10
1.5	Organisation du mémoire	11

1.1 Contexte général

1.1.1 Un intérêt grandissant pour les agents virtuels

Un Agent Conversationnel Animé (ACA) est défini par Cassell et al. (Cassell, 2000) comme une interface homme-machine, potentiellement multimodale qui peut combiner : expressions faciales, gestes corporels et de la parole pour permettre une communication "face à face" avec les utilisateurs.

Certains utilisateurs sont susceptibles de manquer d'expérience quant à l'utilisation des interfaces classiques proposées, ils peuvent alors s'égarer dans une masse considérable d'informations. Il est donc devenu nécessaire de concevoir des interfaces plus ludiques, plus intuitives, plus ergonomiques et plus appropriées à une utilisation grand public. Les personnages virtuels pourraient permettre une communication intuitive avec les utilisateurs, car ils s'appuient sur les modes de communication que nous utilisons quotidiennement lors de nos interactions sociales. L'agent virtuel permettrait donc de répondre à des besoins de l'utilisateur et pallier aux lacunes d'autres sources d'informations peu ergonomiques et présentant trop d'informations.

Ces dernières années ont vu la prolifération des agents virtuels avec diverses applications. Par exemple les agents relationnels sont des personnages virtuels capables de s'engager dans une interaction long terme avec des utilisateurs (T. W. Bickmore & Picard, 2005).

Des travaux tels que (Mehlmann & André, 2015 ; Zaïane, 2002 ; Kenny et al., 2007 ; Lisetti, Amini, & Yasavur, 2015 ; Murakami, Sugimoto, & Ishida, 2005 ; T. Bickmore & Giorgino, 2006) ont observé que les agents virtuels améliorent la réalisation de tâches dans des domaines divers tels que : l'éducation, l'apprentissage, l'assistance médicale, la formation professionnelle ou simplement le divertissement.

À mesure que les champs d'application des agents virtuels se diversifient, les aptitudes attendues de ces derniers grandissent. Afin de répondre aux besoins et aux exigences de nouveaux utilisateurs les agents virtuels s'enrichissent de nouvelles fonctionnalités. Par exemple, avec l'apparition des agents conversationnels animés, des aptitudes émotionnelles se sont avérées nécessaires pour établir une communication naturelle. Apparurent alors dans ces agents virtuels des modèles d'informatique affective pour guider une animation expressive émotionnelle qui soit cohérente avec l'interaction (S. Marsella, Gratch, & Petta, 2010). Des architectures modélisant des différences interindividuelles (tel que des traits de personnalités) ont également été modélisées afin d'augmenter le réalisme des comportements des agents (Allbeck & Badler, 2002 ; Gebhard, 2005).

Les agents virtuels au service de la recherche. De nombreux axes de recherches s'appuient sur des agents virtuels pour explorer diverses problématiques expérimentales. Par exemple, les psychologues sociaux étudient les théories de la communication en modifiant le comportement verbal et non verbal d'un humain virtuel. Des patients souffrant



FIGURE 1.1: Agent assistant médical (Bickmore et al, 2014)



FIGURE 1.2: Agent signeur (Helloir et al. 2008)

de phobies sociales peuvent simuler une situation sociale menaçante dans la sécurité d'un environnement virtuel (Bailenson & Yee, 2005 ; Kang, Gratch, Wang, & Watt, 2008). En effet, un des principaux apports des agents virtuels est de fournir un environnement contrôlé pour des études expérimentales. Ils permettent ainsi à la fois de mieux comprendre la perception humaine et les interactions sociales entre humains. L'agent virtuel, en offrant un cadre d'interaction dont les paramètres sont contrôlables, nous permet d'étudier le poids respectif des différentes variables qui entrent en jeu lors de cette interaction.

Une quête de plus en plus importante du photoréalisme. Les humains virtuels étant souvent déployés dans un environnement 3D, ils ont par conséquent bénéficié des avancées considérables de l'informatique graphique de ces dernières années (Figure 1.3)¹. Cependant les agents virtuels, spécialement ceux utilisés dans les milieux académiques et de la recherche, ne jouissent pas de tout le potentiel de ce domaine qui est en perpétuelle progression.

En effet, le monde de la 3D (par exemple dans le domaine du jeu vidéo, des effets spéciaux et du cinéma) a connu ces dernières années une quête incessante vers le « photoréalisme » (e. i. un rendu visuel tellement détaillé qu'il peut passer pour une photographie). L'un des principaux objectifs dans le développement d'environnements virtuels a été de créer des paysages, personnages et humains de plus en plus réalistes au point de ne plus faire le discernement entre le réel et le virtuel.

1. <http://www.chrisj.com.au/> (mise en ligne SEPTEMBER 26, 2014)



FIGURE 1.3: Exemple de rendu photo-réaliste de visage ¹

Une nouvelle génération de moteurs de rendus a fait son apparition : ils sont dits « unbiased » (sans approximation) et calculent la lumière selon les lois de la physique. Alors qu’au même moment les recherches portant sur l’interaction avec des « Humains virtuels » se contentent de rendu graphique plus rudimentaire, le plus souvent cartoon au du moins non réaliste. Pourtant, des études sur le réalisme (MacDorman, 2010) suggèrent que l’utilisation de personnages plus réalistes visuellement peut augmenter l’engagement des utilisateurs dans l’interaction avec l’agent virtuel.

1.1.2 Clones virtuels

Ces dernières années, le monde des humains virtuels a connu un intérêt croissant pour des agents ayant un aspect familier ou du moins reconnaissable. Cet intérêt s’est surtout manifesté à travers l’industrie audio-visuelle et celle du jeu vidéo (Figure 1.4 - 1.5). Les



FIGURE 1.4: La doublure de l’actrice ”Ellen Page” dans Beyond two souls (Quantic dreams, sortie juin 2012)



FIGURE 1.5: Capture d’écran du film ”La fabuleuse histoire de Benjamin Button”

dernières techniques de rendu 3D et les dernières avancées des scans 3D ont permis la

création de visages virtuels photo-réalistes modélisant des personnes existantes. Grâce aux dernières avancées en informatique graphique (Jimenez, 2012) et les systèmes de scan3D (O. Alexander et al., 2010 ; Fyffe, Jones, Alexander, Ichikari, & Debevec, 2014), il est devenu difficile de discerner un visage humain d'un visage généré par ordinateur (figure 1.6) comme on peut le constater dans « the Emily Project » (O. Alexander et al., 2010) ou dans les jeux vidéo récents comme "Call of duty : advanced warecraft" sorti en novembre 2014 (figure 1.7²). L'industrie du cinéma s'appuie sur cette technologie lors de ses plus grands succès. Les films « avatar » ou « Tron l'héritage » sont des exemples parmi d'autre où une doublure numérique de célébrité telle que « Sigourney Weaver » furent réalisés sans que forcément le spectateur s'en rende compte.



FIGURE 1.6: Emily project(Alexander, 2010)



FIGURE 1.7: Clone de Kevin Spacey dans Call of duty 4 (sorti novembre 2014)

Dans la suite de ce manuscrit, nous définissons un *clone virtuel* comme un personnage virtuel généré par ordinateur qui a été créé numériquement en capturant une personne existante. La phase de capture peut inclure une capture de la géométrie de son visage, son expressivité et des éléments de sa psychologie. Dans ce qui suit, nous désignons le modèle humain dont sont issues toutes ces informations comme étant « *le référent* ». Par exemple la figure 1.8 présente un clone virtuel dont le référent est le président des états unis Barack Obama (Metallo et al., 2015)

2. <http://culturebox.francetvinfo.fr/le-bento/2014/11/06/le-nouveau-role-de-kevin-spacey-call-of-duty-advanced-warfare.html>



FIGURE 1.8: Clonage virtuel de Barack Obama (Metallo et al, 2015)

1.1.3 Attrait des visages familiers

Les visages appartiennent à notre quotidien, ils sont omniprésents dans nos interactions et nos relations sociales, et leur reconnaissance fait partie de ces mécanismes automatiques que l'on ne remarque même plus.

L'identification est primordiale dans nos interactions sociales quotidiennes. L'importance de la perception du visage est révélée de façon spectaculaire quand on perd la capacité de reconnaître des visages connus (ce trouble porte le nom de prosopagnosie). Il se manifeste au quotidien chez les prosopagnosiques par des difficultés importantes à reconnaître les visages familiers, tels que des membres de leur famille, des amis, des collègues de travail, des personnes célèbres dans les médias ou encore même leur propre visage (Bussigny, 2014). De nombreuses expériences sur la perception et la reconnaissance des visages existent dans la littérature. Un élément primordial qui entre en jeu dans ces études est la *familiarité* (Gobbini, Leibenluft, 2004; Persike et al., 2013; Young, Hay 1985; Natu O'Toole, 2011). Les études en neurologie et psychologie ont démontré que notre traitement cognitif face à des visages familiers diffère du traitement des visages inconnus (Lander, 2005, wallraven, 2007, etc.). Ces études avancent, entre autres, que le traitement de visages familiers est cognitivement moins coûteux et que l'attraction du regard par les visages familiers est automatique (Lander, Christie, & Bruce, 1999; Gobbini & Haxby, 2007).

L'effet de la familiarité. La capacité d'identifier une personne que nous rencontrons détermine la façon dont nous devons interagir avec cette personne. Notre comportement change rapidement en fonction de si nous sommes amenés à interagir avec un ami, un membre de la famille, ou notre patron. Ce changement de rapports, dû à notre familiarité avec l'interlocuteur, est totalement automatique.

Les effets de familiarité sont donc importants. Ils ont été utilisés pour expliquer la rai-

son pour laquelle la proximité prédit l'attraction et l'affinité. En effet (Back, Schmukle, et Egloff, 2008; Festinger, Schachter, et Back, 1950; Nahemow et Lawton, 1975; Segal, 1974) avancent que le sentiment de familiarité favorise le sentiment d'appréciation « Likeability ». Harrison et al. soutiennent par exemple que plus un visage ou un nom d'une figure publique est familier, plus cette personne est perçue comme étant sympathique (likeable) (Harrison, 1969). Ce phénomène est expliqué par Zajonc dans (Zajonc, 1986) comme résultant de « L'effet de simple exposition ». Cet effet se caractérise par une augmentation de la probabilité d'avoir un sentiment positif envers quelqu'un ou quelque chose par la simple exposition répétée à cette personne ou cet objet. En d'autres termes plus nous sommes exposés à un stimulus (personne, produit de consommation, lieu) et plus il est probable que nous l'aimions.

1.2 Motivation

Nous avons constaté que le champ de recherche sur les agents virtuels est de plus en plus mature. Les humains virtuels sont d'ores et déjà dotés de plusieurs compétences sociales et ils ont dès à présent plusieurs applications industrielles (même si leurs capacités y restent limitées). De plus l'industrie audiovisuelle est arrivée à générer des visages virtuels de personnes reconnaissables qui défient nos capacités à discerner le fait qu'ils soient virtuels ou réels. Nous avons également constaté que l'effet de la familiarité est une des composantes primordiales de nos interactions sociales.

Pourtant très peu de travaux ont tenté d'étudier conjointement ces deux axes de recherches en tentant d'octroyer aux agents virtuels autonomes une identité d'une personne existante.

L'objectif de cette thèse consiste à explorer cet axe de recherche en examinant le processus de perception d'un visage virtuel familier et d'étudier l'intérêt d'un agent virtuel véhiculant une identité reconnaissable.

Les travaux effectués dans cette thèse portent donc sur la conception, la modélisation et l'animation émotionnelle d'un clone virtuel d'une personne. Des contraintes de fidélité du clone virtuel vis-à-vis de la personne réelle devront donc être formalisées en termes d'apparence, d'expressivité et de comportement émotionnel autonome. Nous nous concentrons sur les dimensions de l'apparence visuelle et les expressions dynamiques du visage, car le visage est une modalité principale pour exprimer des signaux sociaux importants tels que les émotions (Ekman & Friesen, 1977).

1.3 Problématiques et objectifs scientifiques

De notre problématique principale portant sur la modélisation et l'évaluation d'un clone virtuel découlent plusieurs questions de recherche :

1.3.1 Qu'est-ce que la fidélité d'un clone virtuel ?

Dans nos travaux, nous considérons que l'évaluation d'un clone virtuel revient à évaluer la fidélité du visage virtuel avec celui du référent (le personnage humain dont sont issues les données d'apparence et d'expressivité).

La fidélité a plusieurs définitions en fonction de l'axe de recherche exploré, mais il manque dans la littérature existante un cadre conceptuel identifiant ses principales dimensions pertinentes. La fidélité invoque diverses variables et implique de nombreux paramètres extrinsèques et intrinsèques aux visages ou plus généralement aux objets évalués. La première problématique de ces travaux consiste donc en la définition de cette notion avec ses différentes composantes ainsi que les différentes variables qui entrent en jeu lors du processus d'évaluation.

1.3.2 Comment concevoir un clone virtuel fidèle ?

L'informatique graphique connaît des avancées considérables en termes de techniques de rendu et d'animation. Les rendus récents mettent en défi notre perception pour distinguer un visage virtuel (généré en 3D) d'une image réelle. En même temps, très peu d'études se sont penchées sur la question de l'évaluation de ces travaux. Dans les multitudes de possibilités de techniques de rendu, d'animation et d'acquisition de données (scan3d) qu'offre l'informatique graphique, quelle serait la meilleure combinaison pour assurer le plus haut degré de fidélité ? Quels sont les éléments déterminants qui nous permettent d'évaluer la ressemblance avec un référent ?

1.3.3 Comment concevoir un clone virtuel interactif et fidèle ?

Les clones intègrent de plus en plus des applications temps réels. Par conséquent, ils sont amenés à être une source d'interaction. Cependant, nous n'avons pas rencontré jusqu'à présent d'étude qui se soit intéressée à doter des clones virtuels de vraies personnes d'un comportement autonome similaire à celui des personnes clonées.

Certains auteurs (Gobbini, 2006) soutiennent que l'aspect visuel n'est qu'une composante de reconnaissance des personnes familières. D'autres aspects fondamentaux comprennent aussi le traitement des connaissances que nous avons sur cette personne tel que les traits de personnalité et la réponse émotionnelle que nous éprouvons face à cette personne.

Nous remarquons que d'un côté l'aspect visuel d'un clone est généralement issu de données réelles (scan, mocap) leur procurant un caractère unique. D'un autre côté, les modèles computationnels existants qui simulent un comportement autonome sont normatifs (génériques) ou sont spécifiés au mieux pour un trait de personnalité mais pas pour un individu. L'identité dans ces travaux est plutôt assimilée à un rôle donné à plusieurs agents (Di Loreto, Hervouet, et al., 2012). Par conséquent, les modèles actuels d'IA pourraient ne pas être adéquats et intégrables pour des clones puisqu'un modèle normatif pourrait être incongruent avec un visuel personnalisée.

Il nous faut donc définir ou adapter un modèle computationnel individuel des émotions

pour guider l’animation expressive de manière à ce qu’elle soit fidèle aux réactions du référent. De cette problématique découlent plusieurs questions de recherche :

1. Quel élément cognitif pouvons-nous modéliser afin de donner ce ressenti de fidélité ?
2. Simuler une interaction humaine affective implique de doter l’agent d’une représentation interne d’état émotionnel et des raisonnements associés et de permettre à l’agent d’exprimer son état émotionnel interne. Cela implique de choisir parmi les théories et modèles proposés en psychologie les aspects les plus pertinents qui décrivent le mieux un comportement émotionnel expressif fidèle.
3. Enfin, comment évaluer l’apport de ce modèle dans l’évaluation d’un clone virtuel ? À quel point le degré de fidélité psychologique impacte-t-il la perception d’une identité sur un clone virtuel ?

1.3.4 Comment évaluer l’impact de la perception et de l’interaction avec un clone virtuel

Nos travaux sur la fidélité d’apparence, d’expressivité et de comportement autonome ont pour but d’améliorer la perception et l’interaction entre un clone virtuel et des utilisateurs. Il est donc nécessaire d’évaluer nos travaux par des expérimentations perceptives. Plusieurs questions de recherche sur la perception et l’interaction entre un utilisateur et un clone virtuel se posent.

D’une part, concernant la fidélité visuelle du clone virtuel : les humains utilisent-ils les mêmes procédés pour identifier un clone virtuel qu’une photo/vidéo de l’humain réel correspondant ? Sur quelles caractéristiques nous basons-nous pour juger la fidélité de ces clones virtuels ? Accordons-nous la même importance à toutes les composantes de la fidélité ? Les degrés de réalisme de l’apparence et de l’expressivité jouent-ils des rôles équivalents dans l’évaluation de la fidélité d’un clone virtuel ?

D’autre part, nous devons évaluer la fidélité comportementale dans un contexte d’interaction avec un agent familier. Quels sont les apports et les avantages d’interagir avec un clone virtuel plutôt qu’avec un agent ne possédant pas une identité ? Observons-nous une modification du comportement de l’utilisateur, de ses performances ou de ses perceptions subjectives lors d’une tâche interactive ?

1.4 Démarche incrémentale

Cette thèse traite de plusieurs notions sur lesquelles ils n’existent pas de consensus dans la littérature. Il existe différentes définitions de la fidélité, du réalisme, de l’identité, etc. De plus nous n’avons pas rencontré de travaux qui ont soulevé la question de l’évaluation de la fidélité de clones virtuels. Nous avons donc dû aborder la problématique de manière incrémentale sur trois principales étapes successives s’articulant autour de la notion de fidélité.

- Il fallait dans un premier abord proposer un paradigme conceptuel pour donner un cadre à nos travaux sur la fidélité.
- Ensuite, nous nous sommes penchés sur l'étude de la « *fidélité visuelle* » avec un clone virtuel non autonome.
- Finalement, nous avons entamé l'étude de la « *fidélité comportementale* ». Nous avons incrémenté l'architecture définie dans la deuxième étape avec un niveau cognitif pour assurer un comportement autonome à l'agent similaire par certaines caractéristiques à celui de la personne référente.

1.4.1 Paradigme conceptuel pour l'étude de la fidélité

Cette étape a eu pour but de définir les différents concepts et les différentes variables qui sous-tendent le processus de jugement de fidélité entre un clone virtuel et son référent humain. Pour cela, il a été nécessaire d'étudier des domaines de recherches multidisciplinaires issus de la perception de visage, de la psychologie, de la réalité virtuelle, etc.

1.4.2 Fidélité visuelle

Comme le soutiennent les études en sociologie, les humains déduisent l'identité de l'autre en se basant principalement sur des attributs visuels (Goffman, 2002). Cette étape a eu pour but de modéliser, d'implémenter et d'étudier l'interaction entre les différentes composantes de la fidélité visuelle. Pour cela plusieurs sous-étapes furent nécessaires.

Édition d'un clone Dans cette étape, nous décrivons les principaux composants matériels et logiciels permettant de capturer, rendre et animer un clone virtuel qui assure la perception de l'identité du référent d'origine.

Étude perceptive La deuxième sous-étape fut de concevoir une étude perceptive pour l'évaluation des poids respectifs des composantes de la fidélité visuelle des clones virtuels créés à l'étape précédente. Cela nécessita la conception, l'implémentation et l'évaluation de différent degré de fidélité visuelle au sein d'un avatar non autonome.

1.4.3 Fidélité comportementale

Finalement la dernière étape consiste en la proposition d'une adaptation d'un module computationnel émotionnel qui enrichit le modèle prédéfini précédemment pour présenter un avatar autonome et fidèle. Cette étape comporte à la fois l'étude des modèles et des architectures psychologiques de la cognition et de l'émotion, des premières propositions d'implémentation informatique du modèle et des propositions de protocoles d'expérimentation pour l'évaluation de ses apports.

1.5 Organisation du mémoire

Notre démarche incrémentale se reflète dans l'organisation du présent manuscrit comme le montre la figure 1.9 : Le mémoire se compose de 4 parties principales. Pour chaque partie nous introduisons une question de recherche à travers une brève étude de l'existant. Nous essayons par la suite d'apporter des éléments de réponses à ces problématiques par un chapitre "contribution" où nous présentons nos travaux.

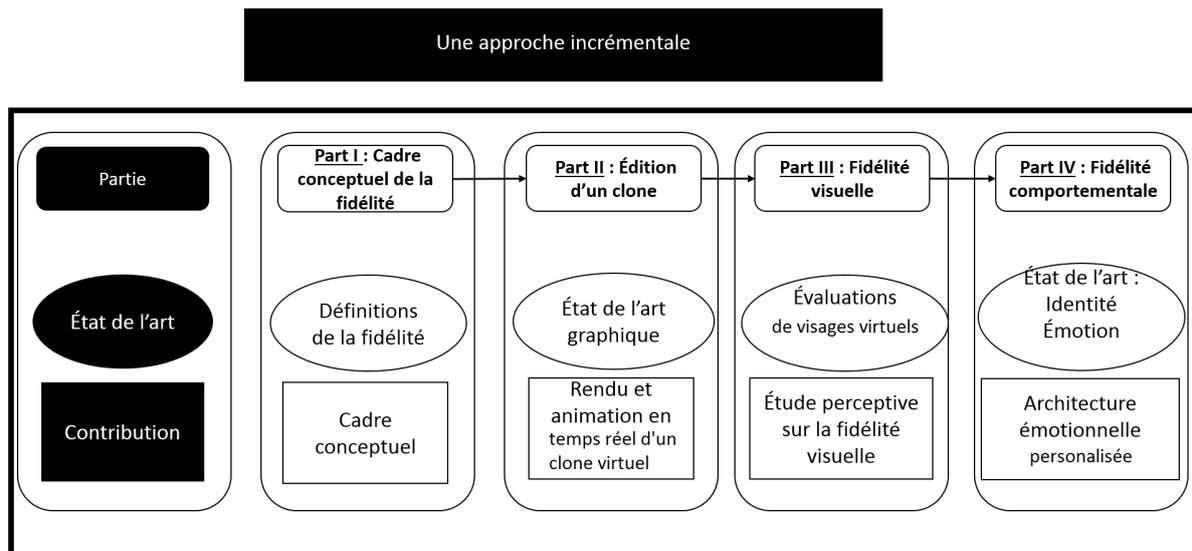


FIGURE 1.9: Plan du manuscrit

Première partie

Paradigme conceptuel pour l'étude de la fidélité

Vers une définition de la fidélité pour des clones virtuels

Sommaire

2.1	Pourquoi s'intéresser à la fidélité ?	16
2.2	Les différentes définitions de la fidélité	17
2.2.1	La fidélité de point de vue de la réalité virtuelle	17
2.2.2	La fidélité vue par la sociologie	18
2.3	Les différentes composantes de la fidélité d'un visage	18
2.3.1	Les caractéristiques du système visuel humain pour la perception des visages	18
2.3.2	Le rôle de l'expressivité dans la reconnaissance des visages	21
2.3.3	Les variables extrinsèques qui influencent la perception des visages	22
2.4	Différences interindividuelles lors de la perception et de l'identification d'un visage	23
2.5	Limites de l'existant & conclusions	24

2.1 Pourquoi s'intéresser à la fidélité ?

L'informatique graphique et la modélisation de visages virtuels ont connu des avancées considérables ces dernières années. En effet, les techniques récentes de rendu et d'animation sont si réalistes que nous avons du mal à discerner un rendu 3d d'une photo. En outre, l'utilisation de visages 3D représentant des personnages familiers est de plus en plus répandue (e.g. les stars sportives, des stars cinématographiques dans les différents jeux basés sur des licences de film).

Néanmoins, à notre connaissance, la question de comment évaluer la réussite de ces travaux ne s'est pas posée dans les milieux de la recherche.

Pour illustrer notre problématique, nous allons donner un exemple simple à travers la figure 2.1¹. Celle-ci représente deux exemples de clones virtuels dans les deux plus célèbres licences de jeux vidéo de football. Ces clones ont été modélisés pour représenter la même personne, pourtant ils sont assez différents l'un de l'autre.

Une des motivations principales qui a initié ces travaux est de comprendre sur quels éléments se baserait un utilisateur pour évaluer de telles réalisations. À la question "quel est l'avatar le plus réussi ?", un utilisateur est amené intuitivement à juger lequel des deux clones est le plus ressemblant. Par conséquent, ce dernier devra juger la fidélité du clone virtuel vis-à-vis de l'image réelle ou de la représentation mentale qu'il a de l'humain d'origine. Ainsi, une des principales questions de recherche de cette thèse est :

Comment un utilisateur évaluerait-il la fidélité entre un clone virtuel et le référent humain ? Quelle est la définition de la fidélité d'un visage de clone virtuel et quelles sont ses différentes composantes ?

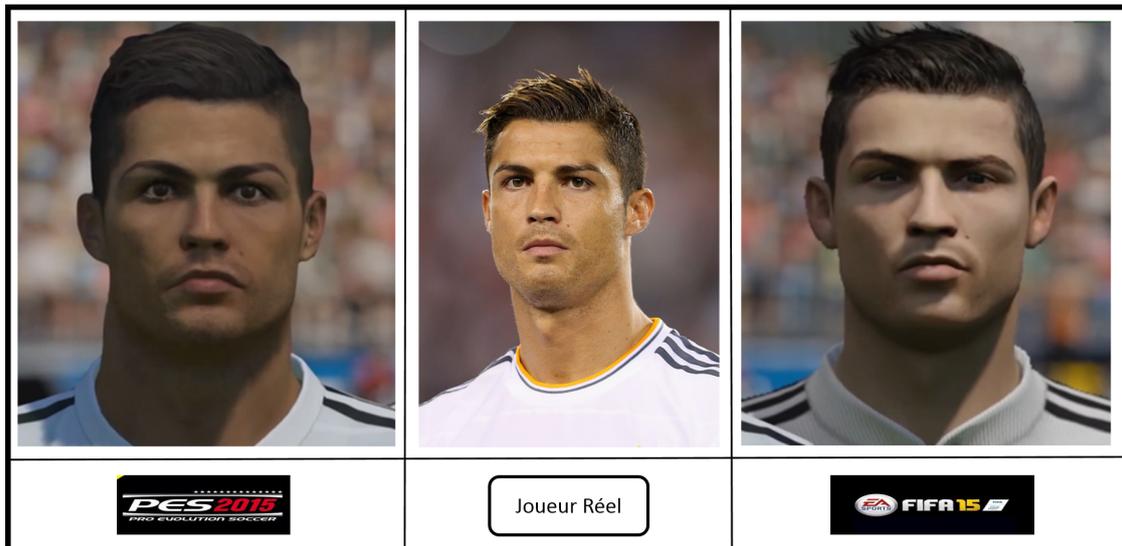


FIGURE 2.1: Exemple de clones virtuels dans le jeu vidéo (sortie 2015)

1. <https://www.youtube.com/watch?v=WeqoPLPqkac> (EA sports 2015)

Dans nos travaux, nous considérons donc la *fidélité* obtenue comme métrique pour évaluer des travaux portant sur un visage de clone virtuel. Dans un premier abord, il nous a fallu donc explorer et définir cette notion de fidélité.

2.2 Les différentes définitions de la fidélité

La fidélité est une notion ambiguë. Cette dernière n'a pas de définition unique et unanime. Les définitions données dans la littérature dépendent de l'axe de recherche exploré. Dans ce chapitre nous allons considérer les différents champs de recherche qui ont abordé la définition de fidélité afin d'en extraire une conception qui soit adéquate à l'étude d'un visage virtuel en trois dimensions.

2.2.1 La fidélité de point de vue de la réalité virtuelle

La réalité virtuelle fut une des premières disciplines qui a tenté de proposer une définition de la fidélité. En effet, cet axe de recherche utilise la fidélité comme unité de mesure et de validation pour ses différentes simulations déployées dans un environnement virtuel (A. L. Alexander et al., 2005).

La fidélité dans le cadre des simulateurs a été définie comme étant "la mesure dans laquelle l'environnement virtuel simule le monde réel" (Allen et al., 1986). Cet axe de recherche a subdivisé la fidélité en plusieurs sous dimensions telles que la fidélité physique, fonctionnelle et psychologique.

2.2.1.1 Fidélité physique

Les auteurs dans (A. L. Alexander et al., 2005) ont défini la fidélité physique comme étant «la mesure dans laquelle un simulateur ressemble d'un point de vue visuel, sonore et tactile à l'environnement opérationnel». Selon Alexander et al (A. L. Alexander et al., 2005), la fidélité physique elle-même englobe un certain nombre de sous-dimensions relatives à différentes modalités : visuel, auditif, vestibulaire, olfactive et proprioceptive.

La fidélité visuelle est particulièrement pertinente pour nos recherches sur la fidélité des visages virtuels. Cette dernière est définie comme étant le degré auquel les composantes visuelles de l'environnement virtuel sont conformes aux composantes visuelles de l'environnement réel.

2.2.1.2 Fidélité psychologique

Dans nos différentes lectures abordant la définition de la fidélité dans un environnement virtuel (simulateur) une autre dimension a retenu notre attention : "La fidélité psychologique". Elle est également définie dans (Wallraven et al., 2005) comme étant le degré selon lequel la simulation reproduit les facteurs psychologiques ressentis lors de la simulation (tel que le stress et la peur).

La fidélité psychologique est fortement liée à la notion de présence et au sentiment d'immersion (Bailenson, Blascovich, & Guadagno, 2008). L'immersion a été définie comme étant «l'illusion que les stimuli sensoriels créés par des dispositifs technologiques sont réels» (Kasap et Magnenat-Thalmann 2009). Plusieurs études ont montré que les personnages virtuels peuvent créer un sentiment de présence sociale (Bailenson, Blascovich et al, 2001 ; Blascovich 2002 ; McQuiggan, Rowe et al., 2008)

2.2.2 La fidélité vue par la sociologie

Dans notre quête pour la recherche d'une définition de la fidélité, nous nous sommes également intéressés à la notion d'identité dans le cadre d'étude sociologique plus particulièrement aux études explorant la définition d'une identité numérique. Par exemple, Goffman et al (Goffman, 2012) définit l'identité comme un modèle mental que chacun a sur lui même. Goffman associe l'identité à un rôle de théâtre permanent que chacun joue quotidiennement.

Goffman dans sa définition de l'identité fait la distinction entre «l'information donnée» c'est à dire une information intentionnelle et préméditée par la personne et une information dite «fuite» sur laquelle la personne n'a ni intention ni contrôles (Goffman, 2002).

Les travaux issus de la psychologie sociale soutiennent donc que l'identité est non seulement déduite par autrui à partir de nos des attributs visuels, mais qu'elle est également évaluée sur des éléments du comportement. Nous présentons également notre identité à travers nos actions et nos interactions.

2.3 Les différentes composantes de la fidélité d'un visage

Un des objectifs clés de nos travaux et de l'informatique graphique est de créer des rendus qui pourraient duper le système de vision humain. Pour cela il nous est apparu impératif de nous intéresser aux résultats des études expérimentales sur la reconnaissance de visages réels. Ces résultats nous donnent un aperçu des indices sur lesquels se basent les humains lors de l'identification. Ils nous permettront par conséquent de dégager les différentes composantes de la fidélité visuelle.

2.3.1 Les caractéristiques du système visuel humain pour la perception des visages

2.3.1.1 L'évaluation holistique des visages

La recherche en psychologie cognitive et en neuroscience soutient que l'évaluation du visage est un processus holistique. Ainsi Tanaka et al. (Tanaka & Gordon, 2011) avancent que lors d'une tâche de reconnaissance de visage, tous les traits qui le composent sont traités

de façon globale, en une seule fois comme une totalité. Le contexte holistique semble affecter la façon dont les caractéristiques individuelles sont traitées. Il a été démontré que l'évaluation des relations géométriques entre chaque caractéristique et le reste du visage est très importante pour la reconnaissance faciale. Déformer cette relation peut fausser la reconnaissance, par exemple, (A. W. Young et al., 1987) ont observé que la superposition des deux moitiés supérieure et inférieure de différents visages génère une forte impression d'un nouveau visage.

L'étude menée par (A. W. Young et al., 1987) est une expérience perceptive où les caractéristiques d'une moitié supérieure d'un visage ont été combinées à la moitié inférieure d'un autre visage. Les deux identités distinctes sont avérées très difficiles à reconnaître comme l'illustre la figure 2.2 (Sinha et al., 2006).



FIGURE 2.2: Évaluation holistique du visage (Young et al,1987)

Le processus holistique comprend principalement l'évaluation de deux facteurs principaux : la forme, et la pigmentation (Russell et al., 2007). Plusieurs études ont été menées sur l'importance des facteurs holistiques pour l'identification. Par exemple, une étude a observé que des images de visage avec un contraste inversé (Yip & Sinha, 2002) ou une forme altérée sont très difficiles à reconnaître. Plusieurs études ont observé que la reconnaissance de visage en couleur est meilleure que la reconnaissance de visage en nuance de gris (Yip & Sinha, 2002).

La primauté de ces deux facteurs de l'un par rapport à l'autre a également été abordée. Il a été observé que la pigmentation de la peau et la forme du visage ont la même importance pour la reconnaissance des visages (Russell et al., 2007). En effet dans une étude de Russell et al (Russell et al., 2007) les auteurs ont fait varier sur un même visage sa forme et la couleur de la peau comme illustrée dans la figure 2.3. Les taux de reconnaissance étaient similaires suivant l'altération des deux variables.



FIGURE 2.3: L'importance de la pigmentation et de la forme du visage lors de la reconnaissance (Russel et al,2007)

2.3.1.2 Évaluation trait par trait

En surcroît de l'évaluation globale du visage, le système de vision humain évalue également les différents traits du visage de façon indépendante (Sinha et al., 2006). Tous les traits du visage ne jouent pas des rôles équivalents lors de la reconnaissance. Des études expérimentales basées sur l'eye tracking (suivi du regard) suggèrent que notre perception se concentre sur quelques parties spécifiques du visage. Afin de déterminer les facteurs du visage les plus importants, plusieurs évaluations perceptives ont été mises en place pour évaluer la fidélité entre deux visages réels. (Ramamurthy et al., 2012 ; Sinha et al., 2006).

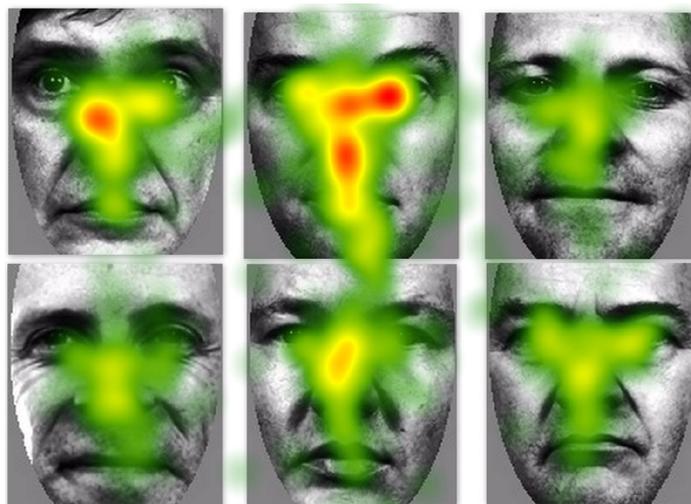


FIGURE 2.4: Les éléments les plus importants lors de l'évaluation de visages (Ramamurthy et al,2012)

Ces études ont montré, comme l'illustre la figure 2.4, que nous nous concentrons sur certaines zones du visage pour évaluer la ressemblance, essentiellement les yeux, suivi de la bouche et enfin du nez.

Sadro et al. (Sadrô et al., 2003) affirment même que les sourcils pourraient être les traits les plus déterminants lors d'une tâche d'identification (surpassant même les yeux). En effet, les taux de reconnaissance de photos dépourvues des yeux (figure 2.5) étaient plus importants que celle dépourvue de sourcils. Selon les auteurs, les sourcils auraient cette importance, car ces derniers sont les éléments les plus susceptibles de véhiculer une importante information de mouvement au sein du visage.

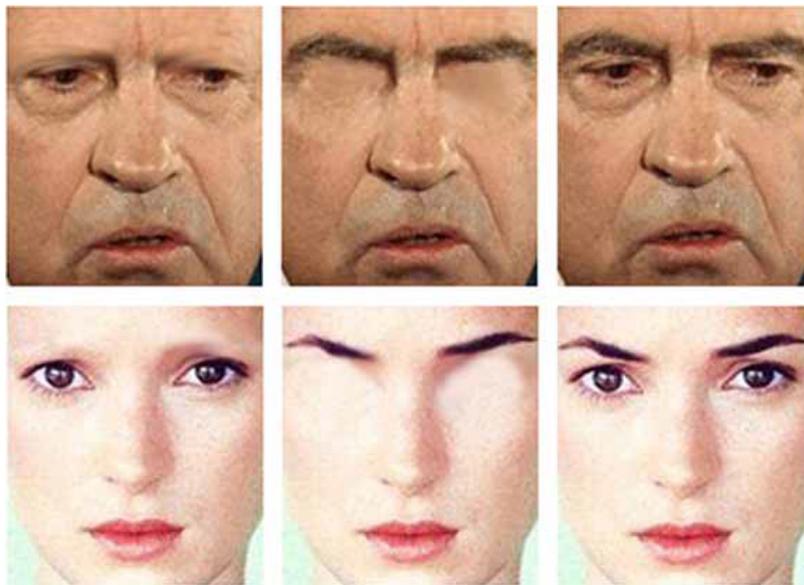


FIGURE 2.5: Les sourcils, éléments les plus importants pour l'identification (Sadro et al., 2003)

Les résultats de (Sadrô et al., 2003) nous ont incité à nous intéresser au rôle du mouvement dans la détermination d'une identité.

2.3.2 Le rôle de l'expressivité dans la reconnaissance des visages

Un des travaux précurseurs du rôle de l'expressivité dans la reconnaissance des visages est celui de Lander et Chuang (Lander et al., 1999). Ils ont observé que les expressions faciales facilitent la reconnaissance d'une personne familière célèbre même quand ces images sont dégradées ou floutées. Leurs résultats suggèrent que chaque visage possède des caractéristiques de mouvement distinctif qui constituent une composante à part entière de l'identité d'un visage.

Une expérience menée par (Knappmeyer et al., 2003) a démontré que sur un visage intermédiaire morphée entre deux identités, l'expressivité de chacun joue un rôle important dans la détermination de la personne (figure-2.6). Dans une étude, plus récente Berisha et

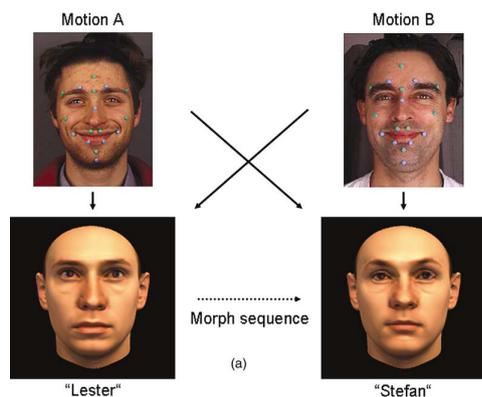


FIGURE 2.7: Les régions les plus importantes

FIGURE 2.6: L'importance de l'expressivité pour l'évaluation de l'expressivité (Berisha et al., 2010) pour l'identification (Knappmeyer et al., 2003) al., 2010)

al. (Berisha et al., 2010) ont essayé de déterminer les parties du visage les plus importantes lorsque nous évaluons son expressivité. Ils ont confirmé les résultats suggérés de Sadrô (Sadrô et al., 2003) sur l'importance des sourcils. En effet grâce à une étude basée sur l'eye tracking, les régions faciales sur lesquelles les regards se concentrent semblent être l'arcade sourcilière puis la bouche comme l'illustre la figure 2.7.

2.3.3 Les variables extrinsèques qui influencent la perception des visages

Mise à part la variabilité de ces propriétés intrinsèques du visage considérée dans les sections précédentes, plusieurs études se sont intéressées également aux phénomènes liés aux conditions environnementales et aux conditions d'observation qui peuvent affecter la perception d'un visage.

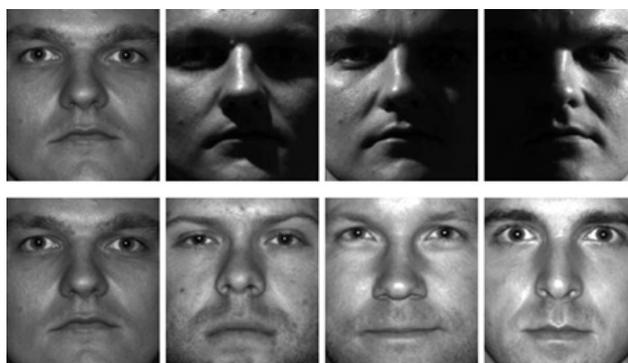


FIGURE 2.8: L'effet des conditions d'illumination sur l'identification des visages (Brajt, 1998)

De nombreuses expériences suggèrent que la perception du visage est sensible aux : conditions d'illumination (Braje et al., 1998) (Figure 2.8), l'angle de vue et la pose du modèle (Figure 2.9) (Wallis & Bühlhoff, 2001). Des études ont été également menées sur la tolérance de la reconnaissance à la dégradation de la résolution de l'image et de l'effet de la distance de jugement. (Sinha et al., 2006)

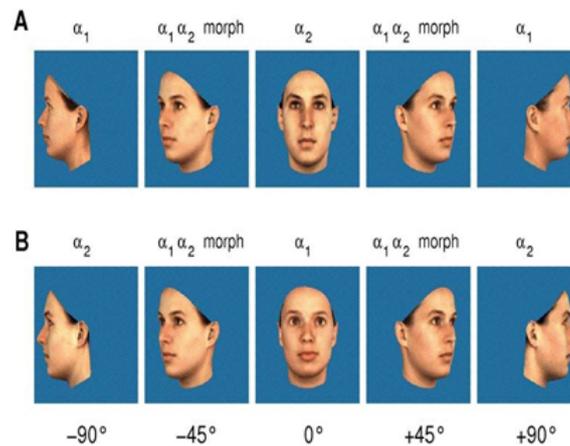


FIGURE 2.9: L'effet de l'angle de vue sur l'identification des visages (Wallis, 2001)

2.4 Différences interindividuelles lors de la perception et de l'identification d'un visage

Étant donné le rôle central que peut avoir le visage dans la cognition sociale, certains concepts sociaux jouent un rôle indéniable lors du processus de perception et d'identification des visages. En effet (S. G. Young et al., 2012) affirment que des groupes sociaux tels que le groupe ethnique, l'âge, le sexe et même le fait d'être en endogroupe ou en exogroupe avec le visage perçu impactent la perception et la mémorisation des visages.

L'ethnicité semble être l'un des facteurs les plus importants qui entrent en ligne de compte lors de l'identification d'un visage. Un visage asiatique est plus difficile à reconnaître et à mémoriser pour une population caucasienne par exemple. Respectivement les gens semblent avoir une plus grande sensibilité pour les visages de leurs propres ethnies (Calder, Rhodes, Johnson, & Haxby, 2011).

Il semble aussi y avoir des différences interindividuelles dans certaines capacités liées directement à la perception de visages. Certains travaux en neurologie ont démontré qu'on n'avait pas tous les mêmes capacités pour la mémorisation des visages ou la reconnaissance des expressions. La reconnaissance faciale est corrélée à notre capacité à coder, stocker et récupérer des représentations de visage dans notre mémoire (Rapcsak, 2003). En outre, les observateurs peuvent être influencés par des expériences perceptives et leurs capacités cognitives. Par exemple Fan et al. (Fan et al., 2012) suggèrent que la perception des visages virtuels dépend des expériences précédentes avec les médias virtuels.

2.5 Limites de l'existant & conclusions

Nous avons exploré différentes disciplines de recherche qui abordent la notion de fidélité et de reconnaissance de visage. Nous pouvons en conclure que la fidélité ne possède pas de définition unique et unanime. Nous n'avons pas non plus rencontré d'études qui se sont intéressées à formaliser une définition pour la fidélité d'un visage.

Malgré le fait que plusieurs aspects et variables ont été étudiés par les psychologues lors du processus de l'identification d'un visage, ces derniers n'ont jamais été formalisés en un cadre pour l'étude de la fidélité d'une identité par rapport à un support de comparaison.

Nous n'avons pas non plus rencontré d'études abordant la fidélité dans le cadre de visage généré par ordinateur. Les études abordant l'identification de visages ont le plus souvent été menées sur la base de photos de vraies personnes. Nous n'avons rencontré aucune étude qui s'est penchée sur la définition des différents éléments sur lesquels nous nous basons pour évaluer la ressemblance d'un visage virtuel avec son référent humain.

Cadre Conceptuel pour l'étude de la fidélité

Sommaire

3.1	Introduction	26
3.2	Principales dimensions de la fidélité	26
3.3	Qu'est ce qu'on évalue ?	27
3.3.1	Fidélité visuelle	28
3.3.2	Fidélité comportementale	28
3.3.3	Synthèse	29
3.4	Les entités impliquées lors du jugement de la fidélité	30
3.4.1	Qui devrait être le référent ?	30
3.4.2	Qui devrait être l'évaluateur humain ?	31
3.4.3	Comment devrait être le clone virtuel	31
3.4.4	Synthèse	31
3.5	Quelles sont les variables qui influencent le jugement de la fidélité ?	32
3.5.1	La relation entre les différentes entités	32
3.6	Synthèse	33

3.1 Introduction

Nous avons vu que la fidélité possède de multiples définitions qui varient considérablement selon l'axe de recherche abordé. La fidélité comprend plusieurs dimensions et aborde plusieurs paramètres. Elle dépend en effet d'un nombre important de variables. Ces variables peuvent être externes dû aux variations de l'environnement d'observation, mais aussi intrinsèques à l'observateur lui-même. À notre connaissance, aucun cadre rassemblant les différentes dimensions de la fidélité pour un visage virtuel n'a été défini.

Dans ce chapitre nous analysons les principales composantes et proposons un cadre conceptuel pour l'étude de la fidélité d'un clone virtuel.

Nous commençons par présenter ci-dessous une vue d'ensemble du cadre conceptuel que nous avons défini et sur lequel nous allons nous baser ultérieurement pour l'étude de la fidélité d'un clone virtuel (Figure 3.1).

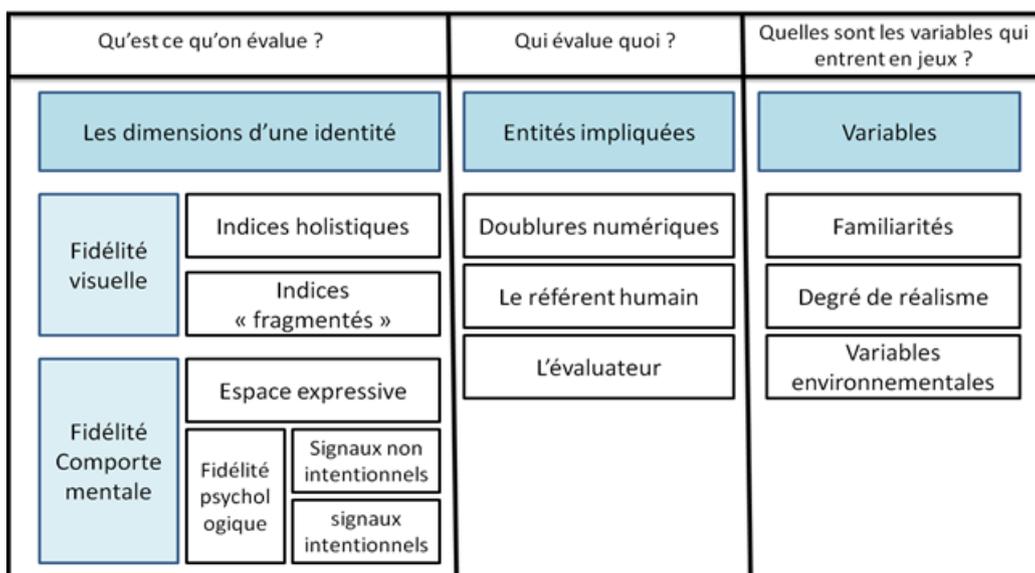


FIGURE 3.1: Cadre conceptuel pour l'étude de la fidélité

3.2 Principales dimensions de la fidélité

Comme suggéré par la définition de la fidélité que nous avons rencontrée dans le domaine de la réalité virtuelle (Allen et al., 1986), il est important de considérer la fidélité comme une notion qui possède plusieurs dimensions.

3.3 Qu'est ce qu'on évalue ?

Notre réflexion a eu pour point de départ une question simple : "Quels concepts sommes-nous amenés à évaluer face à un clone virtuel ?"

L'état de l'art ne répond pas directement à cette question. À notre connaissance, aucun cadre rassemblant les différentes caractéristiques sur lesquelles nous nous basons pour évaluer un visage virtuel n'a été défini.

Nous avons vu que les sociologues abordent la fidélité d'un avatar comme étant la fidélité de son identité (Goffman, 2002). Par conséquent, nous définissons comme premier grand bloc de notre cadre conceptuel "**les dimensions de l'identité**". Cependant quel peuvent être les composantes de cette notion d'identité ?

Selon des études socio-psychologiques (Goffman, 2002 ; Gobbini & Haxby, 2007) nous nous basons sur des éléments visuels, mais aussi des éléments comportementaux pour évaluer l'identité de quelqu'un. Cette conclusion est corroborée par des études en réalité virtuelle (Allen et al., 1986) qui définissent les dimensions physique et psychologique comme composante principale de la fidélité du simulateur.

Parmi les différentes dimensions introduites par la recherche en réalité virtuelle, deux dimensions ont particulièrement retenu notre attention : la dimension "Visuelle" et la dimension "Psychologique". Ces deux dimensions nous ont semblé les plus pertinentes pour l'étude d'une identité d'un clone virtuel.

Comme l'illustre la figure 3.2 le premier bloc de notre cadre expérimental répond à la question : quel élément évaluons-nous ? Il est composé de deux sous dimensions de la fidélité, une dimension visuelle et une dimension comportementale.

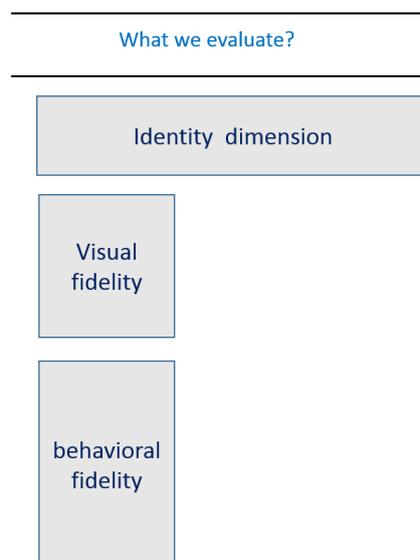


FIGURE 3.2: Cadre conceptuel pour l'étude de la fidélité : les dimensions de l'identité

3.3.1 Fidélité visuelle

Pour étudier plus en profondeur la fidélité visuelle, nous nous sommes intéressés à la façon avec laquelle les gens perçoivent les visages. En effet, appréhender la manière avec laquelle les gens se reconnaissent mutuellement peut contribuer à orienter les efforts visant à développer des avatars plus fidèles qui représentent une vraie personne. Nous avons vu que la perception du visage s'avère être un processus qui est à la fois holistique et configurable (évaluation trait par trait).

3.3.1.1 Élément holistique

Le processus holistique appliqué à un contexte de visages virtuels inclut principalement l'évaluation de deux facteurs globaux (Calder et al., 2011) :

- La structure géométrique du visage : incluant entre autres la forme générale du visage, la distance qui sépare les deux yeux, la longueur du nez par rapport au visage, etc.
- La pigmentation : Par pigmentation, nous faisons référence à toutes les propriétés de réflectance de surface de la peau, y compris l'albédo, la teinte, la spécularité, la translucidité et la variation spatiale de ces propriétés.

3.3.1.2 Les éléments les plus importants lors d'une évaluation trait par trait

L'évaluation des traits du visage semble avoir autant d'importance que l'évaluation holistique (Russell et al., 2007). La littérature nous a renseignés sur les éléments les plus importants sur lesquelles notre attention se focalise lors de l'identification. Ces éléments devront, par conséquent, jouir d'une attention particulière lors de la modélisation et l'évaluation d'un clone virtuel.

3.3.2 Fidélité comportementale

Il a été démontré grâce à des études neurologiques (Gobbini & Haxby, 2007) que nous basons aussi sur des données comportementales pour l'identification d'un visage familier.

3.3.2.1 Le rôle de l'expressivité

Des recherches telles que (Knappmeyer et al., 2003) ont démontré que les visages en mouvement sont plus aisés à reconnaître que des visages statiques, cela prouve que certains visages familiers ont des modèles de mouvements caractéristiques, qui agissent comme un signal supplémentaire à l'identité. Nous référons à cette composante dynamique de l'identité comme étant l'espace expressive. Nous définissons un espace expressif comme étant l'ensemble des caractéristiques des mouvements faciaux d'un être humain. Cette espace comprend à la fois les déformations physiques de cet individu (sa manière d'activer les différents muscles du visage), mais aussi sa manière d'exprimer différentes émotions

en plus des mouvements idiosyncrasiques qui lui sont propres. Les mouvements faciaux fournissent donc des informations importantes pour l'identification et la reconnaissance des visages. L'expressivité serait alors une des principales variables de la fidélité.

3.3.2.2 Fidélité psychologique

L'évaluation d'une identité englobe également le trainement des connaissances que nous avons sur le référent (par exemple sa manière particulière d'évaluer une situation émotionnelle). Nous évaluons par conséquent une identité à travers des comportements traduisant une évaluation psychologique de l'environnement . Au sein de ces comportements (qu'ils soient verbaux ou non verbaux) Goffman (Goffman, 2002) distingue :

- Des signaux intentionnels sur lesquels nous exerçons un contrôle . C'est l'image que nous aimerions donner à notre interlocuteur.
- Des signaux non intentionnels correspondant à des actions réactives qui fuient sans intention ni contrôle de notre part.

La fidélité psychologique est fortement liée à la crédibilité comportementale (Guadagno, Blascovich, Bailenson, & Mccall, 2007). Elle évalue dans quelle mesure les utilisateurs interagissent avec l'agent en ayant l'impression qu'ils observent un être pensant qui a ses propres croyances, des désirs et une véritable personnalité. C'est l'évaluation de la similarité ressentie entre des comportements humains et ceux du clone.

3.3.3 Synthèse

Le premier bloc de notre cadre conceptuel pour l'évaluation de la fidélité est composé de deux sous dimensions (figure 3.3).

- La dimension visuelle qui inclut les deux processus d'évaluation du visage
- La fidélité comportementale qui inclut :
 - un espace expressif qui englobe les idiosyncrasies et expressions particulières du visage
 - les processus psychologiques qui pilotent notre expressivité.

Toutes ces variables doivent être prises en compte lors de l'évaluation d'un visage de clone virtuel.

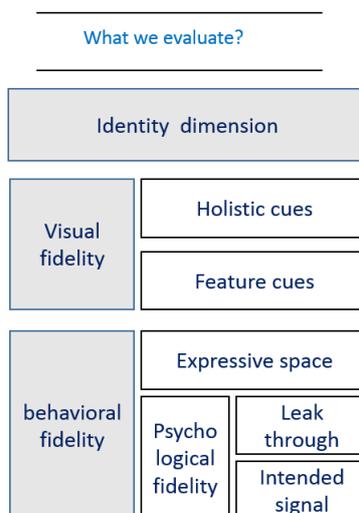


FIGURE 3.3: Cadre conceptuel pour l'étude de la fidélité : Fidélité psychologique

3.4 Les entités impliquées lors du jugement de la fidélité

Alors que le réalisme peut être évalué avec tout visage virtuel (dépourvu d'une identité), le concept de fidélité ne peut être évalué que si l'observateur connaît le visage du référent ou du moins possède un support de comparaison lui présentant le référent réel (photos/vidéo). Des études perceptives sur le jugement de la fidélité apportent donc trois nouvelles questions importantes :

- Qui devons-nous choisir pour être le référent humain ?
- Qui devrait jouer le rôle d'observateur pour juger la fidélité d'un clone ?
- Quel clone virtuel devons-nous déployer ?

3.4.1 Qui devrait être le référent ?

Comme suggéré par les études psychologiques, l'évaluation de la fidélité d'un clone virtuel est fortement liée à la familiarité. La notoriété du référent et sa relation avec l'observateur ont un fort impact sur le jugement de la fidélité.

La perception des visages diffère aussi selon les caractères distinctifs d'un visage. Par exemple, dans une étude menée par (Lander & Chuang, 2005), la reconnaissance d'une identité sur la base de son mouvement est plus importante lorsque le visage observé a une expressivité distinctive. Par conséquent, notre cadre expérimental établit une distinction entre la reconnaissance des visages inconnus et les visages familiers. Les études de psychologie nous suggèrent également de faire varier le choix du référent. Par conséquent pour bien appréhender les processus d'identification de visages virtuels il nous faut faire varier les caractéristiques propres au référent, choisir en même temps des personnes ayant des éléments distinctifs (par exemple des cicatrices, des taches de naissances), ayant une expressivité bien particulière, voire caricaturale. D'autres référents devraient être plus neutres

représentant un visage moyen avec une expressivité plus commune.

3.4.2 Qui devrait être l'évaluateur humain ?

La reconnaissance faciale peut être altérée par les spécificités de l'observateur. En effet plusieurs différences interindividuelles telles que l'origine ethnique, la catégorie sociale, les styles cognitifs et l'expérience préalable avec les supports virtuels (jeux vidéo , film d'animation) jouent un rôle très important lors d'un processus d'évaluation de la ressemblance et l'identification d'un visage familier. En conséquence, nous devons considérer un échantillon représentatif d'évaluateur couvrant ces différences interindividuelles.

Autre perspective intéressante est "la perception de soi". Les auteurs dans (Cook, Johnston, & Heyes, 2012) ont observé que les individus ont la capacité de mieux reconnaître leurs propres expressions faciales que des personnes qui était familières avec ces derniers. Pourtant instinctivement nous sommes portés à croire que nous avons très peu de recul sur nos expressions faciales. En effet, nous avons peu d'occasions où nous pouvons percevoir (visuellement) nos propres expressions (Knappmeyer et al., 2003). Pourtant, des études ont démontré que grâce à des retours moteurs nous avons la capacité de mieux reconnaître nos propres expressions. Cela nous amène à envisager la possibilité de prendre en considération le cas où l'observateur et lui-même le référent.

3.4.3 Comment devrait être le clone virtuel

Comme nous l'avons vu précédemment dans notre état de l'art, la perception des visages est soumise à des variables extrinsèques. Nous devons donc tenir compte de ces variables dans notre plate-forme d'expérimentation. Nous devons contrôler les conditions d'illuminations, l'angle de vision ainsi que la pose et la résolution de l'image du clone virtuel.

De plus, plusieurs directions expérimentales pertinentes peuvent être explorées : comment des données dégradées ou modifiées peuvent-elles véhiculer une identité (par exemple les caricatures) ? Quelles techniques de rendu peuvent elles affecter la perception de l'identité ? À quel point les conditions de perceptions affectent-elles le jugement de fidélité (l'angle de vue, la distance) ? Est-ce qu'il y'a des techniques de rendus plus adéquats selon les conditions de vue ?

3.4.4 Synthèse

Le deuxième bloc qui définit notre cadre conceptuel présente les différentes entités qui entrent en jeu lors de l'évaluation d'un clone virtuel (Figure 3.4). Ce dernier nous renseigne à la fois sur l'échantillon de sujets que nous devons prendre en compte pour une évaluation perceptive d'un clone, sur le choix du référent et sur la manière avec laquelle nous pouvons concevoir le clone virtuel.

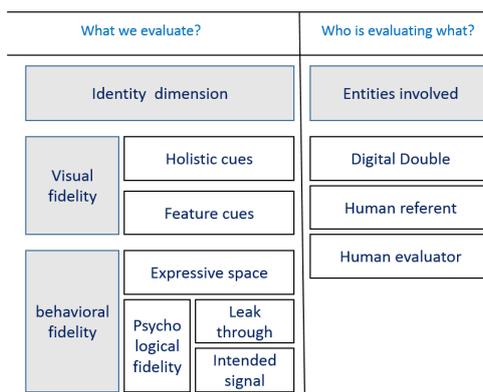


FIGURE 3.4: Cadre conceptuel pour l'étude de la fidélité : les entités impliquées

3.5 Quelles sont les variables qui influencent le jugement de la fidélité ?

Nous avons vu que la perception de visage humain est fortement influencée par de nombreuses variables. Une des plus importantes est la familiarité qu'entretient l'observateur avec le référent (Sinha et al., 2006). En effet, il a été observé qu'il est beaucoup plus aisé de reconnaître un visage familier même si ce dernier est altéré. La familiarité influence fortement le degré de tolérance à la dégradation des stimuli lors d'une tâche d'identification. Cette variable doit donc être prise en compte lors d'une étude perceptive. Nous avons également vu que des rendus non réalistes (p. ex., les caricatures) sont capables de véhiculer une identité avec des taux de reconnaissance équivalents aux photos (Stevenage, 1995). Des études sur "l'uncanny valley" (Mori, 1970) ont aussi démontré qu'un niveau de rendu graphique qui tend vers la photo réaliste sans arriver à l'atteindre suscitait un sentiment de rejet de la part de l'utilisateur. Dans ces circonstances le niveau de réalisme du clone virtuel doit être pris en compte lors de la modélisation et de l'évaluation d'un clone virtuel. Finalement, nous avons vu que la perception de visage est influencée par des variables dites extrinsèques. Ces variables sont dépendantes de l'environnement et des conditions de vue.

3.5.1 La relation entre les différentes entités

Comme illustre dans la figure 3.5 la relation entre les différentes entités est fortement lié aux variables qui influence le jugement de la fidélité.

Par exemple en ce qui concerne le lien entre le référent et l'évaluateur, des études psychologiques tel que (Lander & Chuang, 2005 ; Wallraven et al., 2005) suggèrent que l'évaluation de la fidélité peut dépendre de :

- La familiarité entre le référent et l'évaluateur.
- les signes distinctifs que perçoit l'évaluateur sur le référent.
- Le fait que le référent et évaluateur appartienne au même groupe social.

Les conditions d'environnement et de visualisation entre le clone virtuel et l'évaluateur jouent également un rôle important dans l'évaluation de la fidélité.

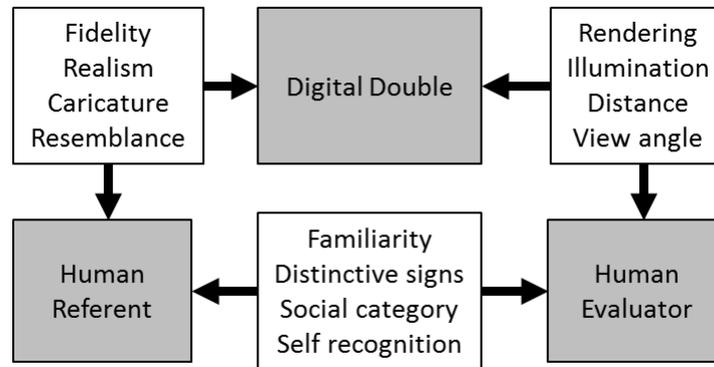


FIGURE 3.5: La relation entre les différentes entités

Finalement, le clone virtuel peut être représenté avec divers degrés de réalisme qui pourraient influencer le jugement de la fidélité.

3.6 Synthèse

Dans cette section nous avons présenté un cadre conceptuel et des variables expérimentales pour l'étude de la fidélité. À travers la définition de ce cadre, nous avons proposé une définition de la fidélité adéquate à l'étude de visages virtuels en 3D et nous avons procédé à l'identification de ses principales dimensions. Nous avons vu que l'évaluation de la fidélité peut être modulée par plusieurs variables intrinsèques aux participants ou dépendantes de l'environnement. Nous définissons, par conséquent, le cadre conceptuel pour l'étude de la fidélité, comme une composition de trois principaux facteurs :

1. Dimensions de l'identité
2. Les différentes entités impliquées dans le processus de jugement
3. Un ensemble de variables externes qui impactent le processus de jugement

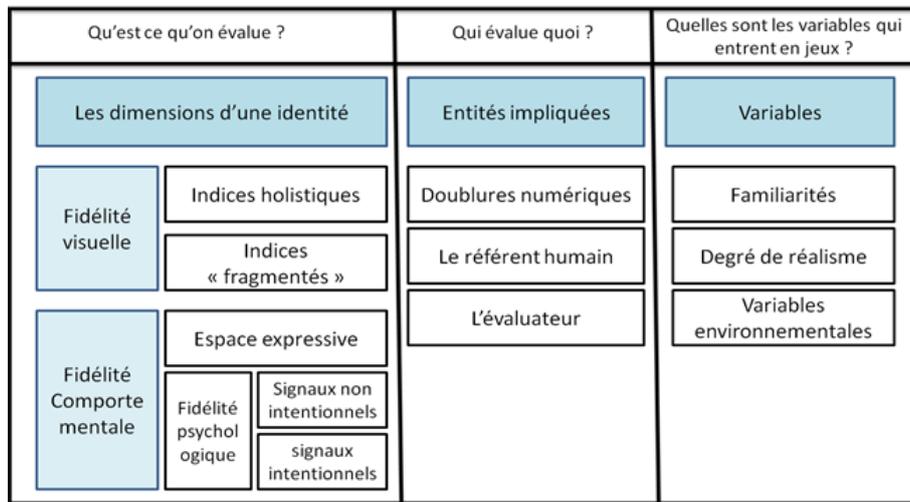


FIGURE 3.6: Cadre conceptuel pour l'étude de la fidélité

Les futures études expérimentales perceptives devront prendre en compte la complexité de ce cadre pour mettre en place des protocoles d'expérimentation qui limiteront les biais lors du jugement de la fidélité d'un clone virtuel avec son référent.

Nous retenons également que la fidélité d'une identité est véhiculée via des indices visuel et comportemental. Nous allons nous baser sur ces résultats afin d'étudier dans la prochaine partie comment modéliser un clone virtuel qui tiendrait compte des éléments de ce paradigme.

Dans le chapitre suivant, nous considérons les solutions offertes par l'informatique graphique vis-à-vis des différentes composantes de la fidélité.

Deuxième partie

Capture, rendu et animation d'un clone virtuel

État de l'art : Conception d'humains virtuels et techniques de numérisation

Sommaire

4.1	Introduction	38
4.2	Création du maillage	38
4.2.1	Technologies de numérisation tridimensionnelle	39
4.2.2	Application des techniques de scan 3D aux visages humains	44
4.3	Rendu	45
4.3.1	Techniques de rendu	46
4.3.2	Rendu de la peau	47
4.3.3	Rendu des yeux	48
4.4	Animation	49
4.4.1	Encodage des expressions faciales	49
4.4.2	Technique d'animation	51
4.5	Limites de la modélisation d'humains virtuels	54
4.5.1	Uncanny valley	54
4.5.2	Interactivité et cout matériel	55

4.1 Introduction

La simulation de visage virtuel représente un axe central des travaux en informatique graphique. Depuis le tout premier visage virtuel qui a été présenté dans les travaux de Parke (Parke, 1972) la recherche portant sur la conception de visage virtuel a connu une évolution spectaculaire comme l'illustre la figure 4.1. Pourtant, il n'en demeure pas moins un problème ouvert qui suscite encore plusieurs questions de recherches.

Ces questions s'articulent autour de deux axes principaux :

- La première problématique concerne le rendu statique du visage virtuel. Cette étape comprend la création du maillage, des textures et les différentes techniques de rendu pour simuler la peau, les yeux et les différents éléments du visage.
- La seconde problématique porte sur les techniques d'animation qui vont guider l'évolution du maillage au cours du temps.



Le visage de synthèse animé de Parke (F. I. Parke, 1974)

Digital ira and beyond: creating real-time photoreal digital actors (Von et al., 2014)

FIGURE 4.1: Évolution de rendu pour les visages virtuels

4.2 Création du maillage

La première étape pour la création d'un visage virtuel est de générer un maillage 3D représentant sa géométrie. Ce maillage peut être entièrement synthétique, issu de l'imagination et du travail manuel d'un graphiste ou bien il peut être obtenu à partir d'un visage réel numérisé. De plus, pour que le maillage puisse être animé en temps réel, la modélisation d'un visage en trois dimensions doit tenir compte de certaines contraintes par exemple une topologie particulière pour le maillage.

Les progrès récents dans la technologie de numérisation ont permis la création rapide de personnages en 3D à partir de sujets humains. La fidélité de ces modèles 3D statiques peut varier en fonction de l'équipement de numérisation, de l'éclairage et des procédés et des

processus manuels utilisés pour affiner le modèle généré.

La construction d'un modèle 3D traditionnel d'un personnage statique par des moyens manuels peut prendre des jours, des semaines ou même plus. Le processus de numérisation permet la production rapide de contenu 3D qui reste indépendant des capacités artistiques d'un designer graphique.

4.2.1 Technologies de numérisation tridimensionnelle

Plusieurs approches ont été explorées selon les propriétés des objets à numériser (statique, dynamique, transparent, etc.) (Amor, Ardabilian, & Chen, 2005). La figure 4.2 donne une classification des techniques existantes de numérisation tridimensionnelle. Elle distingue entre les méthodes avec contact et des méthodes sans contact (transmissive et réfléchive).

Dans cette section, nous restreindrons notre étude aux méthodes dites optiques, qui est une sous-catégorie des méthodes sans contact réfléchives. Elles nous semblent être la plus adaptées à nos besoins en terme de reconstruction faciale 3D. Les méthodes de cette catégorie peuvent être classées elles même en deux sous-catégories qui sont : les méthodes actives et les méthodes passives.

Les deux sections qui suivent dévoileront les techniques développées au sein de ces deux sous-catégories précisant les avantages et les limites de chacune d'elles. Nous abordons également l'utilisation de ces approches dans le cadre de scan de visages.

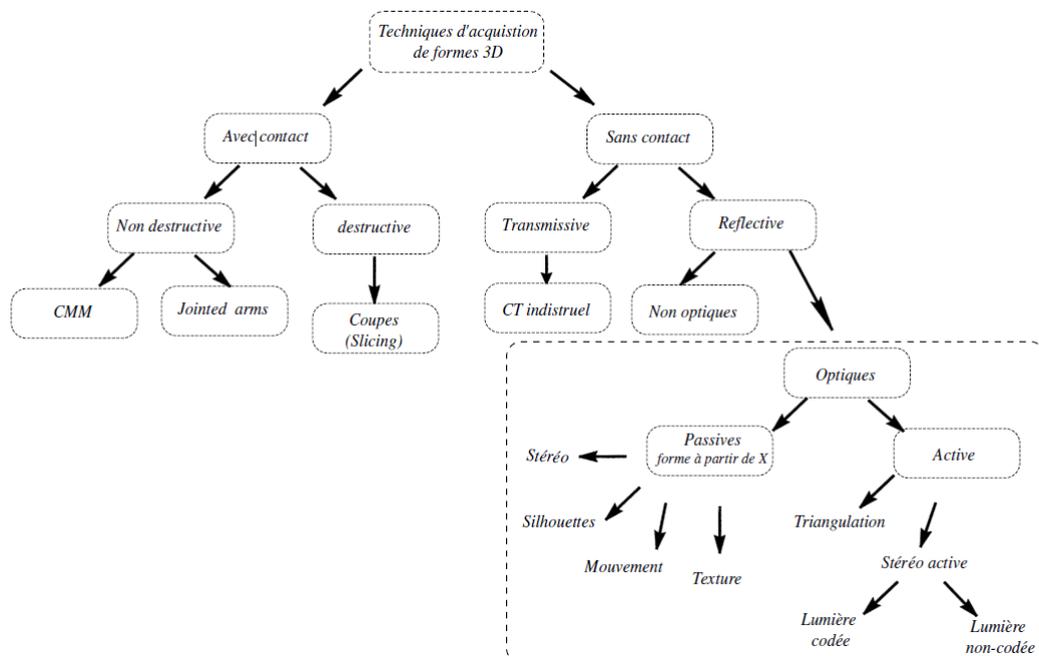


FIGURE 4.2: Taxonomie des techniques d'acquisition de formes 3D (Amor et al., 2011)

4.2.1.1 Techniques de reconstruction active

La vision active consiste à utiliser une source lumineuse avec plusieurs capteurs photographiques dans le but de mesurer les coordonnées 3D des points sur la surface d'un objet (Amor et al., 2005). Dans cette catégorie de technique, deux types de sources lumineuses sont généralement employées : le laser (Blason et al. 2004) ou un schéma de lumière structuré (O. Alexander et al., 2010).

Le laser. Le laser est une lumière monochromatique, qui en traversant une lentille cylindrique, crée un plan lumineux éclairant une seule tranche de la scène. Ceci permet de n'observer, par une caméra associée, que la partie de la scène ainsi éclairée (figure—4.3¹)



FIGURE 4.3: Exemple de système à base de triangulation laser¹

L'équation du plan lumineux projeté sur la scène nous permet de calculer les coordonnées de ce plan dans l'image captée par la caméra associée au dispositif laser, nous pouvons alors calculer la position de la tranche éclairée dans l'espace (Amor et al., 2005) . Ce principe est appelé triangulation active. Le point laser, la caméra associée et l'émetteur laser forment un triangle.

Le laser est pourvu d'un miroir rotatif associé aux émetteurs de faisceau laser qui effectuent un balayage en vue de mesurer les coordonnées 3D de la totalité de l'objet scanné. Cette technique présente de bonnes performances d'acquisition en terme de qualité de maillage reconstruit et de précision de mesures, mais souffre de quelques lacunes. Ces faiblesses sont dues à la discontinuité de la réflectance sur la surface de l'objet, la variation de la forme (sa pose) et les occlusions qui peuvent empêcher le trait laser d'arriver à la surface de l'objet ou par simplement l'absorption du trait laser par la surface de l'objet à numériser. Ces équipements souffrent également de leurs prix très élevés.

Cette méthode a été utilisée depuis déjà plusieurs années pour l'acquisition de la géométrie du corps humain. En effet, une base d'acquisition 3D de corps humain appelée "CAESAR

1. <http://www.rob.cs.tu-bs.de/en/news/david>

anthropometric database” a été développée par Robinette et al. Elle contient plus de 2000 corps d’hommes et de femmes âgés de 18 à 65 ans. Elle a été créée en vue de mesurer les variabilités anthropométriques entre la population d’Europe et celle d’Amérique du Nord (Robinette, Daanen, & Paquet, 1999).

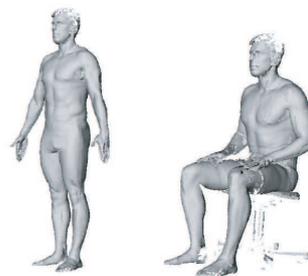


FIGURE 4.4: Extrait de la base de corps humains 3D CAESAR anthropometric database (Robinette, 1999)

Le même type de capteur, à base de triangulation laser, a été utilisé pour capturer des visages humains. Par exemple, dans (Liu, Zhang, Cohen, & Jacobs, 2004) les auteurs ont collecté un corpus de visages 3D d’une centaine de personnes acquises moyennant le scanner laser Cyberware².

Ces appareils spécialisés produisent une numérisation de bonne précision, mais ils nécessitent une étape de post-traitement (intervention manuelle ou automatique) pour éliminer le bruit qui se manifeste par des pics sur le maillage. Ces phénomènes sont présents surtout dans les zones chevelues du visage (sourcils, barbes, moustache) et les régions des yeux où le faisceau laser est très peu réfléchi comme l’illustre la figure 4.5. Les travaux de Liu et al. montrent les problèmes d’occlusion, de bruit d’acquisition et d’absorption du trait laser au niveau de quelques régions du visage qui se traduit par des trous sur le maillage obtenu.

(a) Canal de forme
(range image) (b) Canaux de forme
et de texture



IV2: appareil Minolta VI900/910

FIGURE 4.5: Extrait de la base de visage (Liu et al., 2004)

2. <http://cyberware.com/>

Lumière structurée La seconde technique de méthode active se base sur un motif de lumière structuré qui substitue le faisceau laser. Ces méthodes projettent, sur l'objet à scanner, un motif structurant qui est souvent une grille ou des bandes de lumière parallèles (von der Pahlen et al., 2014).

Le maillage de l'objet scanné est obtenu en analysant les déformations du motif lumineux par rapport au motif d'origine. Cette méthode estime le maillage 3D par décodage du patron distordu sur l'objet. Par exemple, dans le cas de la projection d'un motif simple sur la scène (trait noir sur un fond blanc), il est relativement simple de calculer les coordonnées 3D à partir de la déformation du trait sur l'objet. Une technique similaire à la triangulation est utilisée pour calculer la distance de déformation, et donc la position 3D des points la représentant.

Pour couvrir la totalité de l'objet à scanner, il suffit de réitérer la procédure en modifiant à chaque fois la position du trait dans le motif (Posdamer & Altschuler, 1982 ; Amor et al., 2005).

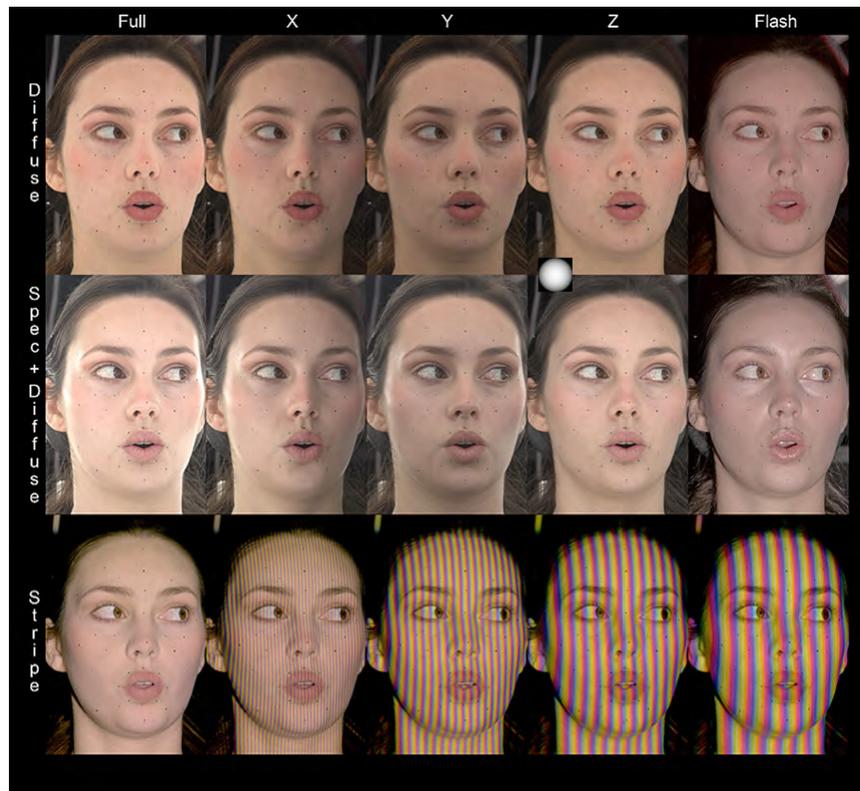


FIGURE 4.6: Exemple de système à base de lumière structurée (Von der Pahlen et al., 2014)

La deuxième alternative consiste à concevoir un motif plus complexe qui peut être une répétition du trait sur l'axe X ou/et Y formant ainsi une grille (O. Alexander et al., 2013) (Figure-4.6). Dans ce cas, la technique de reconstruction consiste à mettre en correspondance les primitives du motif d'origine avec les déformations du motif lumineux sur l'objet. Un spectre très large de travaux basés sur la projection de motifs lumineux a

été élaboré depuis les années 80 jusqu'à présent. Chaque approche propose un schéma de lumière couplé à une méthode de mise en correspondance en vue de reconstruire un objet ou une scène donnée (Zhang, 2010).

4.2.1.2 Techniques de reconstruction passives

Les techniques de reconstruction passives récupèrent l'information 3D en utilisant des images sans aucune assistance d'une source de lumière. La vision passive, contrairement à la vision active, n'utilise aucune structuration particulière de la lumière.

Cette approche reconstruit un maillage 3D de visage à partir d'une ou plusieurs images d'intensité. Les scanners passifs se basent sur des appareils photo simples qui ne dégagent aucun type de rayonnement eux-mêmes, mais reposent sur la détection de la lumière visible. Les seules données disponibles sont une ou plusieurs images prises à partir de plusieurs angles de vue (par une ou plusieurs caméras). Ce procédé également appelé "photogrammie" permet grâce à des filtres de capturer également des textures de propriété de la peau que nous allons voir plus amplement dans la section suivante.

Un exemple de configuration pour la photogrammie est illustré sur la figure 4.7³.



FIGURE 4.7: Une set-up de photogrammétrie de qualité professionnelle (infinite Realities³)

Il existe plusieurs approches pour modéliser des maillages 3D à partir de photos. Cette technique est également dénommée "shape from X". Certaines méthodes peuvent se baser

3. <http://ir-ltd.net/>

que sur une seule image, à condition de disposer d'information supplémentaire telle que des informations d'ombre ou de réflectances. Parmi les méthodes les plus utilisées, il y a "shape from shading" qui se base sur les propriétés de réflectance, "shape from texture" qui se base sur la texture des surfaces, et "shape from motion" qui déduit la forme à partir du mouvement entre plusieurs frames.

Par exemple, la méthode se basant sur les informations d'ombrage, parvient à estimer la forme de l'objet en se basant sur les variations d'intensité dans l'image. L'extraction de l'information de profondeur en chaque point 3D de l'objet est assurée en analysant la brillance de la scène en ce point. Cette profondeur est estimée sur la base de l'intensité de son niveau de gris (Horn et al., 1989). Toutes ces méthodes nécessitent par contre un temps de calcul assez important.

4.2.2 Application des techniques de scan 3D aux visages humains

Diverses approches de captures dédiées au visage humain ont été développées au cours des dernières années. La plupart des approches citées au paravent ont été appliquées pour la captation de visages virtuels. Cependant nous avons remarqué que les dispositifs professionnels les plus récents se basent majoritairement sur l'approche des rayons à lumière structurée et la photogrammétrie (Zhang, 2010).

L'un des travaux leader dans ce domaine sont ceux de la "LightStage" développée par l'université de Californie du Sud⁴(Debevec, 2012).

Debevec et al. (Debevec et al., 2000) ont conçu l'un des premiers systèmes de capture dédiés au visage humain. La première version du système LightStage est composée de plusieurs caméras et sources lumineuses montées sur un axe rotatif permettant de capturer le visage selon plusieurs angles.

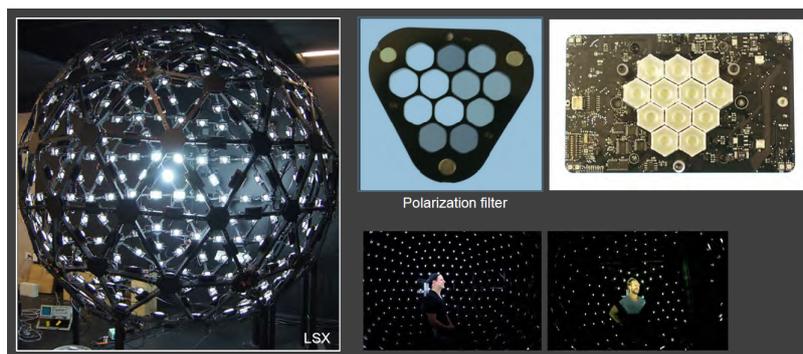


FIGURE 4.8: LightStage Version 5 (Von der Pahlen et al., 2014)

Ce système a ensuite été développé en le dotant d'une structure rigide (Figure 4.8). Le dispositif de la LightStage utilise de la lumière structurée, et permet ainsi de capturer la géométrie faciale du sujet et de mesurer les propriétés de la pigmentation de la peau.

4. <http://gl.ict.usc.edu/LightStages/>

Cette technologie (O. Alexander et al., 2010), est aujourd’hui utilisée dans de nombreux films pour créer des acteurs numériques (Avatar de James Cameron, Benjamin Button ou encore Superman returns).

Ces données sont généralement acquises depuis plusieurs angles de vue, et en considérant plusieurs angles d’incidence de la lumière. L’ensemble de captures lumineuses représente donc une grande quantité de données.

4.3 Rendu

La qualité des données associées au maillage (textures, spéculométrie) est essentielle pour obtenir un rendu de qualité. Grâce à des appareils de mesures spécifiques permettant d’extraire le comportement lumineux de la peau, il est possible aujourd’hui d’acquérir les données spécifiques à chaque région du visage, et ainsi de reproduire le comportement de la lumière de manière fidèle sur l’ensemble du visage (Courgeon, 2011).

Les techniques actuelles considèrent généralement plusieurs types de propriétés lumineuses. Le système LightStage, considère quatre types de comportements lumineux : La lumière réfléchiée, la lumière diffuse directe, la lumière diffusée en surface de la peau, et la lumière diffusée par les couches internes de la peau.

Plusieurs jeux de données sont donc capturés, pour être ensuite recombinaés par le système d’illumination lors du rendu final. La Figure 4.9 montre les différentes composantes lumineuses et leur combinaison, comparée à la photographie du modèle humain.

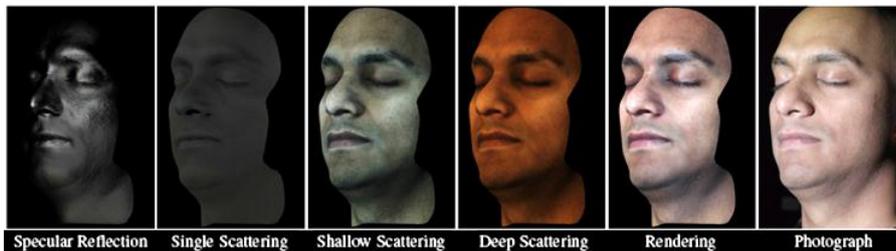


FIGURE 4.9: Composantes lumineuses séparées, combinées, et photo de référence (Ma et al., 2008)

Avec l’augmentation des performances des cartes graphiques modernes et surtout l’émergence de processeurs graphiques programmable (Graphic Processing Unit ou GPU), le rendu d’humain virtuel a connu un nouvel élan permettant de simuler de façon réaliste l’interaction du visage avec les conditions d’illumination.

La qualité d’une image de synthèse repose sur 2 aspects fondamentaux :

1. La qualité des données qui consiste elle même en deux éléments :
 - la définition de la structure géométrique (le maillage). Le maillage est une structure géométrique constituée de vertex reliés ensemble pour former des surfaces.
 - la définition des textures (images correspondant à la couleur de l’objet et ses propriétés lumineuses)

2. La qualité du système d'illumination

4.3.1 Techniques de rendu

Nous pouvons faire la distinction entre les méthodes de rendu temps réel et le rendu pré-calculé (ou non temps-réel), dont les objectifs sont différents. Il existe des algorithmes de rendus pour l'informatique graphique dite pré-calculé nécessitant un temps de calcul très important (plusieurs heures de calcul pour rendre une seconde d'un Film Pixar par exemple). Ces méthodes tel que le projeté de rayon (ray tracing) sont destinés à l'industrie du cinéma et permettent aujourd'hui un réalisme visuel très élevé (figure 4.10⁵). Néanmoins, ces méthodes ne permettent pas l'interaction avec la scène graphique générée. Ces techniques ne sont donc pas pertinentes pour la création d'interfaces homme-machine où l'interaction temps réel est de mise.

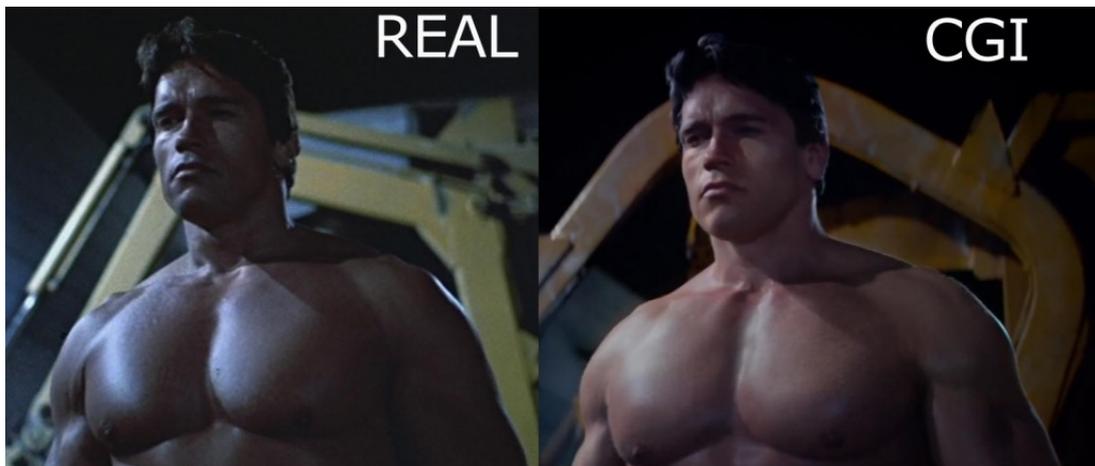


FIGURE 4.10: Extrait du film Terminator Genisys (2014)⁴

En ce que concerne les techniques de rendu en temps réel, le rendu généré est forcément moins réaliste, mais il génère plus de 25 images par secondes. Les algorithmes de rendu temps réel (shaders) tirent partie des capacités matérielles des cartes graphiques (GPU). Il différents Shaders s'exécutent de manière séquentielle sous quatre principales étapes. Ils constituent le "Pipeline Graphique". Dans la partie vectorielle du pipeline, les shaders sont dédiés au traitement de la géométrie 3D. Le Vertex Shader est responsable des changements de référentiels. Il permet de situer les sommets dans l'espace-écran. Le Tessellation Shader, séparé en deux sous étapes, permet de raffiner localement un maillage, et ainsi de générer dynamiquement des détails. Le Geometry Shader permet lui de générer dynamiquement tout type de géométrie. Finalement, le Fragment Shader est dédié au traitement des couleurs, c'est-à-dire le remplissage des pixels (Nguyen, 2007 ; Courgeon, 2011).

5. <http://3danimationcgi.com/arnold-cgi-effects-t5/> (consulté 09/2015)

4.3.2 Rendu de la peau

En infographie, le rendu 3D ultra-réaliste de la peau est une étape assez délicate, la peau est une matière complexe à simuler : imperfections, alternances de surfaces plus ou moins translucides, etc.

Contrairement aux matériaux translucides simples comme le marbre, la peau humaine est composée de plusieurs couches qui définissent comment la lumière est absorbée et dispersée lors de l'interaction avec elle. Deux des couches les plus importantes sont l'épiderme et le derme, qui contiennent de la mélanine et l'hémoglobine chromophores qui définissent la couleur spécifique de la peau. La lumière s'y comporte donc de manière complexe. La couleur de la lumière sortant de la surface de la peau dépendra dans quelle mesure du déplacement de la lumière à l'intérieur de ces couches. Les rayons lumineux sont diffusés et réfléchis plusieurs fois à l'intérieur des couches de la peau et ne sont presque jamais réémis de l'endroit où ils sont entrés (Jimenez, 2012). Cela se traduit par des dégradés de couleur dans les frontières entre la lumière et les ombres. Une main rétro-éclairée est probablement l'un des meilleurs exemples où nous pouvons voir comment cet effet peut paraître sur la peau humaine (Figure-4.11).



FIGURE 4.11: Trajectoire de la lumière dans une surface partiellement transparente (Jimenez, 2012)

Pour la simulation de ce rendu en informatique Jensen et al. (Jensen et al., 2001) se sont rapprochés de la simulation du rayonnement sortant de matériaux translucides en utilisant une approximation des modèles de BSSRDF (Bidirectional Subsurface Scattering Reflectance Distribution Function), mais les performances étaient encore loin d'atteindre celles du temps réel. Par la suite dans les travaux d'Eon et al. (Eon et al., 2007), les auteurs ont proposé une technique où l'approximation de la diffusion peut être estimée par une somme de gaussiennes pour enfin atteindre des performances en temps réel. Par contre, leurs techniques nécessitaient des ressources matérielles importantes.

Jorge Jimenez (Jimenez, Sundstedt, & Gutierrez, 2009) a perfectionné une technique d'effets spéciaux à plusieurs passes baptisées "Screen-space separable subsurface scattering" (Figure-4.12), permettant à la lumière d'interagir de manière ultra-réaliste avec une matière aussi complexe que la peau humaine.

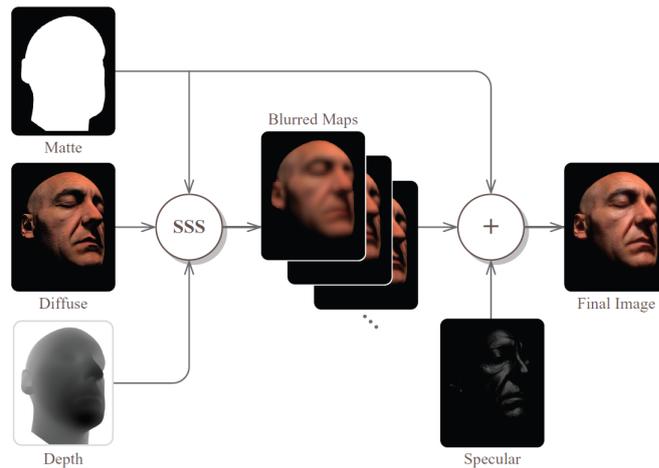


FIGURE 4.12: Vue d'ensemble de l'algorithme de rendu de peau SSSSS : Screen-Space Sub-Surface Scattering (Jimenez, 2009)

4.3.3 Rendu des yeux

Les yeux et leurs mouvements jouent un rôle dans la communication de l'information sociale et émotionnelle. Ils nous renseignent également sur l'endroit où notre attention se porte.

L'oeil est un organe complexe possédant plusieurs propriétés spécifiques de réflexion et réfraction de la lumière. (Argyle & Dean, 1965)

Les travaux les plus récents portant sur le rendu de l'oeil ont été réalisés par Jimenez et al. dans le cadre du projet "digital IRA" (von der Pahlen et al., 2014). Des shaders avancés simulant la réfraction et la caustique de l'oeil humain y ont été développés (la "caustique" désigne en optique l'enveloppe des rayons lumineux subissant une réflexion ou une réfraction sur une surface ou une courbe). Ils ont réalisé une série de shaders pour simuler toutes les propriétés lumineuses de l'oeil humain. La figure 4.13 illustre quelques résultats de leur technique : réflexions, humidité de l'oeil, Réflexion d'Occlusion, réfraction de la vue, la Réfraction de la lumière, L'Occlusion ambiante, Bump Mapping de l'iris, un rendu Limbus et la rougeur des yeux.

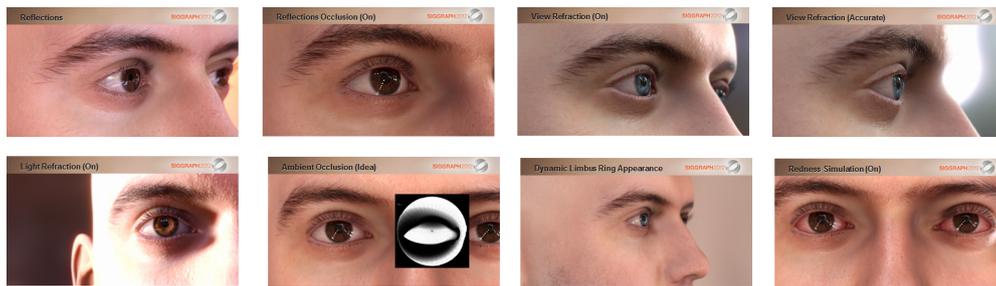


FIGURE 4.13: Ensemble des Shaders pour un rendu photo-réaliste des yeux (Jiminez, 2012)

4.4 Animation

Le visage humain est extrêmement complexe, car composé d'un nombre important de muscles qui interagissent avec une structure osseuse, provoquant ainsi d'innombrables rides sur la peau. Ekman s'est inspiré de la complexité des muscles des visages pour spécifier une base d'action unitaire (Figure-4.14⁶), dont la combinaison sert à décrire nos mouvements faciaux (Ekman & Friesen, 2003).

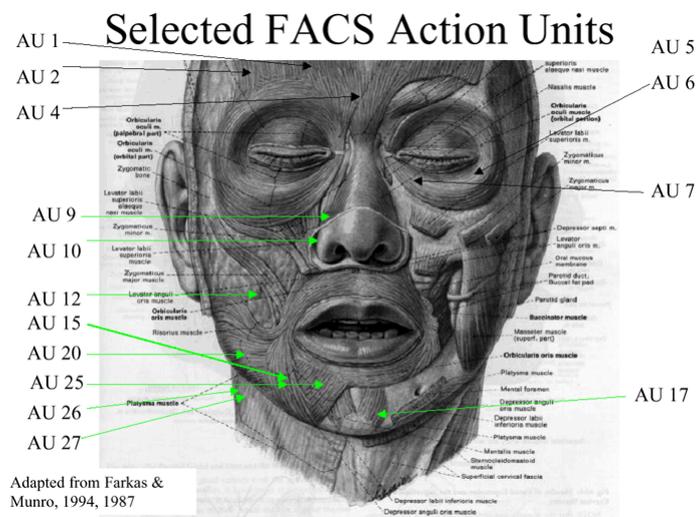


FIGURE 4.14: Ensemble des muscles faciaux (Frakas & Munro, 1987)

4.4.1 Encodage des expressions faciales

Dans cette section nous présentons des méthodes permettant de décrire une expression ou une animation faciale. Certaines sont issues de travaux dans le domaine de la psychologie (Ekman & Friesen, 1977, 2003) et n'ont pas pour vocation initiale d'être utilisées

6. http://www.wedernoch.de/mimik/facs_au.html

en informatique graphique. D'autres ont été proposées uniquement dans ce but tel que le système MPEG-4 (Pandzic & Forchheimer, 2002).

4.4.1.1 Facial Action Coding System (FACS)

Le FACS est un système de codage des expressions faciales développé par les psychologues P. Ekman et W. Friesen en 1978 (Ekman & Friesen, 1977). Il recense 46 changements musculaires unitaires de visage en plus de changement de direction de regard et quelques mouvements de tête. Les expressions faciales sont alors décrites par des combinaisons d'unités d'action et de leurs intensités. La figure 4.15 présente une partie des AU correspondants à la moitié supérieure du visage.

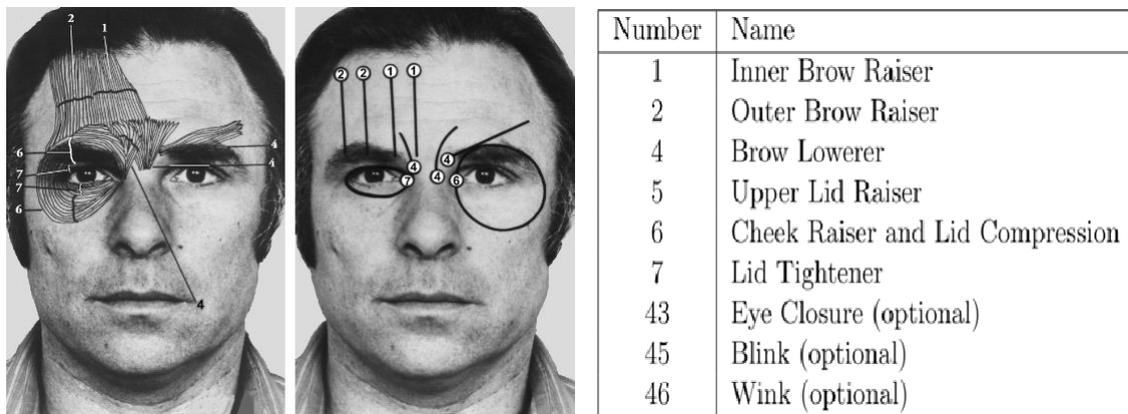


FIGURE 4.15: Les actions unités correspondant à la moitié supérieure du visage (Ekman & Friesen, 1977)

Cette indexation a été beaucoup utilisée pour des travaux sur la reconnaissance des expressions émotionnelles (Puklavage, Pirela, Gonzalez, & Georgiopoulos, 2010 ; Littlewort et al., 2011) mais également pour la génération d'expression faciale (Gunes, Schuller, Pantic, & Cowie, 2011 ; Ochs, Niewiadomski, Brunet, & Pelachaud, 2012).

4.4.1.2 Standard MPEG-4

La norme MPEG-4 contient une partie dédiée à l'animation du corps humain, dont le visage, notamment à travers un système de codage d'animation faciale (Pandzic & Forchheimer, 2002).

La norme MPEG-4 FBA (Face and Body Animation) est une composante de la norme MPEG-4 International Standard (ISO14496) qui s'intéresse à l'animation des personnages humains ou anthropomorphiques. Cette norme définit un ensemble de 84 paramètres qui définissent donc des expressions des visages, mais aussi les visèmes et des paramètres bas niveau pour déplacer la langue ou le cou par exemple. (Figure— 4.16). Cette technique a l'avantage de permettre une large gamme d'expressions faciales que nous pouvons obtenir par combinaisons de différents réglages des paramètres. Cela avec un coût de calcul

relativement faible. Par contre le jeu de paramètre dépendant de la topologie du maillage, un paramétrage générique est impossible. En outre le réglage manuel des différents poids relatifs à chaque point de la norme MPEG-4 est fastidieux.

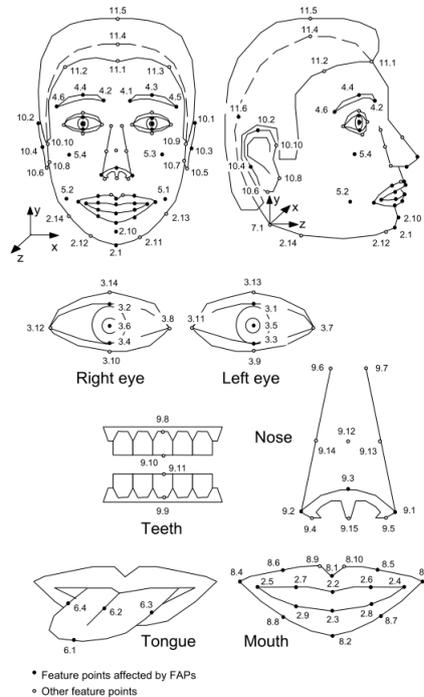


FIGURE 4.16: Norme MPEG-4

Beaucoup de travaux s’appuient sur cette normalisation de l’animation faciale. Des techniques simples à base d’interpolation de données permettent de calibrer un visage générique à partir d’un ensemble restreint de points caractéristiques (Pandzic & Forchheimer, 2002 ; El Rhalibi, Carter, Cooper, & Merabti, 2011). Pasquariello et al. ont proposé un système basé sur des pseudo-muscles compatible avec cette norme (Pasquariello & Pelachaud, 2002).

4.4.2 Technique d’animation

4.4.2.1 Déformations de surface

La technique de paramétrisation par déformation de surface consiste à fixer des paramètres sur la surface de l’objet. Les modifications des poids de ces paramètres entraînent des déformations géométriques du visage. Cela tout en prenant en considération la structure physique sous-jacente du visage. L’animation est réalisée par interpolation entre deux configurations de points-clés qui serviront de paramètres à la déformation du maillage (technique nommée *skinning*).

La méthode du *skinning* consiste à placer chaque point clé à une position précise sur le

visage. Pour chacun d'entre eux, on associe une région d'influence sur la surface du visage. Ainsi, le déplacement de chaque point clé déplace une zone du visage. Ces régions d'influences sont pondérées. Ainsi, les sommets les plus éloignés seront moins impactés par le point clé que les vertices très proches du point clé considéré. La figure 4.17 illustre un exemple de travaux se basant sur la technique d'animation par point clé (Courgeon, 2011). Plusieurs configurations de points clés ont été proposées, mais le modèle MPEG-4 demeure la configuration la plus répandue. Jusqu'à récemment, le skinning était la technique la plus répandue que se soit dans le milieu de la recherche ou celle de l'industrie. En effet, l'animation paramétrique est l'approche la moins consommatrice de puissance de calcul et elle ne nécessite que peu de données en comparaison des deux autres approches. Cependant, avec l'avènement des nouvelles ressources matérielles plus performantes, cette technique est de plus en plus délaissée pour des techniques offrant un degré de réalisme plus important.

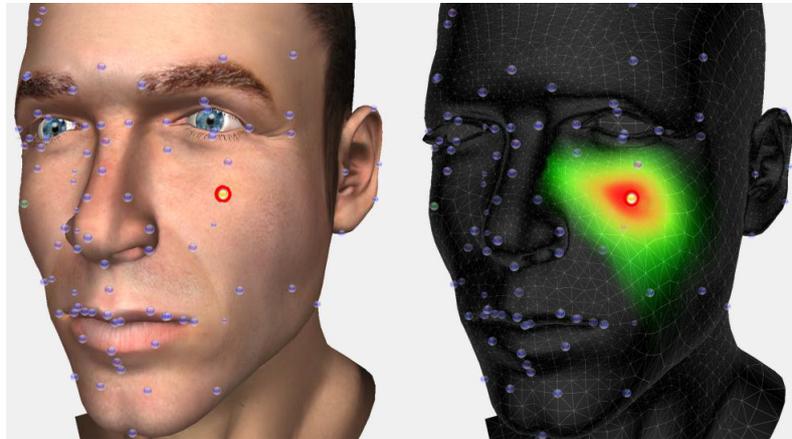


FIGURE 4.17: Exemple de paramètres de Skinning. À Gauche : Positions des points-clés. À Droite : Influence de l'un des points-clés (Rouge : très fort, Vert : Moyen, Noir : Aucune influence) (Courgeon et al., 2011)

4.4.2.2 Simulation musculaire

Pour obtenir des résultats réalistes et physiquement corrects, la recherche s'est également penchée sur des modélisations que nous nommerons « physiques » pour représenter le visage humain et ses déformations. Ces paramétrisations visent à simuler les mécanismes de déformations tels qu'ils existent réellement sur le visage humain.

L'animation par simulation physique consiste à considérer l'ensemble du visage humain, et non pas uniquement sa surface. Les expressions ne sont pas dues à une simple déformation de la peau, mais à une simulation complexe des os, des muscles, des tendons et de la peau (Parke & Waters, 2008).

Ces modélisations portent sur deux aspects : la modélisation de la dynamique de la peau, et la modélisation des muscles et de leurs actions, ces deux modélisations sont très fortement liées. Elles peuvent être définies par des systèmes masses-ressorts, des systèmes vectoriels, ou à partir d'une représentation par éléments finis. Le système masse-ressort proposé en

1981 par Platt et Badler est le premier qui s'intéresse à la modélisation physique du visage (Platt & Badler, 1981). La peau est représentée par un maillage de ressorts sur lesquels les muscles agissent. Par la suite Waters (Waters, 1987) a défini une modélisation devenue populaire des muscles à base de vecteurs. Chaque vecteur est défini par un point d'origine, une direction du champ vectoriel et un point d'insertion au niveau de la peau (le maillage). Son système basé sur FACS permet de définir une expression indépendamment de la morphologie du visage. Ces méthodes offrent d'excellents résultats, cependant, les temps de calcul nécessaires sont très importants, de l'ordre de 8 minutes par frame (Sifakis, Neverov, & Fedkiw, 2005). La figure 4.18 montre les travaux de Sifakis et al. qui se basent sur une simulation de 32 muscles.

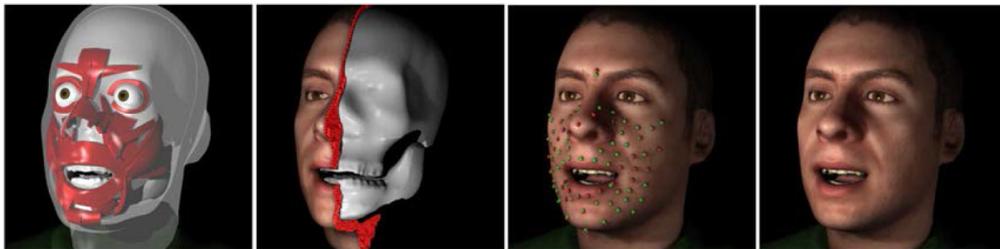


FIGURE 4.18: Modèle d'animation musculaire (Sifakis et al., 2005)

4.4.2.3 Interpolation de formes (Blendshape)

Les techniques de blendshape consistent à créer un espace linéaire d'expressions faciales et à réaliser des interpolations entre les formes de base pour en obtenir une nouvelle. Une expression est alors définie par une combinaison linéaire de n vecteurs, chaque vecteur définissant une forme. Dans le cas d'un maillage, il s'agit des coordonnées de chacun de ses sommets sachant que la topologie entre chaque forme doit être identique. La solution couramment utilisée pour générer de nouvelles expressions avec un modèle contrôlé par blendshape est de manipuler des sliders, qui correspondent aux poids attribués aux formes de base. Ces fonctionnalités existent nativement dans les logiciels commerciaux tels que Maya ou Unity3D. La qualité d'un modèle animé par blendshape est très fortement liée à la qualité et à la quantité des expressions de base fournies. Étant donné qu'elles définissent le domaine des déformations réalisables, elles doivent être créées de sorte que chaque expression désirée puisse être obtenue par mélange des expressions initiales. Ce nombre de formes clés peut-être très important. Le travail nécessaire à leur création est conséquent. Pour minimiser ces efforts de paramétrisation du modèle, certains articles proposent des outils pour faciliter la création des formes de base "Blend" ou les optimiser (Li, Weise, & Pauly, 2010). Les techniques de blendshape ont récemment démontré leurs qualités dans des projets tels que Digital Emily Project et digital IRA réalisés par l'université de Californie du Sud en association avec "Image Metrics"⁷ (O. Alexander et al., 2010; von der Pahlen et

7. <http://www.image-metrics.com/>

al., 2014). Dans le cadre de ces projets, les formes unitaires ont même été scannées à l'aide de la technologie de LightStage 4.19.

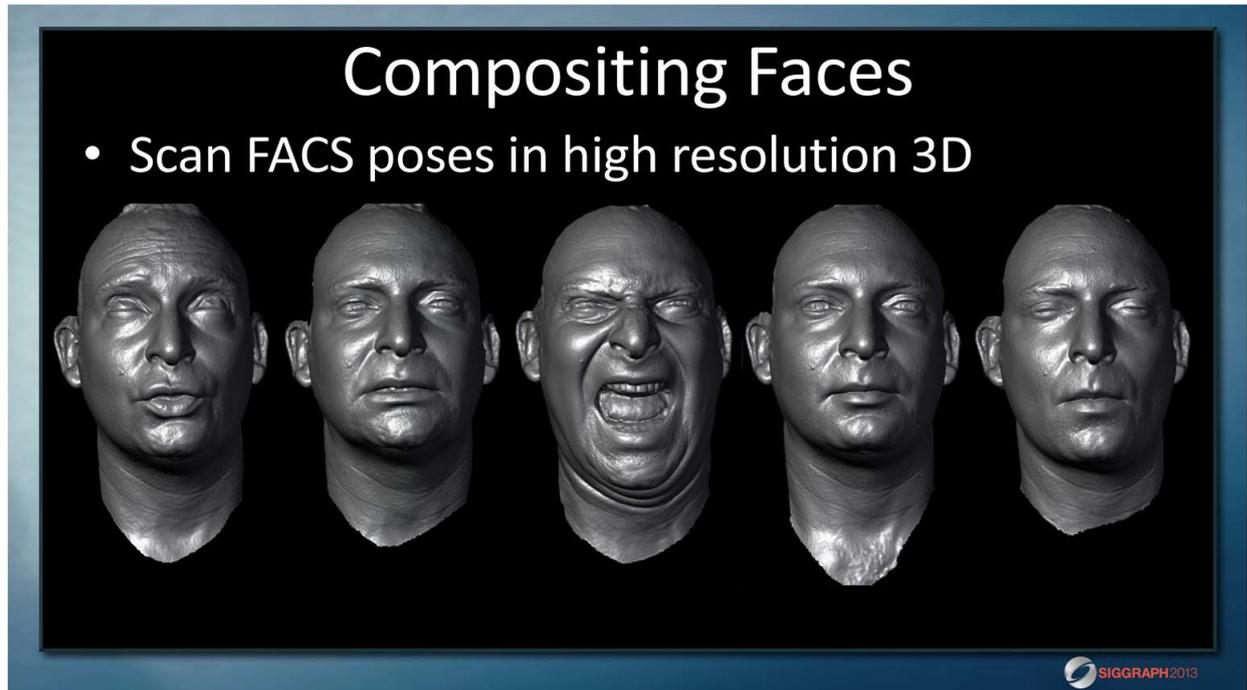


FIGURE 4.19: Scan de position de Blendshapes au sein du projet Digital Ira (Von der Pahlen et al., 2014)

4.5 Limites de la modélisation d'humains virtuels

Malgré les avancées considérables qu'a connues la modélisation de personnage virtuel, ces derniers doivent encore faire face à plusieurs défis.

4.5.1 Uncanny valley

Dés 1906 Ernst Jentsch, un psychologue allemand souligne que le sentiment d'étrangeté naît d'une situation d'incertitude, quand une personne a une difficulté à prendre une décision. La question "est-ce que l'objet que nous apercevons est vivant ou non?" crée un doute dans l'esprit suscitant un sentiment de rejet. (Kaba, 2013).

Le concept initialement abordé en psychologie a été utilisé avec l'apparition des premiers robots humanoïdes (Mori, 1970). Par la suite cette notion a été également étendue pour discuter le réalisme des personnages virtuels.

L'uncanny valley fut donc initialement définie par le roboticien japonais Masahiro Mori. Selon cette théorie, plus un robot androïde est similaire à un être humain, plus ses imperfections nous paraissent dérangeantes (Bartneck, Kanda, Ishiguro, & Hagita, 2009) (figure-

4.20).



FIGURE 4.20: Hiroshi Ishiguro et son double robotique Geminoid HI-1 (Bartneck, 2009)

Avec les niveaux de réalisme d’humain virtuel atteint actuellement, la moindre imperfection a de fortes chances de générer un sentiment de rejet envers l’humain virtuel. D’autant plus que ce phénomène intervient en particulier lorsque différentes modalités de l’humain virtuel possèdent des niveaux de réalisme différents. Par exemple si un visage virtuel photo-réaliste était animé avec une expressivité caricaturale, il tombera dans cette vallée de l’étrange (Walters et al., 2008).

Cet effet concerne également le comportement de l’agent en général. En effet, une apparence très humaine provoque des attentes importantes sur les capacités de raisonnement de l’agent et sur son comportement social. Ces attentes seront moins élevées si l’agent a un aspect plus robotique (Goetz et al., 2003).

4.5.2 Interactivité et cout matériel

Nous avons vu que plusieurs types de scanners 3D existent, ils coûtent généralement entre 35,000 à 500,000\$. La taille et le coût de ces scanners limitent les demandes de modèles de corps 3D. De même, pour les scanners à lumière structurée, ils sollicitent des algorithmes de traitement d’images nécessitant un temps de calcul considérable et par conséquent d’importantes ressources matérielles.

De plus, les scans 3D génèrent des données hautes définitions dont la résolution se mesure en millions de polygones. Par conséquent, ces données ne sont pas adaptées directement pour le traitement en temps réels. Ils nécessitent une phase de post-traitement pour être intégrés dans une application interactive.

Par ces limites, les clones virtuels ont le plus souvent été utilisés dans les rendus pré-calculé pour des productions ayant des budgets élevés (e.g. dans les effets spéciaux au cinéma). Peu d’applications intègrent des clones virtuels dans une application interactive.

Modélisation, rendu et animation d'un clone virtuel

Sommaire

5.1	Introduction	58
5.2	Capture des données statiques	58
5.2.1	Capture de la géométrie	58
5.2.2	Post traitement	60
5.2.3	Édition des textures	61
5.3	Capture des données dynamiques	61
5.3.1	Protocole d'élicitation	62
5.3.2	Session de capture de mouvements	64
5.4	Synthèse	64

5.1 Introduction

Ce chapitre a pour but de faire le pont entre le paradigme conceptuel de fidélité définie dans le chapitre 3 (Figure 3.1) et l'état de l'art de la modélisation d'humain virtuel en informatique graphique. Le but de notre démarche est de modéliser les différentes composantes de la fidélité lors de la création d'un clone virtuel à travers les moyens offerts par l'informatique graphique et les technologies de numérisation tridimensionnelle.

Nous allons décrire étape par étape les différentes composantes nécessaires pour créer un clone virtuel expressif et fidèle.

Nous avons également comme objectif de proposer une solution de capture qui soit abordable et accessible à un grand public. Cela tout en respectant une autre obligation, nos clones étant destinés à une interaction homme-machine en temps réel, nous devons respecter des contraintes en terme de temps de calcul.

5.2 Capture des données statiques

5.2.1 Capture de la géométrie

La première caractéristique discriminante d'un visage est sa forme. En effet, nous avons vu que la forme générale du visage et la distance géométrique entre ces différents éléments sont très importantes lors du processus d'identification (paragraphe-2.3.1.1). Pour préserver ces informations cruciales pour la fidélité, le choix d'un scan 3D nous a paru essentiel (par rapport à l'option d'une sculpture manuelle qui est sujette à l'interprétation de l'artiste graphiste).

Reproduire un vrai visage à l'aide de l'infographie nécessite la capture de la géométrie faciale en haute résolution. Grâce aux technologies de numérisation tridimensionnelle, nous sommes aptes à capturer la taille des différentes caractéristiques et la distance entre les principaux traits (par exemple, les distances entre les yeux ou entre le nez et la bouche, etc.) assurant par conséquent la fidélité visuelle holistique présentée dans notre paradigme conceptuel (Figure 3.1)

Dans le chapitre précédent, nous avons vu différents types de scanner et de techniques de numérisation. Chaque technique a ses avantages et ses inconvénients qui se mesurent surtout en termes de rapport cout/qualité.

Notre choix s'est porté sur une méthode de scan optique "le scan Kinect"¹. En effet, nous avons opté pour une solution dite "low cost" qui est accessible au grand public et qui permet de réaliser rapidement une capture 3D. Les scans par photométrie (lumière structurée) nécessitent des algorithmes de traitement d'image complexes. Quant au scanner laser, ils sont encore actuellement assez onéreux.

1. <https://msdn.microsoft.com/en-us/library/dn188670.aspx>

Scan kinect Nous avons vu qu’avec l’aide des technologies de numérisation 3D, tels que la lumière structurée ou le balayage laser, nous pouvons obtenir des résultats détaillés de visages humains. Cependant, ces dispositifs nécessitent une connaissance approfondie des ressources matérielles et requièrent souvent des compétences importantes en traitement d’image. De plus, ils sont très sensibles aux conditions d’illumination et nécessitent souvent des temps de calcul importants (Weiss, Hirshberg, & Black, 2011).

Un nouveau genre de dispositifs a donc été de plus en plus utilisé basé sur les caméras de profondeur. Récemment la Microsoft Kinect a attiré de plus en plus de travaux dans la communauté de recherche sur les humains virtuels et le traitement d’images (Tong et al., 2012 ; Sun et al., 2012).

La Kinect est un appareil de capture de mouvement développé par Microsoft pour les jeux vidéo. L’appareil utilise une caméra et des rayons infrarouges pour faire ses mesures. L’image est ensuite transmise à l’ordinateur qui va en déduire les distances entre certains points situés à la surface de l’objet et le capteur Kinect. En comparaison avec les scanners 3D classiques, elle est en mesure de capturer des données de profondeurs à la cadence de la caméra RGB. La Microsoft kinect a aussi l’avantage d’être peu sensible aux conditions de lumière ambiante. Elle est de plus bon marché, compacte et facilement manipulable.

Des études sur le scan kinect soutiennent que, pour le scan de corps, il est possible d’atteindre des degrés de précision similaires à un scan laser ou scan à base de lumière structurée (Weiss et al., 2011). Cependant cette méthode ne donnait pas de résultats satisfaisants pour le visage. En effet, le visage comporte des détails beaucoup plus fins que le reste du corps. Des études plus récentes ont utilisé la Kinect pour réaliser des scan3D de visage (Hernandez, Choi, & Medioni, 2012). Dans cette étude Hernandez et al. soutiennent que leur méthode qui consiste à utiliser plusieurs poses afin d’accumuler les sources d’informations manquantes et compenser l’information bruitée donne des résultats avec une qualité équivalente à celles des scanners laser professionnels.

Toutefois, un principal problème persiste. La kinect a une résolution relativement basse selon l’axe le plan X/Y et les données de profondeur manquent de précision nécessitant des algorithmes de traitement d’image pour combler la perte de données ou un post traitement manuel. Avec la sortie de la kinect 2 Microsoft a réduit un peu cette problématique en multipliant la résolution de la caméra par deux.

Malgré une faible résolution de ces deux caméras, la Kinect permet à travers diverses solutions logiciels (ReconstructMe² ; Kinectfusion³) la captation 3D d’un modèle avec une texture d’un visage et cela en à peine quelques minutes (Sun et al., 2012).

Notre choix s’est porté donc sur cette solution. Nous avons utilisé le logiciel ”ReconstructMe”⁴. Les figures 5.1 et 5.2 montrent le déroulement du scan et les données brutes que nous avons obtenues.

2. <http://reconstructme.net/>

3. <https://msdn.microsoft.com/en-us/library/dn188670.aspx>

4. <http://reconstructme.net/>

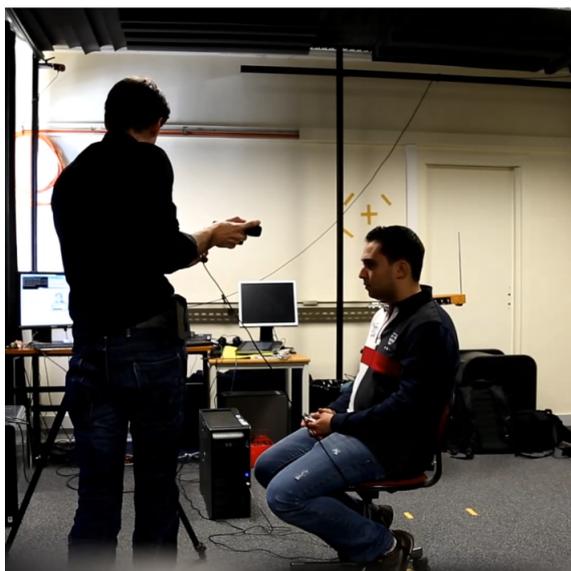


FIGURE 5.1: Déroulement du scan Kinect

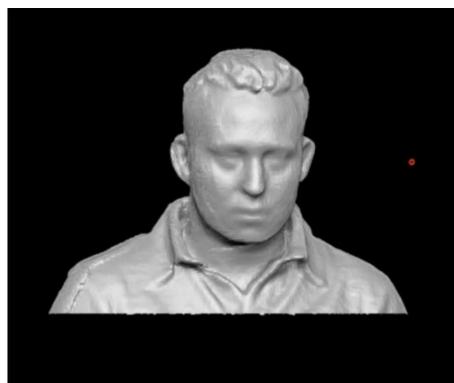


FIGURE 5.2: Données brutes du scan Kinect

5.2.2 Post traitement

La kinect, en contrepartie de sa facilité d'utilisation, offre des résultats qui restent relativement bruités. Un post traitement des données était nécessaire. Pour cela nous avons utilisé le logiciel Zbrush⁵.

Ce post-traitement consistait dans une première étape à creuser les orifices des yeux et de la bouche et à affiner les détails des oreilles. Dans une seconde étape, étant donné que le scan kinect aboutit en un nuage de points triangularisé, le maillage fourni était extrêmement lourd (plus de deux millions de polygones). Or, il nous fallait reconstruire un maillage qui soit apte à être animé en temps réels. Par conséquent, il fallait procéder à une retopologie du maillage (Figure—5.3). Le principe de la retopologie, c'est de créer un modèle par-dessus un autre pour en faire une version adaptée et efficace (pour de l'animation par exemple).



FIGURE 5.3: Phase de retopologie du maillage avec le logiciel Zbrush

5. <http://pixologic.com/zbrush/features/overview/>

5.2.3 Édition des textures

La deuxième composante de la fidélité visuelle holistique est la pigmentation de la peau et ses caractéristiques lumineuses. En effet nous avons vu que la couleur de la peau était le deuxième élément primordial assurant l'identification lors de l'évaluation des caractéristiques holistiques du visage. L'informatique graphique fait référence à toutes les propriétés de pigmentation et de réflectance de surface de la peau et à leur variation spatiale à travers diverses textures y compris l'albédo, la carte des normales et la "displacement map" pour simuler la rugosité de la peau.

Il a fallu donc, dans une première étape, générer une map de texture. Cette étape a nécessité une phase de prise de photos du référent. Afin d'arriver à un résultat le plus optimal possible et pour obtenir une texture indépendante des conditions d'illumination ambiantes, nous avons mis en place un éclairage spécifique à l'aide de plusieurs spots de lumière blanche afin de minimiser les ombres portées. Par la suite, à l'aide du logiciel Zbrush nous avons procédé aux plaquages de la texture photo sur le maillage 3D capturé comme l'illustre la figure 5.4. La texture des normales et la "displacement map" furent par la suite générées à l'aide du logiciel Zbrush.

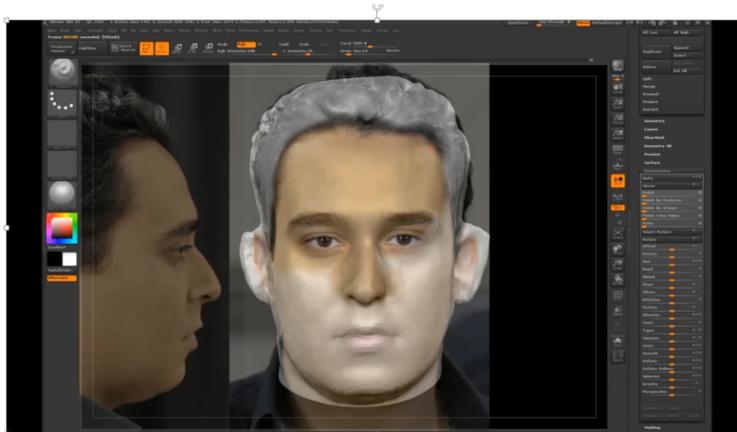


FIGURE 5.4: Plaquage de la texture

5.3 Capture des données dynamiques

Toujours dans l'optique de modéliser notre paradigme de fidélité, nous entamons dans cette section la modélisation de la fidélité expressive dans un clone virtuel en capturant un espace expressif du référent. Nous définissons un espace expressif comme étant l'ensemble des caractéristiques des mouvements faciaux d'un être humain. Cette espace comprend à la fois les déformations physiques de cet individu (sa manière d'activer les différents muscles du visage), mais aussi sa manière d'exprimer différentes émotions en plus des mouvements idiosyncrasiques qui lui sont propres.

Se pose alors la question de comment capturer et animer fidèlement cet espace expressif,

d'autant plus que d'innombrables procédés de capture de mouvement et de techniques d'animation sont offerts par l'informatique graphique comme nous l'avons constaté dans le chapitre précédent.

Dans cette phase de nos travaux où le clone virtuel est non autonome, la meilleure solution qui nous est apparue est le rejeu de données de capture de mouvement. En effet, l'animation faciale par corpus est actuellement l'unique méthode temps réel qui génère des animations dont la dynamique est fidèle à celle du visage humain. Elle consiste effectivement à reproduire les mouvements d'un visage humain à partir d'enregistrements de position 3D qui sont associés par la suite au maillage. Nous nous sommes basés sur une animation combinant technique de Blendshape et données de capture de mouvement. Notre choix s'est porté sur l'animation par Blendshapes car c'est une technique qui est générique si on dispose d'un espace linéaire représentatif des expressions faciales unitaires. Elle offre un degré de réalisme proche de celui de la déformation musculaire pour des temps de calcul plus raisonnables (Deng, Chiang, Fox, & Neumann, 2006). Elle consiste en une interpolation de différent maillage représentant des déformations unitaires pour retrouver l'ensemble du spectre expressif.

5.3.1 Protocole d'élicitation

Pour obtenir une expressivité propre au clone, il est nécessaire de capturer l'espace expressif du référent. Dans nos travaux nous nous concentrons sur l'expression émotionnelle qui est essentielle pour une interaction homme-machine.

Nous avons également besoin de capturer plusieurs expressions émotionnelles dynamiques. Pour cela, nous avons défini un protocole d'acquisition d'expressions faciales émotionnelles dynamiques. Afin de permettre un haut niveau de réalisme de l'expression faciale émotionnelle et de la dynamique associée, nous avons parcouru les différents moyens d'élicitation susceptibles de nous permettre d'obtenir des données expressives naturelles.

Différents moyens sont cités dans la littérature. La première technique consiste à réaliser de l'induction d'états émotionnels par l'utilisation de films (Schaefer, Nils, Sanchez, et Philippot, 2010), de musique (Niedenthal et Setterlund, 1994), ou d'images (Lang et Bradley, 2007). Il est également possible de faire appel à la mémoire du sujet en essayant de lui faire remémorer des souvenirs émotionnels (De Dreu, Baas et Nijstad, 2008).

Cependant, si l'on veut étudier la perception de l'émotion en tant que composante des interactions sociales, il est nécessaire de capturer des expressions émotionnelles en situation d'interaction. Cela nécessite de donner à la personne capturée des instructions pour produire une expression émotionnelle globale, sans se concentrer sur une modalité spécifique (Bänziger, Mortillaro, et Scherer, 2012).

Nous nous sommes donc intéressés à un autre type de procédure nommée "communication effect acting" : l'acteur doit produire une émotion, généralement spécifiée par un label. La signification de ces étiquettes émotionnelles est discutée avec les participants pour éviter des interprétations personnelles de ces dernières (Hawk et al., 2009).

Nous avons finalement opté pour un type de procédure qui peut être considéré comme une combinaison entre une méthode d'induction et la procédure "communication effect acting". Cette dernière procédure est appelée "felt experience enacting procedure".

Elle diffère de la procédure précédente par l'utilisation de techniques d'imagination mentale. Le but est de faciliter l'adoption d'une émotion lors de l'expression. Cette approche est généralement utilisée en recherche, car elle ne fait pas d'hypothèse a priori sur l'existence systématique de motifs connus et fixes pour toutes les émotions. En conséquence, les espaces expressifs obtenus grâce à cette procédure sont susceptibles de présenter d'importantes variations d'expression plutôt que d'un modèle commun prototypique.

Geneva Multimodal Emotion Portrayals (GEMEP) : GEMEP est une collection d'enregistrements audio et vidéo avec 10 acteurs représentant 18 états affectifs, avec différents teneurs verbales et différents modes d'expression. Il a été créé à Genève par Klaus Scherer et Tanja Bänziger en utilisant l'approche "felt experience enacting technique" (Bänziger et al., 2012).

L'intérêt de cette base est que les scénarios d'élicitation d'émotion sont disponibles. L'émotion induite par chaque scénario a été validée avec des études sur un grand nombre de sujets. Cette base présente les 6 émotions primaires avec 12 émotions secondaires. À chacune de ces émotions sont associées trois scénarios possibles.

Les auteurs ont fourni aux acteurs, une ou deux semaines avant les enregistrements, une liste de courtes descriptions explicatives de la signification de l'émotion et trois scénarios pour chacune des 15 catégories d'émotion que l'acteur devra jouer. Un metteur en scène a travaillé avec l'acteur pour choisir un scénario particulier qui était personnellement pertinent pour ce dernier. Avant chaque expression, les scénarios ont été intégrés dans une interaction entre le metteur en scène et l'acteur pour augmenter l'intensité émotionnelle dans le contexte interpersonnel et afin d'avoir l'expression dynamiquement émerger de l'interaction impliquée. L'acteur n'a reçu aucune instruction sur la façon d'exprimer les émotions (Bänziger and Scherer, 2010).

Dans le cadre de nos travaux de thèse, nous nous sommes inspirés de cette procédure lors de nos sessions de capture de mouvements faciaux. Les scénarios GEMEP ont été envoyés au référent quelques jours précédent la session de capture. Le protocole GEMEP suggère trois possibilités de scénario pour chaque émotion. Nous avons donné au participant comme instruction d'opter pour le scénario qu'il évaluait comme étant le plus autobiographique des trois. Nous lui avons également donné la possibilité d'y effectuer de légères modifications pour le rendre plus personnel et augmenter par conséquent le ressenti de l'émotion lors de son expression.



FIGURE 5.5: Disposition des marqueurs lors de la session de capture de mouvements

5.3.2 Session de capture de mouvements

La capture de mouvement a été réalisée avec le setup de mocap Optitrack Arena⁶. Nous avons mis en place un setup de 9 caméras "Prime 13"⁷. Ces caméras possèdent une résolution de 1.3 millions de pixels et offrent une précision 3D de 0.5 mm à longue distance et de 0.2 mm à proximité. Nous avons positionné les caméras infrarouges en arc de cercle à une distance de 60 cm du référent. Pour les marqueurs nous avons utilisé la configuration suggérée par le logiciel Optitrack Arena basé sur 37 marqueurs pour le visage et 4 marqueurs de tête responsables de la capture des mouvements rigides de la tête comme l'illustre la figure 5.5.

5.4 Synthèse

Dans la perspective de réaliser le rendu et l'animation en temps réel des différentes composantes de la fidélité visuelle définies dans notre cadre conceptuel. Nous avons proposé un processus permettant la capture, l'édition et le rendu d'un clone virtuel dans le moteur de jeux Unity3D. Notre technique est basée sur la combinaison de la résolution spatiale à partir de scans 3D et la résolution temporelle à partir des données de capture de mouvement.

Nous avons proposé une solution accessible nous permettant l'édition de clone virtuel animé

6. <http://www.optitrack.com/>

7. <http://www.optitrack.com/products/prime-13/indepth.html>

en temps réel à bas coût et ne nécessitant pas l'utilisation de scanner 3D professionnel. La figure 5.6 résume les différentes étapes du processus d'édition d'un clone virtuel. Des études perceptives sur l'évaluation de la fidélité d'un clone virtuel peuvent être envisagées grâce à ces travaux.

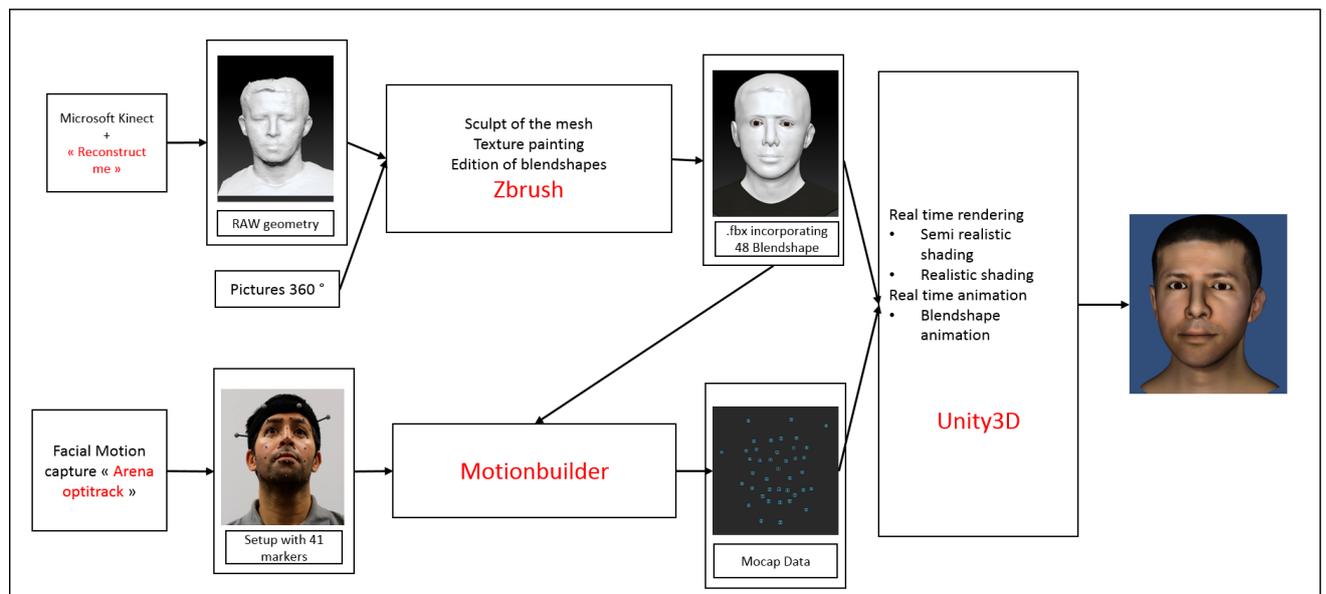


FIGURE 5.6: Pipeline de création de clones virtuels

Troisième partie

Évaluation de la fidélité visuelle

État de l'art : Fidélité visuelle

Sommaire

6.1	Introduction	70
6.2	L'interaction entre l'apparence et l'expressivité lors de l'identification	70
6.2.1	Le rôle de l'apparence dans la reconnaissance	71
6.2.2	Le rôle de l'expressivité dans la reconnaissance	74
6.2.3	La primauté du visuel sur l'expressivité lors d'une tâche de reconnaissance	76
6.3	Le rôle de la familiarité	77
6.4	Des métriques pour l'évaluation de visages virtuels	78
6.5	Synthèse	78

6.1 Introduction

Les études en psychologie ont démontré que notre traitement cognitif face à des visages familiers diffère du traitement des visages inconnus (Lander, 2005, Wallraven, 2007). Ces études avancent, entre autres, que le traitement de visages familiers est cognitivement moins coûteux et que l'attraction du regard par les visages familiers est automatique (Lander, 1999, Gobbini, 2007). Le fait qu'un clone virtuel véhicule une identité, spécialement si cette dernière est familière, peut donc avoir un impact important sur la perception et l'évaluation d'un visage virtuel.

Il est donc important d'appréhender les composantes sur lesquelles nous nous basons pour évaluer la fidélité visuelle d'un clone virtuel. Dans cette section nous allons explorer les différentes études menées sur l'évaluation de visages virtuels abordant les variables apparentées à la fidélité d'apparence et la fidélité expressive.

Bien que de nombreuses études ont analysé les mécanismes d'identification et de reconnaissance de visages familiers sur la base de photos ou de vidéos, l'étude de la perception de visage familier virtuel n'a été abordée que très rarement. En effet, très peu de travaux ont étudié la perception et la reconnaissance des clones virtuels expressifs.

Il existe de nombreuses études abordant l'analyse de visage virtuel, mais pour la plupart de ces travaux, elles ont exploré la perception de différents degrés de réalisme et la perception de rendu artistique (McDonnell, Breidt, & Bülthoff, 2012 ; Wallraven et al., 2005). Pour d'autres, ils ont abordé l'impact des conditions de vue sur la perception de visage virtuel (Wallis & Bülthoff, 2001). Mais rarement une étude s'est portée sur la reconnaissance de l'identité véhiculée par le visage 3D dynamique.

De plus, en ce qui concerne les études portant sur l'identité, la plupart de ces derniers se sont focalisés sur la façon dont l'information statique affecte la perception de l'identité : peu d'étude se sont intéressés au rôle des expressions faciales dynamiques dans l'identification du visage virtuel (Bailenson, 2012).

6.2 L'interaction entre l'apparence et l'expressivité lors de l'identification

Lander et al (Lander et al., 1999) soutiennent que l'information utilisée pour identifier un visage humain se trouve à la fois dans les caractéristiques de la structure invariante du visage (apparence) et dans les mouvements et gestes idiosyncrasiques (expressivité). Mais la question qu'amènent nos travaux est : quels sont les poids respectifs de chaque modalité dans une tâche d'identification ? Est-ce qu'il y a des effets d'interactions entre l'apparence et l'expressivité ?

6.2.1 Le rôle de l’apparence dans la reconnaissance

L’apparence demeure le premier élément discriminant dans la reconnaissance de personnes. Une meilleure compréhension de la manière avec laquelle le cerveau traite l’information visuelle du visage pourrait nous guider à mieux modéliser des clones virtuels. Nous avons préalablement vu que les études en psychologie démontrent que les visages sont traités par analyse des informations des composants (yeux, bouche, nez) et leurs relations spatiales (information holistique).

Il y a eu des travaux sur la perception de visages générés par ordinateur qui ont exploré ces caractéristiques visuelles (Fan et al., 2014). Ces études abordaient la perception du réalisme. Elles n’avaient pas pour vocation d’étudier la fidélité ou l’identification d’un référent. Ils se sont également basés uniquement sur des stimuli statiques. Les auteurs soutiennent que les facteurs relatifs à l’évaluation par trait tiennent une place plus importante pour l’évaluation du réalisme quand la définition de l’image est importante et respectivement les observateurs ont tendance à se baser sur les éléments holistiques quand la résolution de l’image décroît. Au cours de cette étude, les auteurs ont également confirmé que les yeux sont les éléments les plus influents pour le jugement du réalisme.

Plusieurs autres études ont également abordé la perception de visage virtuel mais avec une approche plus informatique, se basant sur des notions plus graphiques, telles que les effets de la modification de la texture, de la modification de la forme et du rendu (shading)(McDonnell et al., 2012; Ring, Utami, & Bickmore, 2014). Une autre étude de Fan et al. (Fan et al., 2012) a tenté d’explorer les variables qui font qu’un visage généré par ordinateur soit perçu comme réaliste (voire d’être confondu avec des photos réelles). Ils ont mené une étude où ils ont comparé la perception de rendu généré par ordinateur et de vraies photos. Ils en ont conclu que le shading (rendu) et la gestion des ombres étaient les caractéristiques les plus importantes pour la distinction des visages réels de ceux qui sont générés en 3D. Leurs résultats suggèrent aussi que la perception de visage virtuel est altérée par les différences inter-individuelles de l’observateur telles que l’ethnicité ou l’expérience avec les médias virtuels.

Dans les sections qui suivent, nous examinons comment les concepts du style de rendu, des textures et les proportions du maillage changent les attitudes envers les agents virtuels.

6.2.1.1 Effet du rendu

Dans le cadre de l’étude de la fameuse théorie de ”l’Uncanny Valley” plusieurs études ont établi différents degrés de réalisme et ont étudié la perception de ces humains virtuels (Seyama & Nagayama, 2007).

McDonnell et al. (McDonnell et al., 2012) ont étudié la mesure avec laquelle différents styles de rendu d’un même humain peuvent affecter la perception des utilisateurs. Ils ont défini une échelle s’étalant sur 10 degrés de réalisme de l’esquisse au crayon jusqu’au rendu photo-réaliste (Figure 6.1). À travers cette échelle ils ont mené des études pour évaluer l’impact du réalisme sur certains aspects sociaux tels que l’appréciation, la confiance et la perception du mensonge.

Leurs résultats suggèrent que le rendu cartoon et le rendu le plus photo-réaliste sont les rendus les plus acceptés par les participants.

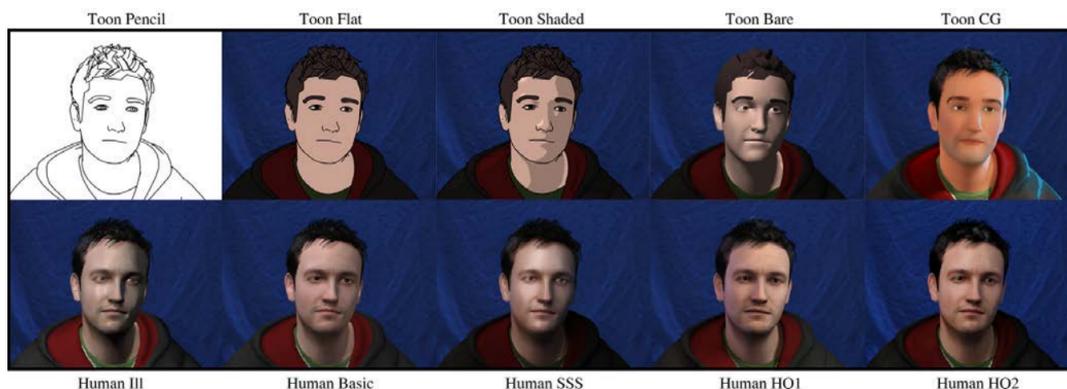


FIGURE 6.1: Exemple de 10 styles de rendu classés du plus abstrait au plus réaliste (McDonnell et Al., 1012)

Une autre étude (Wallraven, Bühlhoff, Cunningham, Fischer, & Bartz, 2007) a rapporté une évaluation perceptive de trois techniques de stylisation différentes. Les auteurs ont associé des expressions faciales dynamiques à ces trois styles de rendu afin d'étudier l'impact d'un rendu non réaliste (artistique) sur l'évaluation subjective des expressions émotionnelles. Ils ont conclu que la variation de rendu n'avait pas d'effet significatif sur le taux de reconnaissances des émotions.

6.2.1.2 Les effets de l'altération de la forme du visage virtuel

Les études en psychologie soutiennent que les humains sont tolérants aux déformations de formes si la personne observée est familière (Hole, George, Eaves, Rasek, et al., 2002). Par exemple, (Sinha et al., 2006) décrit une étude dans laquelle des visages déformés de gens célèbres peuvent être compressés ou au contraire étirés grandement sans aucune perte de reconnaissance. La complexité des informations de forme a été également étudiée par (Wallraven et al., 2005) qui a démontré par contre qu'elle avait un impact significatif sur la perception de l'intensité des expressions faciales.

La modification de la forme et les proportions d'un personnage virtuel animé ont également été étudiées dans le cadre de l'investigation du phénomène de "l'uncanny valley". Ainsi, inspirée par des personnages virtuels populaires célèbres comme celui de "Gollum" dans la franchise des seigneurs des anneaux, des chercheurs ont étudié l'effet d'un maillage disproportionné sur l'acceptation d'humain virtuel.

En effet, beaucoup d'étude telles que (Green, MacDorman, Ho, & Vasudevan, 2008 ; MacDorman, Green, Ho, & Koch, 2009) soutiennent que nous avons plus de facilité à accepter des représentations d'avatar dont les proportions ne sont pas plausibles. Cet effet est, selon (Ring et al., 2014), dépendant de la tâche à laquelle est destiné l'agent (figure-6.2). En effet,

pour des applications professionnelles tel qu’un assistant médical les utilisateurs préfèrent un avatar avec des proportions normales alors que pour des applications de divertissement les utilisateurs ont des préférences pour les avatars avec un rendu cartoon dont les proportions ne sont pas plausibles.

Une autre étude (Green et al., 2008) a comparé les réactions d’observateurs face à des proportions variables de visages réels, de robot et d’humain virtuel. Leurs résultats suggèrent que plus un visage paraissait humain et attrayant moins les observateurs étaient tolérants à la modification des proportions.



FIGURE 6.2: Étude sur la modification de la forme d’un visage 3D (ring et Al., 2014)

6.2.1.3 Impact de la modification de la texture

La peau affiche un large éventail d’informations que ce soit l’âge, le sexe ou l’état de santé. La perception de la peau est donc centrale dans l’identification et les interactions sociales (Giard & Guitton, 2010). Par conséquent, la modélisation et le rendu de la peau ont suscité un intérêt marqué chez les artistes d’images de synthèse. En informatique graphique, de nombreuses études ont étudié la perception du réalisme de la peau sur un visage (Jimenez, 2012 ; d’Eon, Luebke, & Enderton, 2007). Par exemple, selon (Struck, Bohn, Schmidt, & Helzle, 2004) la crédibilité de la peau est principalement basée sur les détails des plis et des pores.

Macdorman et al. (MacDorman et al., 2009) ont étudié l’interaction de la modification de la texture et des proportions. Ils ont mis en place trois niveaux de texture (réaliste, bronze et esquisse de crayon) et ont procédé à des déformations de proportion (e.g. un élargissement de 50 % de la distance entre les yeux). Leurs résultats suggèrent que plus la texture possède un niveau de réalisme important plus le visage virtuel est perçu comme étrange suite aux changements de proportions.

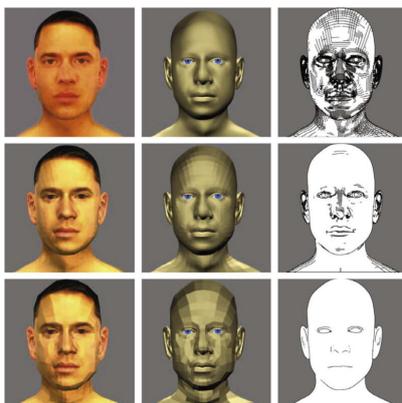


FIGURE 6.3: Texture réaliste, bronze et en trait variant sur 3 niveaux de détail (Macdorman, 2009)

Les études sur la perception des visages ont montré que la modification de la couleur d'une image affecte le taux d'identification. Nous avons vu précédemment l'importance des couleurs pour la reconnaissance de visage en général même sur la base de photos. Des expériences ont confirmé que la reconnaissance des images est plus élevée quand elles sont en couleurs que quand elles sont en noir et blanc (Yip & Sinha, 2002).

Une étude a même prouvé que des observateurs mis face à des photos de célébrités mondialement connues étaient incapables de les reconnaître quand on leur a présenté leur photo en négatif. En outre dans l'expérience de (McDonnell et al., 2012), décrite ci-dessus, les auteurs ont rapporté que la texture a un impact perceptif fort. Par exemple, l'ajout d'une teinte jaune à la texture d'un humain virtuel véhicule une apparence d'un visage malade, provoquant ainsi un sentiment de rejet auprès des participants.

6.2.2 Le rôle de l'expressivité dans la reconnaissance

Les mouvements expressifs non verbaux sont primordiaux pour nos interactions sociales. Mais améliorent-ils le processus d'identification ? Est-ce que notre mémoire visuelle se base sur des informations dynamiques pour la reconnaissance visuelle ?

6.2.2.1 Comment les expressions faciales contribuent-elles à la reconnaissance d'une identité ?

Il y a deux hypothèses concernant le rôle des mouvements faciaux dans la reconnaissance de l'identité :

- L'hypothèse de "l'amélioration de représentation" : Cette hypothèse avance que le mouvement du visage contribue à la reconnaissance de visage en facilitant la perception de la structure en trois dimensions du visage (Pike, Kemp, Towell, & Phillips, 1997 ; Christie & Bruce, 1998). Ainsi l'expression faciale aurait pour rôle de fournir des informations plus précises sur la forme du visage lors du déplacement de certaines de ses régions. Dans cette théorie, l'information dynamique améliore

uniquement l’information d’apparence. Elle ne posséderait pas en elle même une notion d’unicité.

- l’hypothèse de ”l’information supplémentaire” : Cette seconde hypothèse stipule que nous possédons des mouvements faciaux caractéristiques et des expressions faciales individuelles (Brockhaus, 2013). Ces mouvements idiosyncrasiques spécifiques à une personne sont dénommés ”signature faciale dynamique” (Lander et al., 1999). Par exemple, la manière spécifique avec laquelle quelqu’un sourit ou a l’air triste pourrait être utilisée pour l’identification (d’une manière similaire que des sourires d’acteurs célèbres tels que Jack Nicholson ou Tom Cruise sont hautement spécifiques pour eux et facilement identifiables) (O’Toole, Roark, & Abdi, 2002)



FIGURE 6.4: Signature dynamique : exemple d’expression caractéristique de l’hypothèse de ”l’information supplémentaire”

La première hypothèse suppose que le mouvement ajoute et contribue à la qualité de l’information structurelle du visage humain. Cet avantage est assuré par les vues supplémentaires du visage fournies au cours des mouvements faciaux. Cependant, si on présume cette hypothèse comme exclusive, alors la contribution du mouvement pour l’information d’identité serait perdue lorsque l’animation est transférée sur d’autres visages. Pourtant, des études utilisant des techniques de morphing et transfert d’animation ”Retargetting” telles que (Cook et al., 2012 ; Sinha et al., 2006) ont montré que l’affichage d’une animation spécifique à un sujet sur un visage générique affecte le taux d’identification.

Ainsi, nous pouvons conclure que l’hypothèse ”d’amélioration de la représentation” et l’hypothèse ”d’informations supplémentaires” ne sont pas mutuellement exclusives, mais pourraient être complémentaires. L’expressivité contribue à la fois à l’identification en donnant de plus amples informations sur la géométrie du visage (données d’apparence) et contient elle même des informations spécifiques à l’identité telles que des gestes idiosyncrasiques.

6.2.2.2 Composantes des mouvements faciaux

De nombreuses études affirment que le mouvement joue un rôle dans le processus d’identification (Lander & Chuang, 2005 ; Rizzo, Neumann, Enciso, Fidaleo, & Noh, 2001). Cependant, la question que nous nous posons à travers cette section est de savoir si tous les types de mouvement ont des rôles équivalents dans ce processus d’identification.

Les études antérieures sur le rôle des expressions faciales pour l'identification ont fait la distinction entre mouvements rigides et non rigides. Les mouvements faciaux rigides se réfèrent généralement à des rotations et/ou translations de la tête entière (comme le hochement de la tête par exemple). En revanche, les mouvements faciaux non-rigides se réfèrent à des déformations non linéaires de la surface du visage (par exemple, le mouvement des lèvres ou des sourcils).

Plusieurs études (O'Toole et al. 2011 ; Knight et Johnston 1997 ; Lander et Bruce 2000) ont comparé les rôles de ces deux catégories de mouvement. Ils ont utilisé des images dégradées des individus familiers (individus du même environnement de travail que les participants). Ils ont démontré un avantage pour les mouvements non rigides tels que les mouvements des lèvres par rapport aux mouvements rigides lors d'une tâche de reconnaissances. Contradictoirement (Cunningham et Wallraven 2009 ; Hill et Johnston, 2001) ont séparé des données de capture de mouvement en données rigides et non rigides, et ils ont constaté que les données rigides permettaient une meilleure identification par rapport aux données non-rigides (Bülthoff et al. 2011). Ainsi, la prédominance d'un type de mouvement sur l'autre pour une tâche d'identification n'est pas encore clairement établie.

De plus, la catégorisation des mouvements du visage en mouvements rigides ou non rigides pourrait être trop simpliste et pourrait ne pas englober la véritable complexité des expressions faciales. Des mouvements caractéristiques plus complexes (comme une certaine façon de sourire ou de diverses façons d'exprimer la même émotion) fournissent des informations sur des individus distincts. Ainsi, des études expérimentales sont nécessaires pour pouvoir clarifier les mouvements exacts du visage que nous retenons d'une personne et comment ils pourraient nous aider à identifier cette personne.

6.2.3 La primauté du visuel sur l'expressivité lors d'une tâche de reconnaissance

À travers notre recherche du rôle de l'expressivité dans l'identification, nous avons remarqué que la plupart des démonstrations de la contribution des expressions faciales lors d'une tâche de reconnaissance se font à la base d'images dégradées/modifiées. En effet, les stimuli sont pour le plus souvent des images altérées, en négatif ou encore des images floutées. Beaucoup de chercheurs stipulent que l'information dynamique ne joue un rôle pour l'identification que si l'information visuelle est incomplète ou altérée.

Quand les deux modalités sont présentes, des travaux tel que (Wallraven et al., 2005 ; O'Toole et al., 2002) avancent que des évidences psychologiques indiquent que l'information dynamique ne joue un rôle que si les conditions d'apparence ne sont pas optimales. Cela revient à dire que l'information dynamique vient en renfort pour compenser des pertes d'information due à la dégradation de l'information visuelle.

6.3 Le rôle de la familiarité

Comme nous l'avons constaté précédemment en définissant le cadre conceptuel pour l'étude de la fidélité, la familiarité est un élément primordial pour l'évaluation de la fidélité. Beaucoup de travaux ont prouvé l'impact de la familiarité sur la reconnaissance. La familiarité accrue avec le référent permet normalement de reconnaître une personne même dans des conditions de vision dégradées (peu de lumière, ou image floutée)(Sinha et al., 2006). Des travaux se sont également intéressés à l'interaction entre les expressions dynamiques et la familiarité. La littérature en perception de visage soutient que les mouvements faciaux favorisent la reconnaissance de visages familiers. De nombreuses études ont montré que les visages célèbres sont reconnus plus rapidement et plus précisément lors de leur présentation en mouvement par rapport à des vidéos d'inconnus (Knight & Johnston, 1997; Lander & Bruce, 2000).

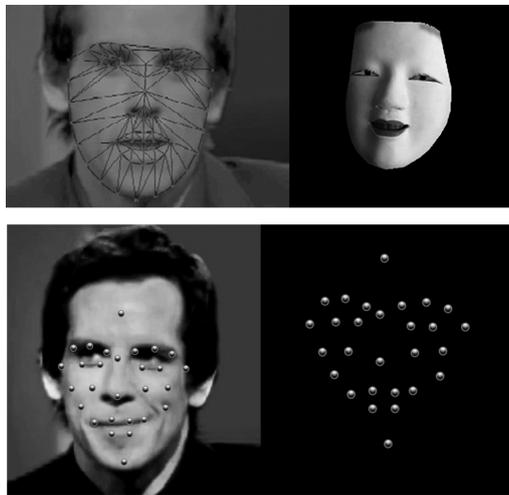


FIGURE 6.5: Étude de l'apport de l'expressivité pour l'identification de visage célèbre (Bennetts et al., 2013)

Cependant, dans une étude menée par Bennetts et al. (Bennetts et al., 2013), les auteurs ont exploré le rôle de la familiarité dans l'identification de visage statique et dynamique. Ils ont transféré la dynamique faciale de personnes célèbres et de personnes non familières sur un avatar générique ou un ensemble de points caractéristiques (figure 6.5). Leurs résultats ne montraient pas un avantage significatif de la familiarité dans la condition dynamique. Le rôle du mouvement dans la reconnaissance de visage familier virtuel reste donc incertain (Bennetts et al., 2013).

De plus, le fait d'être familier avec la personne nous rend experts de cette personne. Si les chances de reconnaissance d'un clone virtuel sont de fait augmentées par notre familiarité avec le référent, la probabilité de relever les moindres défauts et la moindre incohérence avec les caractéristiques réelles du référent s'en trouve aussi augmentée. La familiarité pourrait donc constituer une autre dimension de l'uncanny valley.

6.4 Des métriques pour l'évaluation de visages virtuels

Les informations suffisantes et nécessaires pour soutenir une identification demeurent inconnues et incertaines. Il est difficile pour un humain de décrire précisément sur quels éléments il s'est basé pour évaluer la ressemblance entre une représentation virtuelle et le visage réel. Aucune des études antérieures n'a clairement décrit l'information exacte qui est nécessaire ou suffisante pour reconnaître une personne en particulier. Concernant les visages virtuels, il est également difficile d'évaluer ou de mesurer le succès d'une technique particulière de rendu (souvent très dépendante de la subjectivité du designer) ou d'un style d'animation pour véhiculer l'identité de quelqu'un.

Dans nos lectures nous avons dégagé trois approches pour l'évaluation de visages virtuels en général. Ces approches, bien qu'elles n'abordent pas directement la problématique de l'évaluation de la fidélité, peuvent nous donner des indications sur les éléments sur lesquels nous devrions nous baser pour formuler des protocoles d'évaluation :

1. Dans la première approche, qui est basée sur l'introspection, les observateurs sont invités à décrire leurs sentiments subjectifs au sujet de certains aspects d'une image. Par exemple, Ring et al. (Ring et al., 2014) ont évalué les attitudes des utilisateurs vers un agent virtuel à l'aide d'un questionnaire basé sur une échelle de Likert (après chaque interaction avec un agent). Des exemples d'aspect étudiés furent par exemple : le réalisme, le charme "appeal", la familiarité, l'étrangeté, la convivialité, la fiabilité, la facilité d'interagir avec, le désir de continuer à travailler avec, l'appréciation (likability), la compassion, la pertinence et la qualité du mouvement.
2. La seconde approche consiste à évaluer la performance d'un utilisateur dans un contexte d'une tâche dépendante. Par exemple plusieurs études (Bartz & StraBer, 2006 ; Gooch, Reinhard, & Gooch, 2004 ; Wallraven et al., 2005) ont évalué l'impact d'un rendu non réaliste sur l'apprentissage de nouveaux visages. La tâche d'identification ou de ressemblance n'est pas directement posée, mais évaluée à travers les scores obtenus grâce à la tâche d'apprentissage de nouveaux visages.
3. Finalement, des travaux tel que (Santella & DeCarlo, 2004) ont présenté une 3eme approche basée sur l'évaluation par suivie du regard. Dans cette étude, Santella et al. ont mené une analyse statistique des zones de fixations sur des scènes 3D avec différentes techniques de rendu. Bien que les résultats semblent prometteurs, nous ne savons pas ce que cette technique mesure réellement et comment nous pourrions comparer ses résultats avec les autres approches.

6.5 Synthèse

Les techniques de rendu récentes arrivent à défier notre capacité à discerner la différence entre des visages générés par ordinateurs et des visages réels (sur la base de photos ou vidéos). Pourtant, en dépit de la maturité que connaît l'informatique graphique en terme

de création et de rendu de visages virtuels, et malgré la tendance actuelle lors de la création de visages photo-réalistes de se baser sur des données réelles (von der Pahlen et al., 2014 ; O. Alexander et al., 2010), très peu d'études évaluent la qualité des résultats obtenus. Cela est d'autant plus vrai pour ce qui concerne l'évaluation de la fidélité du modèle 3D par rapport au référent d'origine. En effet les clones virtuels ont l'avantage de proposer un support de comparaison. Pourtant cette opportunité n'est pas bien exploitée dans le domaine de recherche sur la perception d'humain virtuel.

L'objectif de cet état de l'art est de mettre en exergue deux aspects importants lors de la perception du visage virtuel. Premièrement, la communauté de psychologues soutient que l'expressivité joue un rôle lors d'une tâche d'identification sur base de photos ou de vidéo. Ce rôle prend de plus en plus d'importance quand les informations relatives à l'apparence sont partielles ou altérées. Cependant le rôle de l'expressivité pour l'identification de visage virtuel a été très peu exploré et demeure même incertain selon Bennets et al. (Bennetts et al., 2013). Le deuxième point important est que l'interaction entre la familiarité, l'apparence et l'expressivité joue un rôle crucial dans le processus de perception de visages, mais très peu d'études ont exploré cet aspect dans une étude perceptive d'humain virtuel. Ces caractéristiques doivent impérativement être prises en compte lors de l'évaluation perceptive de la fidélité visuelle d'un clone virtuel.

Étude perceptive pour l'étude de la fidélité visuelle

Sommaire

7.1	Introduction	82
7.2	Mise en place des niveaux de réalisme d'apparence	82
7.3	Méthode	83
7.3.1	Variables	84
7.3.2	Participants	85
7.3.3	Procédure	85
7.3.4	Hypothèses	86
7.4	Résultats	87
7.4.1	Évaluation de la fidélité	87
7.4.2	Impact de la familiarité sur l'évaluation de la fidélité	89
7.4.3	Évaluation de l'appréciation (Likeability)	90
7.5	Discussion et synthèse	92
7.6	Conclusion	93

7.1 Introduction

Plusieurs chercheurs ont souligné les effets des choix de conception visuelle sur la perception d'un agent virtuel par les utilisateurs. Par exemple, beaucoup de travaux ont exploré les effets de différentes stylisations de rendus d'humain virtuel.

Le but des rendus stylisés (artistiques) est de fournir une représentation abstraite d'une image qui met en évidence certaines caractéristiques de l'information visuelle. (Wallraven et al., 2007).

Basés sur les résultats de l'état de l'art du chapitre précédent, nous allons essayer d'extraire les éléments les plus pertinents sur lesquels nous allons nous baser pour définir une échelle de fidélité pour un clone virtuel.

7.2 Mise en place des niveaux de réalisme d'apparence

La question qu'amènent nos travaux est : qu'elle est la meilleure technique de rendu (ou combinaison de techniques) qui favorise la fidélité d'un clone virtuel ?

À travers les différentes approches d'évaluation de visage virtuel qu'il soit réaliste ou basé sur des rendus artistiques, nous avons constaté qu'il est difficile de savoir quel canal d'information est pertinent lors de l'évaluation.

Dans le but d'étudier plusieurs variables impliquées dans l'identification des clones virtuels, nous avons décidé de définir trois degrés de fidélité. L'échelle de fidélité s'inspire de nos conclusions sur l'état de l'art de l'évaluation de la perception des visages virtuels qui soutient que la couleur et donc la texture, les techniques de rendu et la gestion des ombres ont la plus grande influence sur l'évaluation de visages 3D. Par la suite, nous nous sommes basés sur les résultats des travaux de (McDonnell et al., 2012) qui clament que l'avatar Toon-shading (rendu cartoon) et le rendu le plus photo-réaliste (figure-6.1) sont les mieux acceptés et les plus appréciés. Ainsi nous avons inclus deux niveaux similaires en tant qu'élément de notre échelle de fidélité.

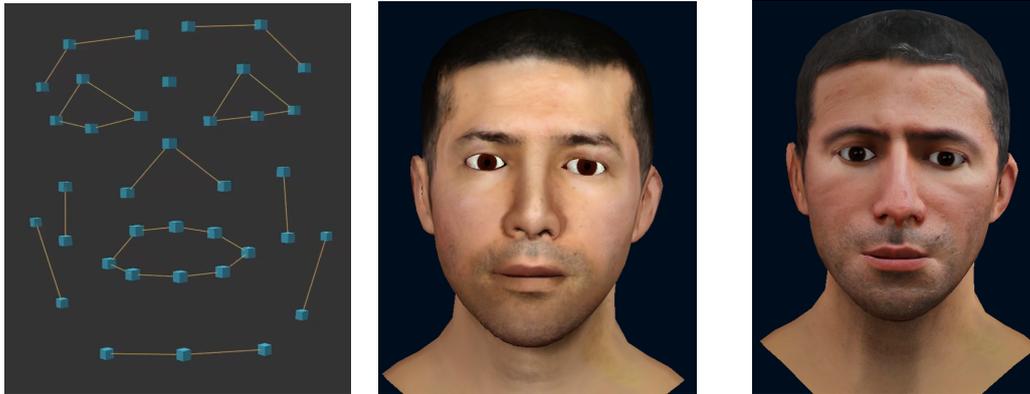
Pour le troisième niveau de notre échelle de fidélité, nous avons opté pour un niveau qui ne véhicule que l'information dynamique : un niveau qui serait dépourvu de toute information d'apparence holistique. Le but de ce niveau de fidélité est d'étudier le poids de l'information expressive seule sur l'évaluation de la fidélité.

Ainsi nous avons défini trois niveaux de fidélité d'apparence comme l'illustre dans la figure-7.1 :

1. (M) Données de capture de mouvements : Ce niveau correspond au plus haut niveau d'abstraction de l'information d'apparence. Nous avons mis en place un niveau de clone virtuel qui se constitue seulement de points correspondant à l'ensemble des marqueurs de capture de mouvements.
2. (SR) Clone semi-réaliste : un niveau intermédiaire d'abstraction s'inspirant des techniques de rendu plutôt cartoon ; les textures ont été peintes à la main grâce au logiciel

”Zbrush”. Les yeux et les techniques de rendu de peau sont volontairement assez basiques.

3. (R) Clone réaliste : Pour ce niveau, nous avons conservé le plus d’informations relatives au référent. La texture est basée uniquement sur les photos réelles du référent. Nous avons également généré des textures de ”normal” et de ”displace” grâce au logiciel ”Zbrush” pour conserver toutes les rugosités de la peau. Des techniques de shading avancé tel que le ”Sub surface scattering” ont été implémentées pour rendre le visage du clone aussi réaliste que possible.



(1) Motion capture data (M) (2) Clone Semi-réaliste (SR) (3) Clone Réaliste (R)

FIGURE 7.1: Exemples de stimuli statiques utilisés dans l’expérimentation

L’objectif de considérer ces trois niveaux de fidélité est d’étudier si un niveau dépourvu d’information d’apparence peut quand même véhiculer un sentiment de fidélité. Par la suite en définissant deux niveaux de réalisme à travers le clone semi-réaliste et le clone réaliste, nous avons pour but de déterminer si la quantité d’information d’apparence supplémentaire est nécessaire pour véhiculer un sentiment accru de fidélité.

7.3 Méthode

La recherche en psychologie suggère que les humains se basent essentiellement sur des informations statiques pour identifier des visages familiers. L’information dynamique ne jouerait un rôle que lorsque l’information visuelle est incomplète (Bennetts et al., 2013). Cette évaluation est fortement liée au degré de familiarité avec le visage observé.

Ces conclusions ont été faites suite à des expériences sur la base de photo/vidéo de visages réels. La question que nous nous posons est de savoir si elles sont applicables sur des visages de clones virtuels. Plus précisément, nous examinons si la dynamique expressive améliore l’évaluation de la fidélité entre le visage du référent et celui du clone virtuel en utilisant différentes techniques de rendu que nous avons définie dans notre échelle de fidélité.

Très peu d'études se sont également penchées sur l'interaction de l'apparence et l'expressivité dans la transmission du sentiment de fidélité. Quelles peuvent être les contributions respectives de chacune de ces modalités pour véhiculer l'information d'une identité ?

Nous avons donc mené une étude perceptive qui tente d'apporter des réponses à trois questions de recherches :

- Quels sont les poids respectifs de l'apparence et de l'expressivité lors du processus de jugement de la fidélité ?
- Dans quelles mesures les changements d'apparence affectent-ils le jugement de la fidélité expressive ?
- Quel est l'impact de la familiarité dans l'évaluation perceptive d'un clone virtuel ?

7.3.1 Variables

Pour notre expérimentation nous avons opté pour un design mixte entre une expérimentation intra/inter sujet.

7.3.1.1 Variable indépendante

Variables intra-sujets. Nous avons défini deux variables intra-sujet. Ces deux variables sont relatives au clone virtuel.

1. **Le rendu** : Ce facteur englobe toutes les techniques utilisées sur les visages du clone virtuel afin d'aboutir aux trois degrés de fidélité définie ci-dessus
 - (M) Les marqueurs de la session de capture de mouvements
 - (S-R) Clone semi-réaliste
 - (R) Clone réaliste
2. **La dynamique** : les clones virtuels ont été présentés sous deux conditions de dynamique comme suit :
 - (S) Une condition statique où le clone a été présenté affichant l'apex d'une émotion.
 - (D) Le clone virtuel est affiché avec une animation d'une émotion sélectionnée.

Ainsi, il y a six conditions (trois niveaux de rendus \times deux niveaux de dynamique). Chaque condition a été répétée six fois.

Variable inter-sujets. Au vu de l'importance de la familiarité dans un processus d'évaluation d'une identité. Nous avons décidé de prendre en considération le degré de familiarité que possèdent les différents participants vis-à-vis du référent humain. Les participants ont été divisés en trois catégories selon leur réponse à un questionnaire leur demandant d'évaluer leur familiarité avec le référent sur une échelle de Likert à sept points.

7.3.1.2 Variables dépendantes

Les variables dépendantes comprenaient deux mesures introspectives. Pour chaque condition, nous avons demandé aux participants d'évaluer :

- La fidélité perçue entre le clone virtuel déployé et le référent humain
- L'appréciation ("Likeability ") du clone virtuel

7.3.2 Participants

Au total, 36 participants ont été inclus dans l'étude. Ils étaient âgés entre 23 et 55 ans (âge moyen = 37.37 ans). Notre échantillon de sujets comprenait 12 femmes et 24 hommes.

- Douze participants étaient non familiers du référent
- Dix participants étaient modérément familiers avec le référent
- Quatorze participants étaient familiers avec le référent

7.3.3 Procédure

Les participants ont été informés qu'ils allaient réaliser une expérience de perception utilisant des clones virtuels. Ils ont été installés dans une pièce où ils étaient seuls à une distance de 50 cm d'un écran de 44,7 x 33,5 cm.

L'expérience a débuté avec un petit questionnaire où il devait renseigner leur âge, leur sexe et nationalité. Nous leur avons également demandé d'évaluer s'il avait l'habitude de percevoir des visages générés par ordinateur sur une échelle de Likert à sept points.

Dans l'étape suivante, les participants ont été invités à évaluer leur familiarité avec le référent sur une échelle de 1 à 7 (1 correspondait à "je ne connais pas cette personne" et 7 correspondait à "j'ai l'habitude de voir cette personne plus de 10h par semaine depuis un an").

Par la suite, les différentes instructions décrivant l'expérience ont été affichées sur l'écran. Nous avons ajouté à la fin des instructions une définition pour éclaircir et désambiguïser nos deux mesures de la fidélité et de l'appréciation "*Likeability*" du clone virtuel.

La phase d'instruction s'achève par une phase de test où nous avons demandé au participant d'évaluer un stimulus donné en exemple.

Par la suite, l'expérimentation réelle débute. Celle-ci est composée de 36 blocs représentant l'ensemble des combinaisons de nos variables indépendantes et où chaque condition a été répétée à six reprises (3 niveaux de rendu × 2 niveaux de dynamique × 6 expressions émotionnelles).

Un bloc est constitué de trois phases. Dans un premier temps une vidéo du référent réel jouant une émotion est affichée à l'écran suivi par un écran noir de transition de 200ms.

Par la suite nous affichons, d'une manière aléatoire, un rendu en temps réel d'un clone virtuel sous les différentes conditions établies.

Dans la dernière phase, après chaque fin de stimuli les participants étaient invités à répondre à un questionnaire évaluant le rendu qu'il venait de visualiser. Les participants étaient donc invités à répondre aux deux questions suivantes sur une échelle continue de 0 à 100.

- Comment évaluez-vous la fidélité entre le clone virtuel et la vidéo réelle du référent humain ?
- À quel point avez-vous apprécié cette version du clone virtuel ?

7.3.4 Hypothèses

Nos hypothèses sont les suivantes :

H1 : Basé sur les résultats des recherches en psychologie et en perception du visage (Lander & Chuang, 2005 ; O'Toole et al., 2002) qui stipulent que les visages sont mieux identifiés lorsqu'ils sont en mouvement, nous émettons l'hypothèse que les scores de fidélité seront plus importants en condition dynamique qu'en condition statique. Cela devrait être d'autant plus vrai que le participant est familier avec le référent. Nous nous attendons à ce que le bénéfice de la dynamique du mouvement sur le jugement de la fidélité soit plus important chez les participants familiers avec le référent grâce à la reconnaissance des mouvements idiosyncrasiques du référent.

H2 : Nous avons vu que quand les deux informations d'apparence et de dynamique sont disponibles, l'humain se base essentiellement sur l'information d'apparence. Nous avons vu également que l'information dynamique joue un rôle uniquement quand l'information d'apparence est incomplète sur les vrais visages. Ainsi, nous nous attendons à ce que plus le rendu est de qualité inférieure plus l'information dynamique a un impact sur le jugement de la fidélité. Par conséquent, nous émettons l'hypothèse que la différence de notation entre la condition statique et dynamique sera plus importante quand le niveau de rendu est bas.

H3 : Basé sur les études autour de "l'uncanny valley" (Mori, 1970) et tenant compte du niveau d'abstraction de nos animations (dû à des limitations techniques et le nombre de marqueurs de capture de mouvement). Nous émettons l'hypothèse suivante concernant l'appréciation ("Likeability") du clone virtuel : nous nous attendons à un phénomène de rejet au niveau de rendu réaliste pour la catégorie des familiers. En effet nous supposons qu'à cause de leur connaissance accrue du référent ils seront plus sensibles aux différences du niveau de réalisme entre le rendu d'apparence et le réalisme de l'animation.

Nous émettons donc l'hypothèse que les participants qui sont familiers avec le référent vont noter la "likeability" du clone virtuel dans l'ordre décroissant suivant :

Condition 4 : Rendu semi-réaliste - condition dynamique (SR-D) ;

Condition 3 : Rendu réaliste - condition dynamique (RD) ;

Condition 6 : Rendu réaliste - condition statique (SR-S) ;

Condition 5 : Rendu réaliste - condition statique (RS) ;

Les participants qui ne sont pas familiers avec le référent devraient évaluer la "likeability" dans un ordre différent, car ayant une connaissance moindre du référent et de ses expressions particulière, ils devraient avoir des exigences plus faibles notamment vis-à-vis du rendu réaliste. Ainsi, les participants qui ne sont pas familiers avec le référent devraient

noter la "likeability" dans cet ordre décroissant :

Condition 3 : Rendu réaliste - condition dynamique (RD) ;

Condition 4 : rendu semi-réaliste - condition dynamique (SR-D) ;

Condition 5 : Rendu réaliste - condition statique (RS) ;

Condition 6 : Rendu réaliste - condition statique (SR-S) ;

7.4 Résultats

Pour analyser les notations de la fidélité et de l'appréciation des clones virtuels, nous avons effectué une anova à trois facteurs (rendu \times dynamique \times familiarité) à mesures répétées. Les effets d'interaction ont été analysés avec des tests posthoc.

7.4.1 Évaluation de la fidélité

Les résultats de l'ANOVA ont démontré un effet principal sur les variables de rendu et de dynamique.

Pour les effets du rendu les résultats de l'ANOVA étaient $F(2, 189) = 46.284, p < 2e - 16$. Une analyse post hoc des interactions entre les différents niveaux de rendu a démontré une différence significative entre le niveau marqueurs de capture de mouvement(M) et les deux autres niveaux (R et SR).

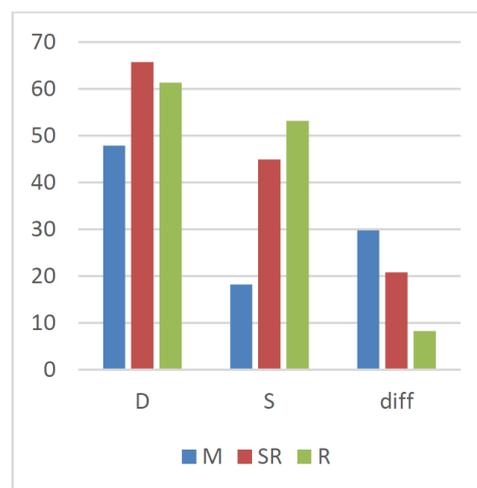


FIGURE 7.2: Impact de la variable (D)dynamique sur l'évaluation de fidélité

Selon les tests du Tukey HSD :

Le résultat de l'interaction entre le niveau (R) clone réaliste et le niveau (M) données de capture de mouvement avait pour p-value ($p < 0 : 005$).

L'interaction entre clone semi-réaliste (SR) et donnée de capture de mouvement (M) était

($p < 0 : 05$).

Ces résultats démontrent l'importance de l'information de l'apparence dans les tâches d'évaluation de la fidélité. Cependant, il n'y avait pas de différence significative entre le rendu semi-réaliste et le rendu réaliste.

En ce qui concerne la variable dynamique, les résultats d'analyse de variance étaient $F(1; 189) = 110,495; p < 2e - 16$. Comme le montre la figure 7.2, les stimuli en condition dynamique ont reçu des scores de fidélité plus élevés que sous la condition statique.

Nos résultats ont également démontré des effets significatifs de l'interaction entre le rendu et la dynamique : $F(2, 52) = 11,844; p = 5,77e - 05$. Par conséquent, comme indiqué dans l'hypothèse H2, plus le niveau de rendu est bas plus l'information dynamique avait un impact important dans la notation de la fidélité. En effet, comme l'illustre la figure 7.2, nous pouvons constater que la différence dans la notation de la fidélité entre la condition dynamique et statique pour le même niveau de rendu est plus importante pour les données de capture de mouvement (M) que pour le clone semi-réaliste (SR) qui est elle-même plus importante que pour le clone réaliste (R).

Nous avons également constaté que, sous la condition statique, la fidélité du rendu réaliste (R) a été notée comme plus élevée que celle du rendu semi-réaliste. En revanche, sous la condition dynamique cette classification s'inverse. Le rendu semi-réaliste est jugé plus fidèle que le rendu réaliste.

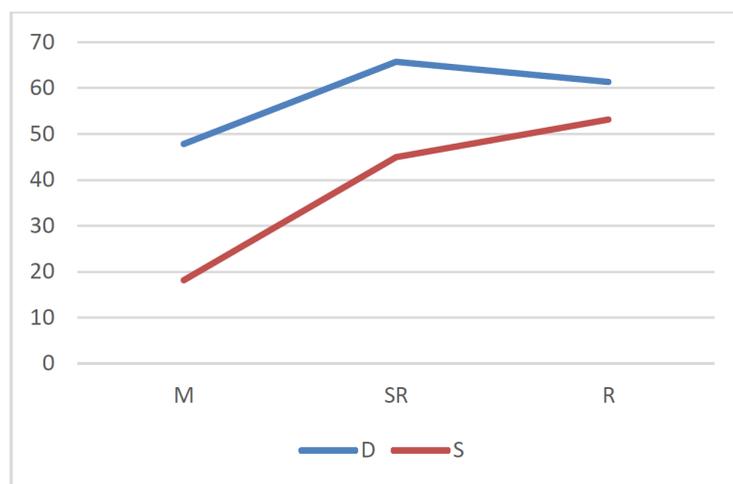


FIGURE 7.3: Effet de l'interaction entre les variables dynamique(D) et rendu(R) sur l'évaluation de fidélité

Une analyse post-hoc de l'interaction entre le rendu et la dynamique a démontré qu'il n'y avait pas d'effet significatif entre le rendu semi-réaliste (SR) et le rendu réaliste (R) dans la condition dynamique ($p = 0.77$).

Cette similarité dans la notation de fidélité pour ces deux niveaux qui sont pourtant bien distincts visuellement, peut être expliquée par le fait que les utilisateurs ont basé leur

jugement de la fidélité sur les informations dynamiques. En effet, les expressions faciales affichées pour les différents niveaux de rendu ont été obtenues à partir des mêmes données de capture de mouvement. Ainsi la différence de niveau dans les indices d'apparence a pu être inhibée par une information dynamique similaire.

La figure 7.3 montre également que l'évaluation de la fidélité pour le rendu réaliste était inférieure aux rendus semi-réalistes sous les conditions dynamiques. Cela peut s'expliquer par un effet lié à la théorie de l'uncanny valley (Mori, 1970) vu que l'apparence et l'expressivité ne sont pas au même niveau de réalisme.

7.4.2 Impact de la familiarité sur l'évaluation de la fidélité

Les résultats de l'ANOVA ont démontré un effet principal de la familiarité sur l'évaluation de la fidélité $F(2, 156) = 8,754, p = 0,0002$.

Les tests de post hoc ont démontré une différence significative entre les groupes des familiers et des modérément familiers avec le groupe des non-familiers. Les scores obtenus sont respectivement ($p = 0.011$) et ($p = 0.017$). Par contre, il n'avait pas de différence significative dans l'évaluation entre le groupe des familiers et le groupe des modérément familiers.

Cela suggère que nous avons la capacité de reconnaître des expressions idiosyncrasiques sans forcément être extrêmement familiers avec la personne : quelques interactions peuvent être suffisantes pour assimiler la signature dynamique d'une personne.

Les résultats des tests Post Hoc ont montré une interaction significative entre la familiarité et le niveau de rendu : $F(2, 52) = 4.93, p = 7.10e - 04$. De plus, comme l'illustre la figure 7.4, une plus grande familiarité avec le référent conduit à une évaluation plus sévère de la fidélité du clone. Plus notre connaissance de la personne est importante plus nous serions exigeants lors la notation de la fidélité.

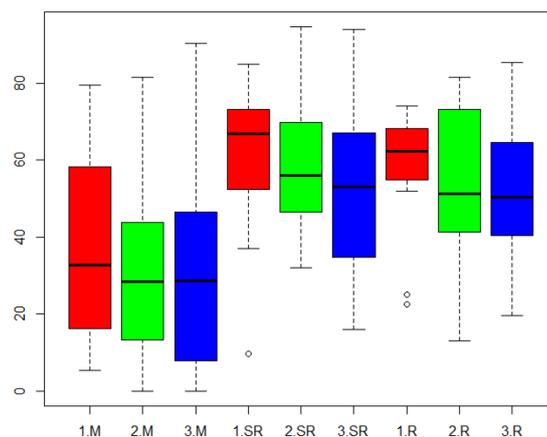


FIGURE 7.4: Évaluation de la fidélité sur la base du rendu et de la familiarité avec le référent(1= non-familier, 2= modérément familier, 3= très familier)

Nous avons également constaté (7.4) que notre hypothèse H1 a été confirmée. En effet, la contribution du mouvement est d'autant plus grande que le niveau de familiarité avec la personne est important.

En outre, les tests de post hoc de chaque sous-ensemble de la familiarité ont montré que les différences entre les conditions (R-S et M-D, SR-S et M-D) ne sont pas significativement différentes au sein du groupe des très familiers contrairement aux autres groupes. Cela pourrait suggérer que l'expressivité pourrait compenser les limitations dues aux pertes des informations visuelles et véhiculer un sentiment équivalent de fidélité auprès de la catégorie des gens qui connaissent très bien le référent.

Pour le groupe des non familier, les évaluations de la fidélité étaient équivalentes pour les stimuli (SR-D et SR-S) et entre (R-D et R-S). L'expressivité joue donc ici un rôle moins important dans l'évaluation de la fidélité.

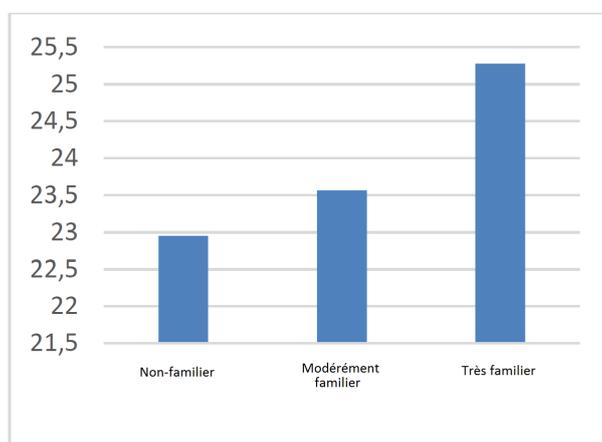


FIGURE 7.5: La différence entre l'évaluation de la fidélité dans la condition statique et la condition dynamique par groupe de familiarité

7.4.3 Évaluation de l'appréciation (Likeability)

Les résultats de l'analyse de la variance ont démontré un effet principal des variables rendu, dynamicité et familiarité. Les résultats sont respectivement $F(2, 144) = 117,635, p < 2e - 16$, $F(1, 144) = 56,640, p = 5,22e - 12$, and $F(2, 144) = 5,693, p = 0,003$.

L'effet de l'interaction entre le rendu et la variable dynamique était aussi significatif $F(2, 48) = 3,765, p = 0,03$. Cependant, l'interaction entre la variable rendu et familiarité n'était pas significative ($p=0,08$).

Nos résultats suggèrent, comme l'illustre la figure 7.6, que la dynamique améliore les taux d'évaluation de fidélité pour l'ensemble des niveaux de rendu.

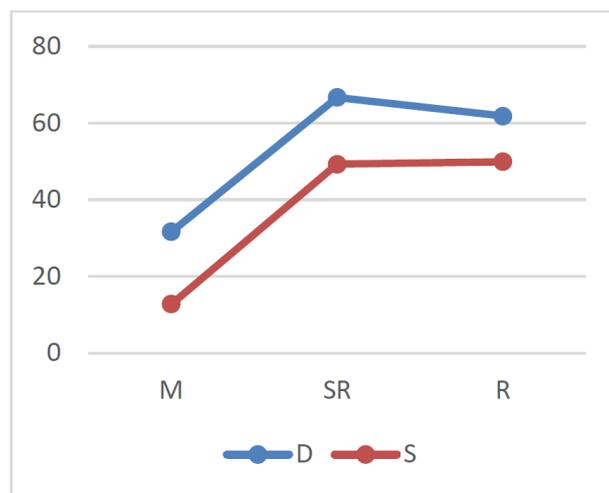


FIGURE 7.6: Interaction du rendu et de la dynamique dans le jugement de la "Likeability"

La figure 7.7 ci-dessous présente l'impact de la familiarité vis-à-vis du référent sur les taux d'évaluation de la likeability avec les différents niveaux de rendu. Nous pouvons constater que le groupe des très familiers préfère le rendu semi-réaliste (SR). Les deux autres groupes (modérément familier et non familier) ont attribué au niveau réaliste le plus haut score de "Likability".

Ces résultats restent du moins à confirmer par des plus profondes investigations. En effet, la différence de notation entre le niveau SR et R est très légère. Les tests posthoc n'ont pas démontré de différence significative de l'interaction entre la familiarité et le rendu rendu. Nous avons seulement obtenu une tendance avec $p=0.08$. Par conséquent, plus de tests avec un plus grand nombre de sujets sont nécessaires pour confirmer ces résultats.

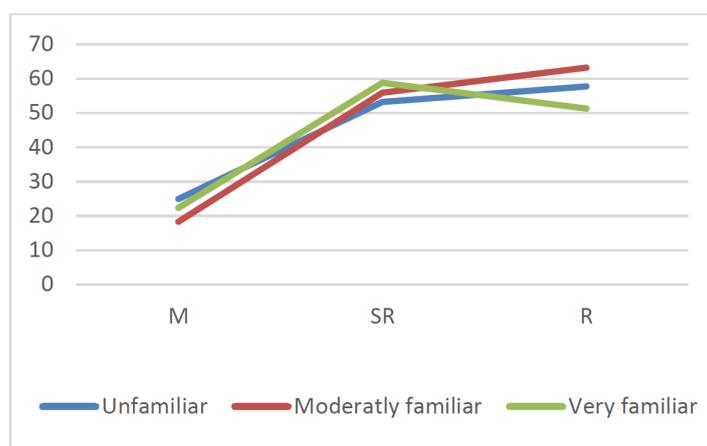


FIGURE 7.7: L'impact de la familiarité sur l'évaluation de la "Likeability"

7.5 Discussion et synthèse

Cette évaluation confirme la spécificité de la perception d'un visage du clone virtuel. En effet, contrairement aux études classiques sur la perception d'humain virtuel anonyme, l'affichage de visage virtuel représentant une identité implique des variables externes spécifiques telles que la familiarité du participant avec le référent.

Nous avons constaté qu'une partie de nos résultats confirment et sont en accord avec certains résultats qu'on retrouve dans la littérature psychologique.

En effet, nous avons confirmé l'hypothèse que l'expressivité renforce un jugement de la fidélité. Les stimuli en condition dynamique jouissent de taux de fidélité supérieurs à une expression statique. Ces résultats sont dans la continuité des études psychologiques telles que ceux de Lander et al. (Lander et al., 1999).

En modifiant le degré d'abstraction de l'information d'apparence, nous avons également remarqué que l'impact de l'expressivité diffère selon la qualité de l'information d'apparence mise en place

Nous avons confirmé que plus l'information d'apparence est altérée et incomplète plus l'information dynamique jouait un rôle important.

De plus en nous basant sur un niveau dépourvu presque entièrement d'information d'apparence (niveau de marqueur de capture de mouvement (M)), nous avons démontré que l'information dynamique seule peut véhiculer un sentiment de fidélité. Par conséquent, les signatures dynamiques idiosyncrasiques du référent jouent un rôle important lors d'une tâche d'évaluation de visages virtuels. Ces résultats sont cohérents avec l'hypothèse de "l'information supplémentaire" vue précédemment dans l'état de l'art (O'Toole et al., 2002).

Cependant, après un débriefing informel avec les participants suite à l'expérience, environ 65% de ces derniers n'ont pas remarqué qu'il y avait trois niveaux de rendu. Les participants ont estimé que les niveaux de rendu semi-réaliste et réaliste étaient identiques. Les résultats des analyses statistiques ont aussi démontré que la perception de ces deux différents niveaux de rendu ne sont pas significativement différentes. Suite à cela, nous pouvons émettre l'hypothèse que les participants ont donné plus d'importance à la modalité expressive pour l'évaluation. Ainsi, contrairement à ce qu'avancent les études plus classiques sur des visages réels, l'expressivité a été la principale modalité dans le jugement de la fidélité. Comme suggéré par (Ring et al., 2014), la perception d'un visage virtuel est fortement dépendante de la tâche. Pour notre tâche bien particulière d'évaluation de la fidélité sous les conditions de notre protocole où on présente d'abord une référence dynamique, l'information visuelle devient secondaire.

Nous pensons donc que l'hypothèse qui stipule que l'information dynamique ne joue qu'un rôle secondaire dans l'identification et l'évaluation de la fidélité est relative. Les modalités sur lesquelles nous nous basons peuvent dépendre du contexte et de la tâche que nous testons.

De plus, nos résultats suggèrent que la familiarité avec le référent a un impact im-

portant sur l'évaluation de la fidélité et l'appréciation d'un visage de clone virtuel. Ainsi, les participants qui étaient familiers avec le référent étaient plus sensibles à la condition dynamique. En effet plus l'utilisateur était familier avec le référent plus l'écart de notation du même niveau de rendu sous les conditions dynamiques vs. statiques étaient important. Cela confirme l'avantage que procurent les caractéristiques idiosyncrasiques de l'expression faciale qui sont forcément plus connu des personnes familières.

Une autre conclusion, formulée suite à cette étude, concerne la relation entre niveaux de réalisme dynamique et le réalisme de l'apparence. Nous avons en effet constaté que l'appréciation de l'avatar est le jugement de sa fidélité par rapport au référent sont fortement corrélé. Par conséquent, si le stimulus a tendance à inspirer un effet relative à la théorie de l'uncanny valley du a une disparité entre le niveau de réalisme visuel et de réalisme dynamique, l'appréciation va sans aucun doute diminuer. Et cela affectera l'évaluation de la fidélité.

Nos résultats suggèrent de plus que cette hypothèse est d'autant plus vraie que la personne est plus familière du référent. Par exemple nous avons observé que la condition avec les clones virtuels les plus réalistes en condition dynamique est bien accueillie par les participants les plus familiers avec le référent.

7.6 Conclusion

Nous avons étudié la manière avec laquelle les participants perçoivent des clones virtuels d'une personne existante. En nous basant sur notre cadre conceptuel pour l'étude de la fidélité et sur l'état de l'art de l'évaluation de visages virtuels, nous avons proposé plusieurs variables comme étant les plus pertinentes pour l'étude de la fidélité. Par la suite, à partir de ces variables nous avons défini 3 niveaux de réalisme d'apparence.

Nos résultats fournissent de nouvelles perspectives sur la manière avec laquelle les humains perçoivent un visage virtuel reconnaissable. En particulier, nos résultats suggèrent que nous devrions donner la même importance aux indices dynamiques que ceux de l'apparence pour renforcer le sentiment de fidélité chez des observateurs.

Nous avons également constaté que l'affichage des niveaux équivalents de réalisme d'expressivité et de rendu est essentiel pour la transmission de l'impression de fidélité. Un décalage du niveau de réalisme d'une modalité par rapport à l'autre nuit au sentiment de "likeability" qui s'est avéré être fortement corrélé au jugement de fidélité. Autrement dit, si nous déprécions un rendu il est perçu comme étant moins fidèle.

Nos résultats peuvent aussi suggérer des pistes de solutions pour adapter des rendus de faibles résolutions dans des conditions de ressource limitée. Un visage de clone virtuel avec une capacité de rendu limité peut être compensé par l'expressivité pour donner un plus grand sentiment de fidélité. En effet, un rendu simpliste comme les points de marqueurs de capture de mouvement peut véhiculer un sentiment de fidélité si son expressivité est fidèle à celle du référent.

Nous considérons que notre analyse de la perception de clones virtuels peut contribuer à

plusieurs domaines. En effet, la compréhension des facteurs importants impliqués dans la perception de la fidélité peut aider les artistes infographistes à créer des clones virtuels plus reconnaissables.

Nos travaux pourront également aider les chercheurs en sciences cognitives et psychologique à identifier les mécanismes perceptifs et cognitifs qui sous-tendent la reconnaissance du visage en permettant le contrôle précis des différentes caractéristiques et variables d'un visage virtuel et explorer différents degrés de rendu.

Limite et perspective. Dans notre conception expérimentale, nous avons simplement fait varier la familiarité vis-à-vis du clone virtuel dans notre échantillon de participants. D'autres différences inter-individuelles devraient également être prises en considération. Par exemple l'âge, le sexe, l'origine ethnique et l'expérience préalable avec des médias virtuels sont des variables potentiellement importantes pour l'évaluation de la fidélité. De nombreuses études antérieures ont prouvé qu'il avait un impact important sur l'évaluation de visages virtuels (Fan et al., 2012 ; Sinha et al., 2006).

En outre, nous ne considérons pas les conditions environnementales telles que différents points de vue ou des changements de l'éclairage. De même, nous ne considérons pas la variabilité d'expression telle que d'autres techniques d'animation faciales, la vitesse de la dynamique de l'expression, etc.

Dans des recherches futures, nous visons à établir un profil d'expression qui décrit la signature de mouvement d'un référent. Nous aurions à faire varier les composantes de ce profil pour identifier les principaux indices qui assurent la fidélité expressive.

Nous prévoyons également de considérer d'autres clones virtuels pour évaluer la fidélité dans un contexte dépendant d'une tâche (par exemple, une tâche de reconnaissance avec des identités différentes en fonction de plusieurs clones virtuels).

Quatrième partie

Vers une fidélité comportementale

État de l'art : Fidélité cognitive

Sommaire

8.1	Introduction	98
8.1.1	Contexte	98
8.1.2	Motivation	98
8.1.3	Problématiques	99
8.2	Les différentes définitions d'une identité psychologique	100
8.2.1	Identité vs self	100
8.2.2	Identité vs personnalité	101
8.2.3	Synthèse	101
8.3	La structure affective de l'identité	102
8.3.1	L'émotion	102
8.3.2	Différences inter-individuelles dans l'expérience de l'émotion	103
8.4	Approches, théories et modèles des émotions en psychologie	106
8.4.1	Approche catégorielle	106
8.4.2	Approche dimensionnelle	106
8.4.3	Approche cognitive	107
8.4.4	Synthèse	110
8.5	Modèle computationnel des émotions	111
8.5.1	Modèle de génération d'expression faciale basée sur la théorie du CPM	113
8.6	Synthèse	114

8.1 Introduction

8.1.1 Contexte

Les clones virtuels sont de plus en plus répandus dans diverses interfaces homme-machine. Après avoir été largement utilisés dans l'industrie du cinéma et les cinématiques dans des jeux vidéos, ils intègrent maintenant de plus en plus des applications interactives en temps réel. Les exemples sont multiples de célébrités dont le clone virtuel est interactif dans un jeu vidéo.

De plus, dans le but d'atteindre un certain degré de réalisme (von der Pahlen et al., 2014), la tendance actuelle pour modéliser des humains virtuels, même plus anonymes, est de se baser sur des techniques d'acquisition de données réelles (issues de scans et de sessions de capture de mouvement). Les clones virtuels seront donc de plus en plus amenés à être une source d'interaction.

Pourtant, très peu d'études s'intéressent à doter des agents virtuels d'un comportement personnalisé. En effet, l'approche actuelle de l'IA tend en général à définir un comportement normatif moyen qui a pour aboutissement de simuler un comportement générique, ou du moins celui d'un groupe (un groupe social, ethnique, etc.). Les travaux d'IA actuels ont de surcroît tendance à assimiler l'identité à un rôle d'un agent (par exemple un rôle de pompier dans un système multi agent) (Di Loreto et al., 2012).

Plusieurs études en neurosciences témoignent de l'importance et la nécessité de s'intéresser à doter un clone virtuel d'un comportement cognitif fidèle. En effet, certaines études neurologiques (Gobbini & Haxby, 2007) soutiennent que l'aspect visuel n'est qu'une composante de la reconnaissance des personnes familières. D'autres aspects fondamentaux comprennent aussi le traitement des connaissances que nous avons sur cette personne tels que les traits de personnalités et la réponse émotionnelle que nous éprouvons face à cette personne. En effet selon les mêmes auteurs, des études neurologiques suggèrent que des zones du cerveau qui sont associées à la théorie de l'esprit et aux processus cognitifs et sociaux s'activent lors d'une tâche de reconnaissance d'une personne familière. L'évaluation comportementale de l'individu serait donc aussi importante que l'évaluation de son aspect visuel dans une tâche d'identification.

8.1.2 Motivation

Ce chapitre a pour but d'explorer la dernière composante des dimensions de la fidélité définie dans notre cadre conceptuel : *la fidélité psychologique*. En effet, nous avons vu que des efforts considérables sont réalisés pour modéliser l'apparence de clones virtuels. Pourtant, très peu de travaux se sont intéressés à les munir de comportements autonomes.

Ce décalage entre l'aspect visuel d'un humain virtuel et l'aspect simulation de comportement pourrait amener un sentiment de non-congruence. Ce sentiment est à la base de la fameuse théorie de l'uncanny valley (Mori, 1970).

Ce constat est d'autant plus vrai, qu'une autre variable entre en jeu avec l'apparition de

clones virtuels. Cette variable est le degré de *familiarité* que nous avons vis-à-vis du référent. Nous avons constaté dans notre expérience perceptive sur l'évaluation de la fidélité visuelle que la familiarité avec le référent a un impact sur le jugement de la fidélité et l'appréciation du clone. Il s'est avéré que plus notre degré de familiarité est important avec le référent plus nous sommes exigeant sur la fidélité et la qualité du rendu.

Une des hypothèses émises dans le cadre de l'étude de *l'uncanny valley* est l'implication d'autres dimensions autres que visuelles dans l'émergence du sentiment de rejet. Par exemple, Groom et al. soutiennent que l'évaluation du réalisme comportementale joue un rôle important dans l'acceptabilité et l'appréciation des agents (Groom et al., 2009).

La question additionnelle qu'amènent nos travaux est : Est-ce que l'évaluation de cette dimension comportementale ne relève-t-elle pas aussi de la familiarité avec le référent ? Autrement dit, est-ce qu'un comportement non congruent aux comportements du référent (en contradiction avec son comportement dans la vie réelle) peut provoquer également un sentiment de rejet ?

Par ses données graphiques, un clone virtuel est porteur d'une identité bien précise que nous pouvons éventuellement connaître dans la vie courante. Une approche classique de l'IA qui se base sur la création de profils moyens (normatif) pourrait générer un sentiment d'étrangeté dû à l'incongruence entre un visuel individuel et une IA générique. Nous pensons donc que les modèles d'IA actuels pourraient être inadéquats et aboutiraient à une altération du sentiment de fidélité du clone virtuel par rapport à son référent.

Pour illustrer cette question, prenons l'exemple d'un clone virtuel de "Gandhi" auquel on associe un modèle cognitif simulant une personne colérique. Les comportements issus de ces modèles peuvent être réalistes et parfaitement plausibles sur un agent anonyme, mais est ce que ces comportements créeront un sentiment d'incohérence avec la représentation mentale que nous possédons du personnage de Gandhi ?

8.1.3 Problématiques

Nous avons donc vu la nécessité de modéliser un comportement autonome qui prend en considération certains éléments psychologiques du référent. Notre but est de conserver quelques informations d'identité dans un comportement qui soit cohérent avec l'identité visuelle.

Cette problématique soulève plusieurs questions de recherches. Quels sont les aspects comportementaux qui sont déterminants d'une identité ? Quels éléments devront nous choisir pour modéliser un comportement qui traduise quelques aspects du comportement réel du référent ?

Cette thèse aborde le comportement non verbal émotionnel. Nos efforts se sont par conséquent focalisés sur l'étude de la structure affective de l'identité et son expression à travers les expressions du visage. Nous nous intéressons par conséquent à définir ou adapter un modèle computationnel individuel des émotions pour guider l'animation expressive de manière à ce qu'elle soit fidèle aux réactions du référent.

De cette problématique découlent plusieurs questions de recherche : simuler une interaction humaine affective implique de doter l'agent d'une représentation interne d'états émotionnels

et des raisonnements associés et de permettre à l'agent d'exprimer son état émotionnel interne. Cela implique de choisir parmi les théories et modèles proposés en psychologie les aspects les plus pertinents qui décrivent le mieux un comportement émotionnel expressif fidèle. Et par la suite, comment concevoir un modèle computationnel fonctionnel à partir d'un modèle théorique, souvent sous-spécifié, issu de la psychologie tout en préservant quelques composantes de sa psychologie qui permettrait son identification ?

8.2 Les différentes définitions d'une identité psychologique

Dans un premier abord, il nous fallait définir et extraire les éléments déterminants qui nous permettraient d'évaluer le comportement d'autrui pour la reconnaissance d'une identité. Autrement dit, sur quels aspects autres que visuels, nous nous basons pour reconnaître une identité familière ? Pour répondre à cette question, nous allons explorer les différentes définitions données à l'identité par différentes perspectives de recherche.

L'étude de l'identité a été très largement explorée en Sciences Humaines, notamment en science sociale. "L'identité sociale est le point où se focalisent les composantes sociales et psychologiques à l'intérieur **d'une structure affective** et cognitive de la personnalité" (Tajfel, 1981). Elle donne lieu à un type de représentation de soi et à des échanges avec le monde social qui l'entoure en fonction de l'image que chacun se fait de lui même (Levita & Ruemke, 1965).

Le concept d'identité développé en psychologie sociale explore la manière avec laquelle se construit l'image que nous avons de nous-mêmes par rapport aux contextes sociaux dans lesquels nous évoluons et des apprentissages sociaux dans lesquels nous sommes impliqués (Ashforth & Mael, 1989).

La sociologie n'a pas réussi à définir de distinction précise entre l'identité individuelle et l'identité collective. Les deux sont parfois opposées et parfois confondues.

8.2.1 Identité vs self

Quand nous explorons les définitions d'une identité, on se retrouve souvent face à la notion de *self*. Dans un sens général, les termes « Self » et « Identité » sont souvent interchangeable (Bosma and Gerlsma, 2003). Il s'agit d'un processus psychologique de représentation qui se traduit par le sentiment d'exister en tant qu'être singulier et d'être reconnu comme tel par autrui. Ces deux concepts sont fortement corrélés. Beaucoup de recherches (Goffman, 2009) sur le self considèrent l'identité comme une définition, une interprétation du self. L'identité peut être définie comme l'interprétation des différentes composantes du self à un instant donné.

Le self social par exemple est constitué selon (Markus & Nurius, 1986) de l'interaction de plusieurs caractéristiques intrinsèques propres au self tel que :

- La connaissance de soi "Self knowledge"

- Le concept de soi "Self concept"
- Les intentions et objectifs à long terme

Les travaux sur le concept de self (Smith, 2011) mettent en relief que les individus n'ont pas de réponses dépendantes uniquement du contexte, mais avancent que leurs réponses sont aussi modulées par des composantes qui les définissent à un moment donné (Oyserman, 2007, 2009a, Smeesters, Wheeler, Kay, 2010).

8.2.2 Identité vs personnalité

Les concepts d'identité, de soi et de personnalité sont souvent associés. Ce sont des notions très proches les unes des autres, elles sont ardues à définir et sont souvent employées comme synonymes d'après (Bégin, Bleau et Landry, 2000). La personnalité a été longuement étudiée dans la littérature. Quelques définitions de la personnalité suggèrent qu'elle peut être le concept qui assure notre caractère unique. Par exemple Allport et al. ont défini la personnalité comme étant "l'organisation dynamique des différents systèmes psychologique au sein de l'individu qui détermine ses comportements et ses pensées caractéristiques" (Allport, 1961).

Par conséquent dans la personnalité réside aussi l'individualité de nos réactions émotionnelles. Il y a eu de nombreux travaux sur les liens entre personnalité / individualité et agents virtuels (Damian, Endrass, Huber, Bee, & André, 2011 ; Arellano et al., 2011).

8.2.3 Synthèse

Nous pouvons retenir de ces études que l'identité est un concept complexe. Il est parfois assimilé à la personnalité et au concept de self.

Il n'y a pas de consensus quant aux phénomènes auxquels ce terme se réfère. En effet, l'identité peut être perçue comme une identité sociale, ethnique, personnelle, etc.

Fearon,(Fearon, 1999) dans un examen de cette définition changeante de l'identité, affirme qu'elle est actuellement considérée comme étant :

- (a) Une catégorie sociale, définie par les règles d'adhésion, d'attribut caractéristique et de comportements attendus (présumé par les autres)
- (b) Une interprétation du soi, la façon dont un individu s'appréhende cognitivement lui-même.
- (c) la réunion des deux.

La question de l'identité demeure donc une question complexe et vaste. Elle a été abordée selon plusieurs perspectives. En effet, l'identité peut être perçue comme une identité sociale, ethnique, personnelle, etc.

Dans le cadre de nos travaux pour caractériser une identité dans une communication non verbale nous avons décidé de focaliser nos recherches sur les différences inter-individuelles dans la structure affective de l'identité et d'examiner ses liens avec le concept de self et de personnalité.

Un autre point important que nous retenons de ces définitions, c'est que l'identité est avant tout un concept social. Comme précisé par De Levita (Levita & Ruemke, 1965), la question qui définit l'identité n'est pas "Qui suis-je?", mais "Qui suis-je aux yeux des autres?".

Nous retenons surtout que l'identité intègre une structure affective dans sa composition. C'est sur cet aspect que nous allons focaliser nos travaux afin de pouvoir imaginer un modèle computationnel guidant l'expressivité émotionnelle d'un clone virtuel.

8.3 La structure affective de l'identité

Dans le cadre de nos travaux pour caractériser une identité dans une communication non verbale, nous avons décidé de focaliser nos recherches sur les différences inter-individuelles dans la structure affective de l'identité.

Comme nos travaux de recherche se focalisent sur l'interaction non verbale avec un visage virtuel d'un clone, nous allons nous restreindre à étudier la notion de l'identité et l'inter-individualité des expressions faciales émotionnelles.

8.3.1 L'émotion

Chacun sait ce qu'une émotion, jusqu'à ce qu'on lui demande d'en donner une définition. À ce moment-là, il semble que plus personne ne sache. (Fehr et Russell, 1984).

Avant de s'intéresser aux différences inter-individuelles qui peuvent exister chez les humains lors du traitement émotionnel, il nous fallait dans une première étape définir ce qu'est une émotion. Une grande confusion règne sur la terminologie de l'émotion, ceci est en grande partie dû à la nature même de l'émotion. L'émotion étant une part importante de nos vies quotidiennes elle conduit à des interprétations intuitives de ces termes qui sont souvent différentes de leurs définitions données par la recherche en psychologie. De plus, même dans la communauté de la recherche, il n'existe pas un vrai consensus sur la définition donnée aux émotions (S. Marsella, 2010). Nous allons par conséquent aborder cette notion du point de vue de son rôle.

L'émotion joue un rôle primordial dans la vie de l'être humain, en effet, l'émotion

influence le comportement humain de différente manière.

D'un point de vue introspectif : La signification émotionnelle d'un stimulus perçu module le déroulement des processus cognitifs tels que la perception, attention et la prise de décision (Damasio, 1994, Phelps, 2006, Clore, 2009, Loewenstein, 2003, Gros, 2010) (Rodríguez & Ramos, 2014). Ce phénomène appelé régulation affective permet aux individus de développer des réponses motrices dirigées par l'émotion ainsi que de concentrer leurs ressources internes sur les éléments les plus saillants dans l'environnement.

Extérieurement, l'état émotionnel des individus détermine la configuration de leurs expressions faciales, des postures du corps, et l'intonation de la voix lors de l'interaction avec les autres, révélant, par le comportement non verbal, leur état affectif interne et attitudes à l'égard des situations, des objets et des gens (Scherer, 2003, Planalp, 1996).

8.3.2 Différences inter-individuelles dans l'expérience de l'émotion

Dans notre étude nous nous focalisons sur les éléments émotionnels distinctifs d'une personne. Nous allons essayer d'établir le lien qui existe entre un individu en particulier et ses réactions émotionnelles. Nous cherchons à comprendre les processus qui font que nous avons des réactions émotionnelles bien distinctes, voire antagonistes, face à un même événement.

Des preuves neurologiques ont été avancées dans plusieurs travaux sur les différences interindividuelles de l'expérience émotionnelle (Hamann & Canli, 2004). Un stimulus émotionnel donné peut évoquer un large éventail de réactions émotionnelles entre les individus, et ces différences individuelles peuvent s'observer sur les bases neuronales du traitement des émotions. (Hamann & Canli, 2004).

La cognition sociale désigne l'ensemble des processus cognitifs (perception, mémorisation, raisonnement, etc.) impliqués dans les interactions sociales chez l'homme (Fiske & Taylor, 2013). Ces éléments jouent un rôle important dans l'émergence des émotions. Le caractère unique de la réponse émotionnelle y réside, en effet chacun a une expérience propre et des déclencheurs émotionnels qui résultent de ses expériences passées.

Au-delà des traits caractéristiques dispositionnelles et d'adaptations, les vies humaines varient par rapport à la manière dont ils intègrent leurs expériences passées. Les individus construisent une identité et un récit personnel "*Integrative Life Narratives*" qui influence la manière dont ils traitent un événement émotionnel (McAdams & Pals, 2006).

8.3.2.1 Profil émotionnel

Depuis quelques années, de nombreux travaux se sont intéressés à l'étude des caractéristiques émotionnelles stables des individus. Ces recherches avancent qu'il existe des traits émotionnels sur lesquels les individus diffèrent. Les recherches psychologiques et en neuroscience ont rarement pris en compte ces caractéristiques intrinsèques aux individus. Pourtant ces caractéristiques peuvent largement moduler les effets des émotions. En plus de ces avancées théoriques, de nombreuses mesures de ces caractéristiques émotionnelles individuelles ont été développées (Zenasni et al., 2003).

”Un trait émotionnel se définit comme une caractéristique émotionnelle intrinsèque à l'individu. Comme pour les conduites de la personnalité, ces traits sont considérés comme une disposition interne, relativement générale et permanente, plus ou moins marquée selon les individus et ayant une valeur explicative. Le trait s'oppose donc à la notion d'état émotionnel qui fait référence à une situation très transitoire de l'individu dans un contexte spécifique. Ce sont des dispositions ou des tendances à réagir avec une ou plusieurs émotions.” (Zenasni et al., 2003)

Goldsmith (Goldsmith, 1994) définit les traits émotionnels comme des caractéristiques individuelles liées à la manière dont les émotions basiques sont vécues et exprimées. ”Un trait émotionnel est alors défini comme une représentation moyenne de l'ensemble des états émotionnels d'une personne à travers une variété de situations représentatives de celles que l'on rencontre dans la vie quotidienne” (Mehrabian, 1996).

Dans (Zenasni et al., 2003) Zenasni et al. présentent quelque exemple de trait émotionnel :

- Intensité affective
- L'anxiété
- L'ambivalence émotionnelle
- La créativité émotionnelle
- l'expressivité émotionnelle
- L'alexithymie
- la clarté émotionnelle
- l'attention portée à ses émotions
- Le contrôle émotionnel

Une étude de (Zenasni et al., 2003) a examiné la relation de ces traits émotionnels avec des performances cognitives. Leur étude suggère qu'il existe des liens directs entre certaines de ces dimensions émotionnelles et les performances face à une tâche cognitive de mémorisation.

Cette étude cite également les différents questionnaires qui ont permis la récolte de ces traits émotionnels auprès des utilisateurs. Ces questionnaires sont re-

TABLE 8.1: Les questionnaires de mesure des traits émotionnels

Questionnaire	Auteurs	Mesures
trait meta-mood scale (TMMS, 30 items)	Salovey, Mayer, Goldman, Turvey, et Palfai (Salovey et al., 1995)	permet d'évaluer la disposition du participant à clairement identifier, distinguer et décrire les émotions spécifiques (clarté), à prêter son attention à ses différents états émotionnels (attention) et à les contrôler (contrôle).
Emotional Expressivity Scale (EES, 17 items)	Kring, Smith et Neale (Kring, Smith, & Neale, 1994)	mesure la disposition d'un individu à exprimer ses émotions (expressivité).
Affect Intensity Measure (AIM, 40 items)	Larsen et Diener (Larsen & Diener, 1987)	évalue la disposition du sujet à ressentir faiblement ou fortement ses émotions (Intensité affective).
Emotional Creativity Inventory (ECI, 30items)	JR Averill (Averill, 1999)	mesure les prédispositions d'une personne à vivre des états émotionnels différents de ceux des autres
Ambivalence Scale (AS, 28 items)	King et Emmons (King & Emmons, 1990)	évalue la tendance des individus à être ambivalents au niveau émotionnel (Ambivalence); un individu est ambivalent lorsqu'il exprime des émotions différentes de celles qu'il ressent.

groupés dans le tableau 8.1.

8.3.2.2 Synthèse

il a été supposé que les individus diffèrent dans leur prédisposition à vivre certaines émotions (Revelle & Scherer, 2009). Nous avons constaté à travers la littérature psychologique sur l'émotion que *l'identité émotionnelle* s'articule autour de multiples différences inter-individuelles qui impacte l'expérience des émotions.

Les principaux facteurs qui ont été étudiés dans la littérature sont :

- Personnalité (McAdams & Pals, 2006 ; Mehrabian, 1996).
- Trait émotionnel (Zenasni et al., 2003).

Mis à part les concepts que nous avons explorés au cours de ces travaux, d'autres aspects responsables des différences inter-individuelles dans l'expression de l'émotion ont été étudiés par exemple :

- Ethnie (Eid & Diener, 2001).
- Profil d'appraisal qui dépend d'une mémoire émotionnelle (Scherer, 2009).

Scherer (Scherer & Ekman, 1984) soutient qu'il y'a autant d'émotions différentes

que de profil d'appréciation différent.

8.4 Approches, théories et modèles des émotions en psychologie

Le fonctionnement interne du processus émotionnel est complexe. Depuis plusieurs décennies plusieurs théories ont tenté de répondre aux questions du fonctionnement de l'émotion. Ces théories ont étudié l'émotion dans ces différentes phases de l'apparition du stimulus jusqu'au ressenti de l'émotion en proposant des schémas cognitifs ou/et physiologiques de l'activité émotionnelle. il existe trois principales approches de l'émotion.

8.4.1 Approche catégorielle

Cette première approche est une des plus anciennes. Elle s'inspire des théories évolutionnistes de Charles Darwin. L'émotion y possède une fonction adaptative, c'est-à-dire, qui sert à définir le comportement approprié en fonction d'un stimulus. Cette théorie stipule que les émotions sont innées. Il existerait donc des schémas neuronaux spécifiques pour chaque émotion, qui seront activés lors de l'apparition d'un stimulus. L'approche catégorielle soutient qu'il existe un ensemble d'émotion de base. Par exemple, Ekman et Friesen en dénombrent six : la joie, la surprise, la peur, la colère, la tristesse et le dégoût (Ekman & Friesen, 1977). Cette liste diffère selon les chercheurs Frijida et Izard incluent par exemple la honte et la culpabilité (Frijida, 1987 ; Izard, 2013). Ces émotions de base sont censées être déclenchées sur les conditions spécifiques par des programmes internes universels. Des processus neuromoteurs sont déclenchés automatiquement et sont relativement résistants aux changements.

8.4.2 Approche dimensionnelle

Contrairement à l'approche précédente, l'approche dimensionnelle stipule que les émotions ne sont pas des états discrets et indépendants les uns des autres. Les émotions constituent un espace continu à plusieurs dimensions. Les émotions peuvent être différenciées sur la base de paramètre dimensionnel telles que l'excitation, la valance ou la dominance. Plutchik (?, ?) propose un espace circumplexe en 3 dimensions pour représenter les émotions. La Figure 8.1 illustre une représentation graphique de cette approche à 3 dimensions.

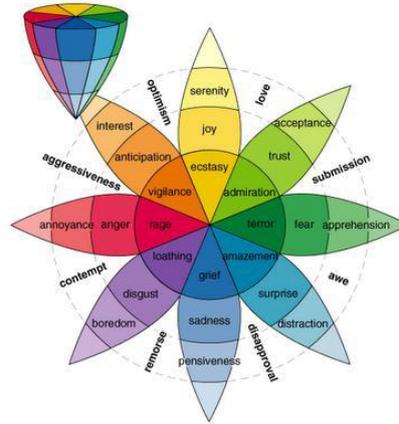


FIGURE 8.1: Le Circomplexe de Plutchik

8.4.3 Approche cognitive

Déjà dans l'allégorie du chariot de Platon modélisant un conducteur de char essayant de guider deux chevaux représentant l'intellect (cognition) et l'émotion pour se déplacer ensemble, l'idée qu'émotion et cognition sont fortement corrélées a été largement explorée. La cognition à travers l'évaluation «l'appraisal» est essentielle pour l'émergence de l'émotion (Lazarus, 1982) et l'émotion est essentielle pour la cognition et la réflexion (Damasio, 1994).

Les théories de l'évaluation expliquent le déclenchement et la différenciation des émotions sur la base de la relation entre les individus et leur environnement. Les émotions découlent de l'évaluation des situations, des objets et des agents existants dans l'environnement. L'émotion influe par la suite directement ou indirectement les objectifs, les buts, et les croyances de l'individu (S. Marsella et al., 2010) (roseman, 1990). Cette évaluation de la relation de l'individu par rapport à son environnement est effectuée en utilisant une série de "dimensions d'évaluation. Les différentes théories de l'évaluation diffèrent les une des autres par le nombre, la répartition et la définition de ces variables d'évaluation et la manière avec laquelle ces variables se combinent entre elles pour générer une émotion labellisée ou pas. Mais elles s'accordent toutes à rejeter l'hégémonie d'une source unique des émotions. Les notions de connaissances et de réflexions prennent une part importante dans cette vision des émotions en opposition aux approches décrites précédemment qui prônent que l'émotion possède un déclenchement automatique, universelle et innée. Selon Arnold (Arnold, 1950), l'évaluation cognitive est systématiquement à l'origine de la séquence émotionnelle et génère à la fois les actions appropriées et l'expérience émotionnelle en elle-même.

Il est important de noter que le processus d'évaluation n'exige pas nécessairement un calcul cognitif complexe, mais se produit souvent d'une manière automa-

tique, inconsciente, et sans effort.

Plusieurs modèles ont émergé de la théorie cognitiviste des émotions. Deux des plus emblématiques sont :

8.4.3.1 Le modèle OCC

Ce modèle présente une taxonomie des émotions selon un arbre de conditions déclenchantes. Ortony (Ortony, 1990) considère les émotions comme la combinaison d'une réaction de valences face à des objets (like vs dislikes), les préférences de l'agent (plaisir), et les conséquences de l'événement (approbation ou désapprobation) (Figure 8.2). Leur taxonomie englobe 22 émotions.

Dans le cadre de l'approche cognitive d'évaluation, le modèle OCC proposé par Ortony et al (Ortony, 1990) est le modèle le plus implémenté dans le cadre d'une architecture computationnelle des émotions (Hudlicka, 2011 ; S. Marsella et al., 2010).

8.4.3.2 Le modèle CPM Component Process Model

L'autre représentant le plus emblématique de l'approche cognitive des émotions est le modèle d'évaluation séquentielle de Scherer (Scherer, 2001). Cette approche établit quatre phases successives dans lesquelles une série d'évaluation de contrôles (SECs) décrit la relation environnement/agent. Comme le montre le diagramme 8.3, le CPM suggère que l'événement et ses conséquences sont évalués avec un ensemble de critères sur plusieurs niveaux de traitement (les composantes de l'évaluation). Le résultat de l'évaluation résulte en général d'un effet motivationnel en plus d'une réaction physiologique qui influe elle-même sur l'évaluation cognitive de l'événement. (Scherer, 2009). Le modèle CPM suggère qu'il y a quatre étapes principales d'évaluation pour réagir de manière adaptative à un événement. Pour atteindre ces objectifs, l'organisme évalue l'événement et ses conséquences en un nombre de critères définis par Scherer comme les (SEC Stimulus Evaluations Check).

Les quatre phases séquentielles de modèle CPM sont :

1. Pertinence (relevance) : Quelle est la pertinence de cet événement pour moi ? Est qu'il m'affecte directement ou affecte-t-il mon groupe social de référence ? Les SECs impliqués dans cette phase sont la nouveauté (soudaineté, la familiarité et la prévisibilité), le plaisir intrinsèque, et la pertinence par rapport au but.
2. Implications : Quelles sont les implications et les conséquences de cet événement ? Comment influent-ils mon bien-être et mes objectifs immédiats ou mes objectifs à long terme ? Les SECs impliqués dans cette phase sont la causalité motif/agent, probabilité des résultats, écart par rapport aux attentes et l'urgence.

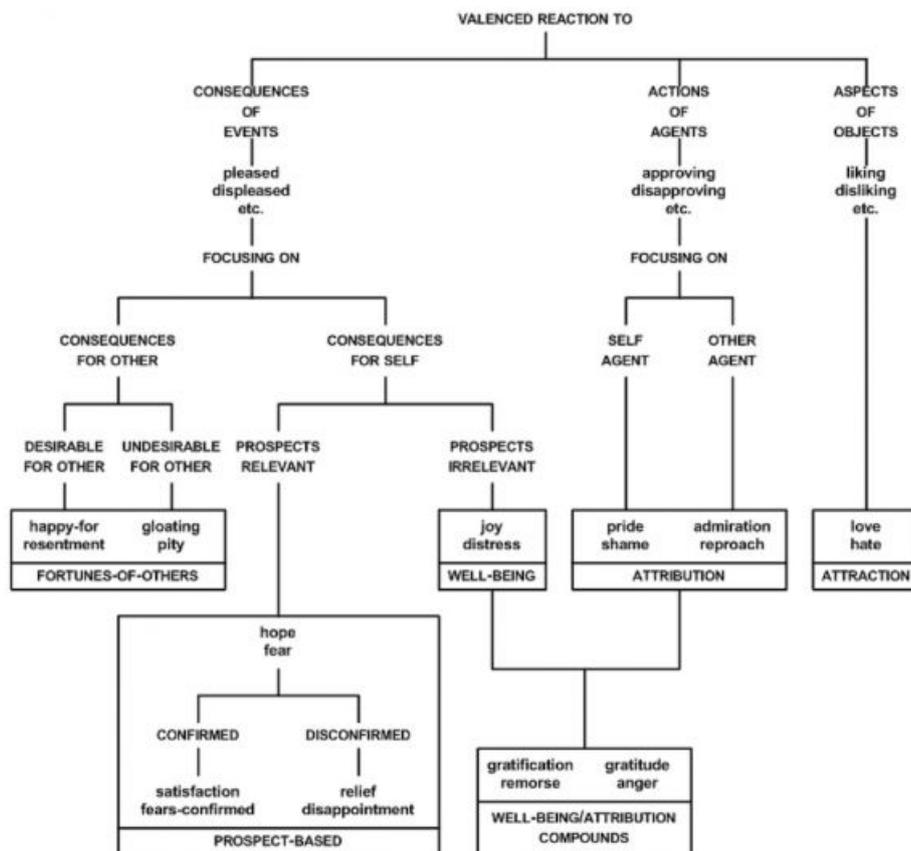


FIGURE 8.2: Modèle OCC (Ortony et al., 1990)

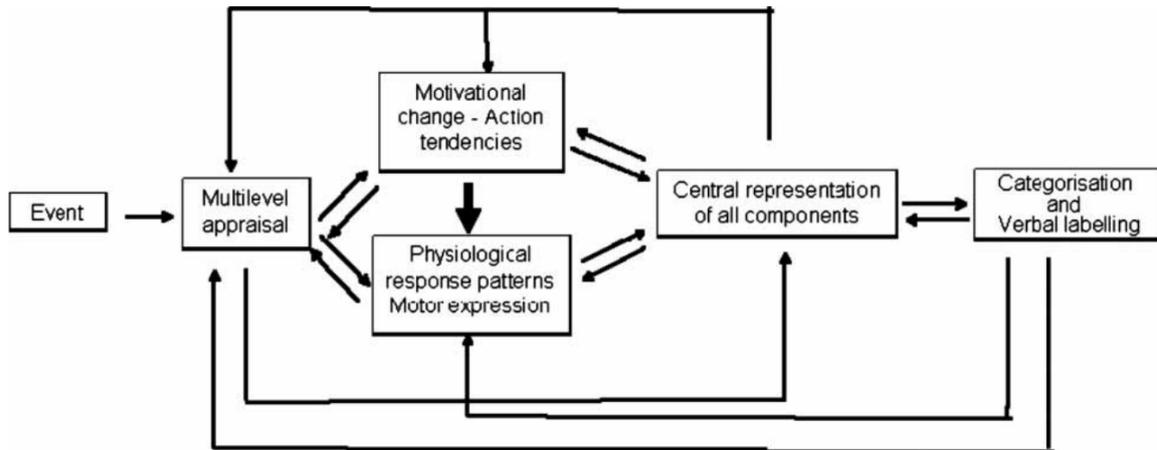


FIGURE 8.3: architecture dynamique du CPM (Scherer, 2005)

3. Stratégie d'ajustement (coping) : Comment puis-je faire face ou m'adapter à ces conséquences ? Les SECs impliqués dans cette phase sont le contrôle, le pouvoir et l'ajustement
4. Norme : Quelle est la signification de cet événement vis-à-vis de mes concepts, mes normes sociales et mes valeurs ? Les SECs impliqués dans cette phase sont la compatibilité avec mes normes internes et externes.

Dans ce modèle, pour des raisons de dépendances logiques, ces quatre phases sont traitées en séquence et suivent un ordre fixe (l'implication d'un événement ne va pas être évaluée si l'événement en lui-même n'est pas pertinent).

Comme l'illustre la figure 8.3 le modèle CPM inclut l'aspect dynamique et récursif qui caractérise une émotion. Les différentes sous étapes tiennent compte des besoins, des buts et des valeurs d'un individu en particulier.

8.4.4 Synthèse

Le processus émotionnel humain a été largement étudié dans plusieurs disciplines à travers diverses perceptives et à différent niveau d'abstraction (Frijda, 1987 ; Scherer, 2001 ; Panksepp, 1998). Ces initiatives multidisciplinaires ont abouti à la formulation d'un grand volume de théories et de modèles qui tentent d'expliquer les diverses facettes de l'émotion. Comme mentionné ci-dessus, les théories d'émotion soulignent l'importance du traitement affectif pour le développement d'un comportement rationnel chez l'homme et suggèrent la nécessité d'inclure ce phénomène dans des interactions homme-machine. Notamment ces théories suggèrent l'importance de munir des agents virtuels de comportements affectifs pour être en mesure de montrer un comportement crédible et cohérent.

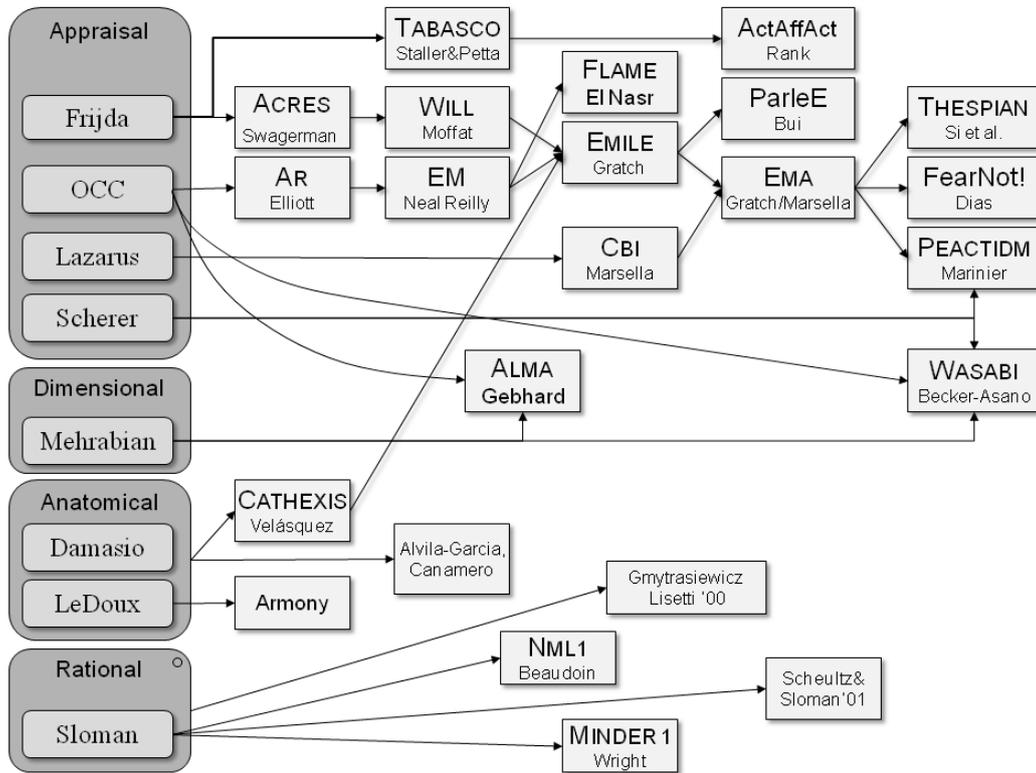


FIGURE 8.4: Modèles computationnels de l'émotion (Marsella, 2010)

8.5 Modèle computationnel des émotions

Ces résultats théoriques issus du domaine de la psychologie ont posé une série de défis dans la communauté de l'intelligence artificielle. Le développement d'un modèle computationnel de l'émotion implique une variété de choix de conception à faire qui doit être abordé afin de parvenir à un modèle qui imite certaines facettes du processus de l'émotion humaine. La figure 8.4 liste les modèles les plus importants avec leur inspiration théorique issue de la psychologie. (S. Marsella et al., 2010) En général le modèle computationnel des émotions inclut trois étapes (S. Marsella, 2010) :

1. Des mécanismes pour l'évaluation d'un stimulus.
2. La sélection de l'émotion
3. La génération d'une réponse émotionnelle

Nous avons vu dans la section précédente que les théories de l'émotion sur lesquelles se basent ces différents modèles ont des conceptions substantiellement très différentes. Les théories de l'appraisal semblent dominer les modèles computationnels de génération d'émotion. Cela est dû à sa considération de l'émotion comme une notion qu'on peut déduire suite à raisonnement logique.

Model	Theoretical foundations	Appraisal dimensions	Emotions
EMA [38]	Appraisal theory by Smith and Lazarus [44]	Relevance, perspective, desirability, likelihood, expectedness, causal attribution, controllability, and changeability	Surprise, hope, joy, fear, sadness, anger, and guilt
Flame [45]	Appraisal theory by Ortony et al. [12] and Roseman et al. [42]. Inhibition model by Bolles and Fanselow [46]	Desirability, expectation, causal attribution, and standards (or social norms)	Joy, sad, disappointment, relief, hope, fear, pride, shame, reproach, and admiration. <i>Complex emotions</i> : Anger (sad + reproach), gratitude (joy + admiration), gratification (joy + pride), and remorse (sad + shame)
Mamid [47]	Appraisal theories such as those by Lazarus [43] and Smith and Kirby [48]. Personality models such as the five factor model [49]	<i>Universal</i> : novelty, unexpectedness, intensity, threat level, desirability. <i>Individual and context-dependent</i> : individual history, experience, and expectation and goal congruence	Anxiety/fear, anger/aggression, negative affect (sadness, distress), and positive affect (joy, happiness)
Alma [39]	Appraisal model by Ortony et al. [12], five factor model of personality [49], and PAD temperament space by Mehrabian [50]	Desirability of events, praiseworthiness of actions, appealingness of objects, liking of objects, likelihood of an event occurs, and realization that an event has occurred	Admiration, anger, disliking, disappointment, distress, fear, fears confirmed, gloating, gratification, gratitude, happy for, hate, hope, joy, liking, love, pity pride, relief, remorse, reproach, resentment, satisfaction, shame
Cathexis [37]	Diverse psychological [42, 51] and neuropsychological [1] theories	The model does not report used appraisal dimensions but states that they are based on Roseman et al. [42]	<i>Primary emotions</i> : anger, fear, sadness/distress, enjoyment/happiness, disgust, and surprise. This model generates <i>secondary emotions</i> but does not provide an explicit model for the labeling of them
PEACTIDM [52]	Appraisal theory by Scherer [13] and physiological concepts of feelings by Damasio [1]	Suddenness, predictability, goal relevance, intrinsic pleasantness, outcome probability, causal agent, causal motive, discrepancy from expectation, conduciveness, control, and power	This model implements the model by Scherer [13] for the mapping of appraisal dimension values to specific model emotions
WASABI [36]	Appraisal theory by Scherer [13], PAD space by Mehrabian [50], and physiological concepts by Damasio [1]	Intrinsic pleasantness, expectation deviation, goal conduciveness, dominance, causal (mis) attribution, and coping	<i>Primary emotions</i> : angry, annoyed, bored, concentrated, depressed, fearful, happy, sad, surprised. <i>Secondary emotions</i> : hope, fears confirmed, relief

FIGURE 8.5: Les principaux modèles computationnels des émotions se basant sur la théorie de l'appraisal (Rodriguez, 2014)

Les théories de l'appraisal gagnent de plus en plus les architectures simulant une intelligence artificielle chez les agents virtuels puisqu'elle stipule que l'émotion découle d'un processus cognitif. Nous allons par conséquent nous intéresser au modèle d'évaluation, car ce dernier nous semble le plus pertinent pour une modélisation d'une fidélité cognitive (nous justifions plus amplement ce choix dans le chapitre suivant). Nous allons citer quelque exemple dans cette section.

Beaucoup de modèles s'inspirent de l'approche cognitive pour générer un comportement émotionnel. La table ci-dessous Figure-1 8.5 décrit les différents modèles se basant sur les théories de l'appraisal, ainsi que les processus d'évaluation à laquelle elles font appel ainsi que l'ensemble des émotions générées. Dans cette section nous allons discuter les modèles qui nous ont semblé les plus emblématiques et les plus pertinents pour nos travaux.

EMotion and Adaptation (EMA) (S. C. Marsella & Gratch, 2009) : Dans ce modèle, la relation entre l'environnement et l'individu est représentée par une structure appelée l'interprétation causale de l'agent. Elle englobe, entre autres,

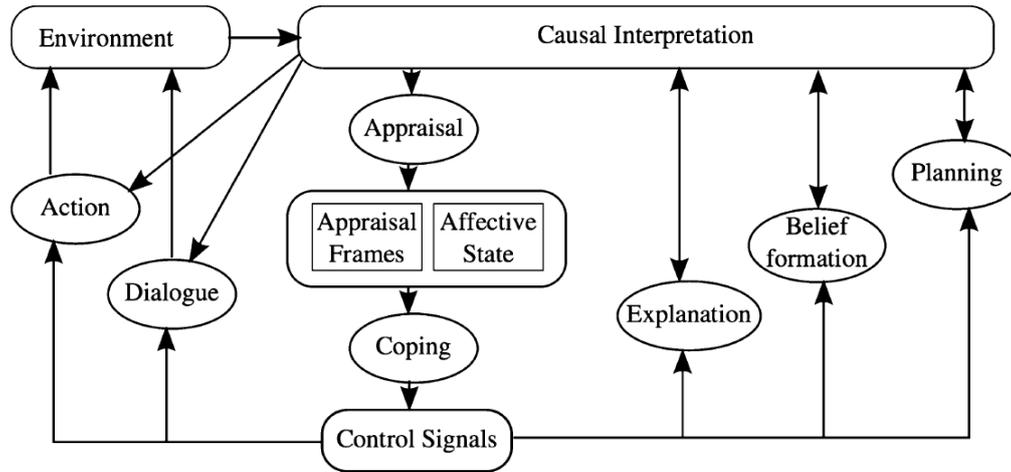


FIGURE 8.6: L'architecture EMA (Marsella, 2009)

des représentations explicites de l'état actuel de l'environnement ainsi que les croyances de l'agent, ses désirs et intentions.

Cette structure est construite et mise à jour par des processus cognitifs de perception. Et évaluée par des processus d'évaluation qui identifient et caractérisent en continu tous les événements significatifs en termes d'une série de variables d'évaluation. Ces variables d'évaluation sont en outre organisées en structures de données appelées cadres d'évaluation. Selon la configuration particulière des valeurs de ces variables d'évaluation, chaque trame est marquée avec un type spécifique d'émotion et son intensité. Enfin, afin de générer l'état affectif de l'agent, tous les cadres d'évaluation sont ajustés selon l'humeur active de l'agent. Parmi les différentes émotions résultantes, celle qui a la plus grande intensité est choisie comme étant l'état émotionnel de l'agent. Une abstraction de l'architecture de EMA est présentée dans la figure-8.6 ci-dessous.

PEACTIDM Marinier et al (Marinier, Laird, & Lewis, 2009) proposent un modèle computationnel des émotions qui combine le modèle CPM de Scherer (Scherer, 2001) et la théorie PEACTIDM de Newell et al. (Newell, 1994). Ce modèle spécifie un cycle composé de huit fonctions cognitives visant à produire un comportement affectif. Durant ce cycle, une série de processus d'évaluation sont distribués dans le but d'extraire l'information émotionnelle.

8.5.1 Modèle de génération d'expression faciale basée sur la théorie du CPM

Quelques travaux se sont basés sur la théorie de Scherer (Scherer, 2001) pour proposer un système d'animation faciale sur différente plateforme. Par Exemple,

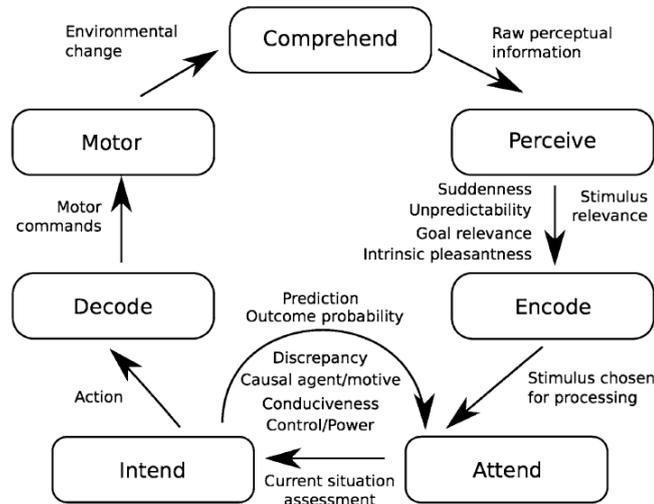


FIGURE 8.7: Modèle cognitif pour un agent virtuel intégrant une fonction affective basée sur CPM (Marinier et al., 2005)

Paleari et al. (Paleari, Grizard, & Lisetti, 2007) ont mis en oeuvre un système qui affiche des expressions temporaires lors de différentes phases de l'évaluation cognitive du modèle CPM (Figure 8.8).

Malatesta et al. (Malatesta, Raouzaïou, Karpouzis, & Kollias, 2009) ont également abordé la dynamique des expressions faciales en se basant sur la théorie cognitive CPM. Ils ont proposé deux modes d'animation. Une première approche où ils effectuent une accumulation successive des expressions des différentes évaluations cognitives. Dans la deuxième approche chaque nouvelle expression substitue l'expression précédente. Leurs résultats suggèrent que la méthode cumulative est mieux acceptée par les observateurs. Cependant cette étude met en relief certaines limites de la théorie du CPM. En effet, certaines activations des Aus définis dans le modèle de Scherer (Scherer, 2001) sont en contradiction lors de l'évaluation séquentielle d'un événement.

8.6 Synthèse

Nous avons étudié la notion d'identité et nous avons découvert que l'identité est un concept des plus hétéroclite qui englobe un nombre important de notions que ce soit les self-concepts, la personnalité, les croyances et les buts, mais aussi la construction sociale.

L'identité possède diverses dimensions qui s'influencent mutuellement. En effet, l'identité peut être personnelle, ethnique, sociale, mais aussi émotionnelle. Nous nous sommes attardés sur l'identité émotionnelle à travers l'étude des différences inter-individuelle dans l'émergence et l'expression des émotions.

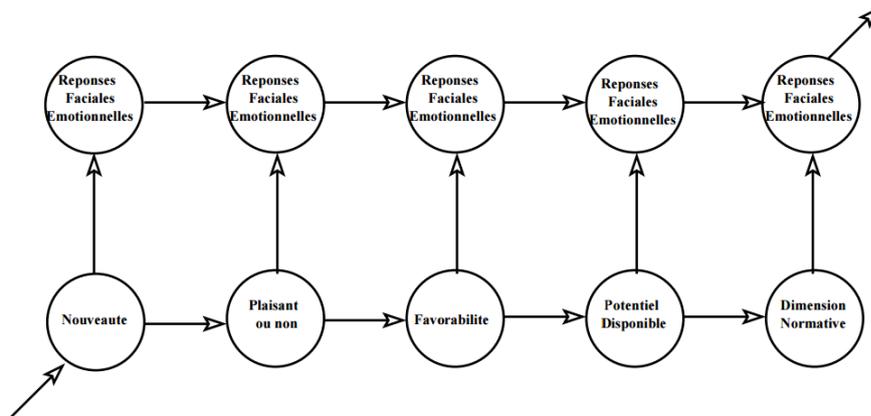


FIGURE 8.8: Représentation computationnelle du processus émotionnelle faciale (Grizard, 2006)

Par la suite nous nous sommes intéressés à regarder les approches théoriques de l'émotion ainsi que quelques modèles computationnels. Cela nous permit de constater que dans le monde de l'intelligence artificielle c'est l'approche d'évaluation cognitive qui est la plus traduite en modèle informatique. Dans le chapitre suivant, nous allons regarder la manière avec laquelle nous pouvons traduire et adapter un modèle computationnel des émotions pour qu'il véhicule un sentiment de fidélité lors de l'expression de l'émotion.

Vers une fidélité comportementale : réflexions sur une architecture d'évaluation émotionnelle personnalisée

Sommaire

9.1	Introduction	118
9.2	Choix du cadre théorique des émotions	118
9.3	Justification du choix du modèle CPM	119
9.4	Architecture d'un modèle computationnel d'émotion personnalisé	121
9.4.1	Adaptation du modèle CPM	121
9.4.2	Caractéristiques cognitives	123
9.4.3	Profil émotionnel	123
9.4.4	Synthèse	123
9.5	Cas d'étude : Protocole d'élicitation GEMEP . . .	124
9.5.1	Annotation manuelle des scripts GEMEP	125
9.5.2	Déroulement d'un exemple	126
9.6	Conclusion et perspectives	126

9.1 Introduction

Au vu de l'état de l'art auquel nous avons procédé dans le chapitre précédent, nous avons constaté que certaines études psychologiques soutiennent que la fidélité comportementale serait aussi importante lors d'une tâche d'identification que l'information visuelle.

Il nous fallait donc étudier la manière par laquelle nous pouvons incorporer quelques éléments de la psychologie de la personne clonée afin de l'intégrer dans une architecture d'un agent autonome. Pour cela nous nous sommes intéressés à la définition d'une identité psychologique. Nous en avons déduit que l'identité est une notion complexe qui englobe, entre autres, le concept de "personnalité", de "self", de mémoire, des croyances et des buts à long terme.

La fidélité comportementale est un champ large et complexe. Elle implique la totalité de nos capacités cognitives, notre vécu et notre construction sociale. Dans le cadre de nos travaux, nous allons nous restreindre à l'étude de la fidélité de la structure affective d'une identité. Dans le but de simuler un comportement fidèle, nous allons nous focaliser sur la modélisation d'une architecture émotionnelle personnalisée.

La finalité de nos questions de recherches abordées dans ce chapitre est de savoir si nous pouvons simuler un modèle computationnel des émotions qui est capable de guider une expressivité faciale émotionnelle de telle manière qu'elle soit fidèle aux réactions émotionnelles du référent. Par exemple, serait-il possible de reconnaître une identité sur un visage virtuel expressif dépourvu de toute fidélité visuelle, par exemple, un visage virtuel générique ?

Ce chapitre a pour but de présenter les premières briques d'une adaptation d'un modèle computationnel de l'émotion intégrée à notre architecture de clone virtuel. Cette architecture assurera des réactions émotionnelles similaires à celles que pourrait avoir le référent dans le monde réel. Pour cela il fallait tout d'abord opter pour une approche de la modélisation de l'émotion et par conséquent un choix de cadre théorique.

9.2 Choix du cadre théorique des émotions

En psychologie, beaucoup de chercheurs ont proposé des mécanismes et des théories expliquant la production des émotions. L'approche cognitive d'évaluation des émotions (appraisal) semble la plus pertinente pour modéliser un comportement émotionnel fidèle. En effet, ces théories se basent sur l'hypothèse que l'émotion résulte d'un processus d'évaluation de la situation par apport à l'individu (Arnold, 1960 ; Scherer, 2001). Une des principales hypothèses sur lesquelles se basent ces théories est que des personnes différentes peuvent réagir de manières distinctes à un même stimulus, et ce en fonction de leur évaluation

de cet événement.

Cette approche, qui assure une réponse émotionnelle personnelle face à un même événement, est particulièrement pertinente pour caractériser un comportement fidèle aux réactions émotionnelles du référent.

9.3 Justification du choix du modèle CPM

Une fois l'approche de l'évaluation "appraisal" retenue, il nous fallait choisir parmi les nombreux modèles adoptant cette approche. Le modèle CPM (Scherer, 2001) a particulièrement retenu notre attention due à sa capacité potentielle à gérer les différences individuelles dans la réponse émotionnelle. Nous présentons ci-dessous les arguments qui ont appuyé notre choix :

Preuves empiriques de la considération par CPM des différences inter-individuelles. Nous avons constaté dans le chapitre précédent que le modèle CPM prend en compte plusieurs différences inter-individuelles et que ces différences prennent forme tout au long de la dynamique émotionnelle (figure-9.1).

"CPM a l'avantage d'être particulièrement utile pour la conceptualisation et la prédiction des différences individuelles dans les réactions émotionnelles dans le cas où l'événement déclenchant est objectivement le même" (Scherer, 2009).

Dans son modèle Scherer insiste sur le fait que l'émotion est un processus émergent dynamique basé sur une évaluation individuelle. Il soutient que le modèle CPM est capable de prédire les différences individuelles dans la réponse émotionnelle ce qui est particulièrement pertinent pour nos travaux sur la fidélité comportementale (Scherer, 2001).

Scherer (Scherer, 2009) soutient que le modèle des composantes d'évaluation est le modèle qui prend le plus en compte les différences inter-individuelles. Par exemple, Scherer relate une étude où ils ont analysé les différences de réactions sur 103 personnes face à un événement identique (perte de ses bagages dans un aéroport).

Cet exemple prouve selon les auteurs l'utilité de la théorie des composantes d'évaluation pour modéliser et prédire les différences individuelles. Les auteurs soutiennent que cette étude prouve que les émotions sont provoquées entièrement sur la base d'évaluations subjectives d'un individu de l'événement qu'il perçoit. Celui-ci est en effet évalué sur la base des SECs : des objectifs, des valeurs de l'individu ainsi que sur son potentiel d'adaptation (Scherer, 2009).

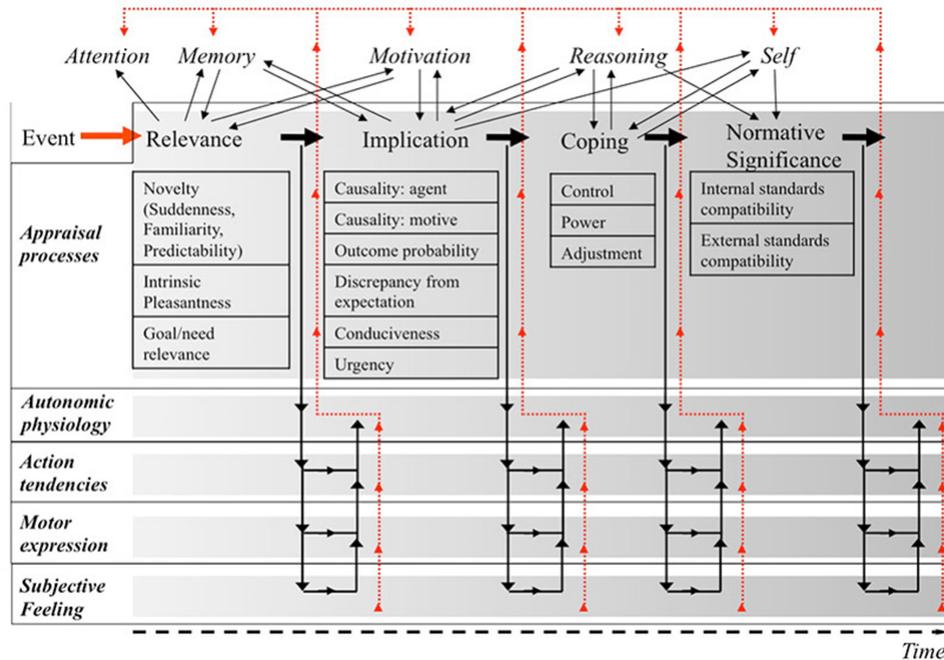


FIGURE 9.1: Architecture séquentielle du modèle CPM (Sander et al., 2005)

Formalisation des réactions musculaires. Le modèle CPM a également pour intérêt de proposer une formalisation des réponses motrices à chaque sous-étape de l'évaluation de l'événement déclencheur de l'émotion. CPM fait des prédictions précises en ce qui concerne les effets qui découlent de l'évaluation de certains SECs sur le système nerveux automatique et somatique indiquant quels changements physiologiques et quelle expression motrice caractéristiques sont attendus. En particulier, CPM prédit des mouvements faciaux qui accompagnent l'évaluation de certains SECs nous permettant ainsi de prédire une animation faciale individuelle basée sur une évaluation personnalisée.

Espace infini d'émotions potentielles. Contrairement aux théories de base, le CPM ne suppose pas l'existence d'un ensemble limité d'émotions discrètes. Cela a l'avantage de ne pas produire des émotions prototypiques. Les travaux d'Ekman (Ekman & Friesen, 1977) suggèrent l'existence d'un sous-ensemble d'expressions universellement reconnues. Nous pensons que l'approche catégorielle n'est pas la plus adaptée pour représenter des différences inter-individuelles. En effet, bien que cette théorie prône aussi la multiplicité des expressions pour une même émotion (Ekman & Friesen, 2003), elle se focalise sur des expressions labellisées et sur les caractéristiques expressives communes de ces dernières. Le modèle CPM considère la possibilité d'un nombre infini de différents types d'épisodes émotionnels (Scherer, 2009) puisque les différentes étapes d'évaluation

ne résultent pas forcément en une émotion labellisée, mais plutôt en une suite de réactions physiologiques.

9.4 Architecture d’un modèle computationnel d’émotion personnalisé

9.4.1 Adaptation du modèle CPM

Le modèle CPM est un modèle issu de la psychologie qui vise à simuler le processus émotionnel. Scherer définit par conséquent des notions complexes qui sont sujettes à interprétation. Par exemple le concept de pouvoir, de contrôle ou norme sociale englobe plusieurs notions et peut être interprété de plusieurs manières. Une simplification de ces concepts et leur traduction dans un contexte bien précis sont nécessaires pour une simulation informatique.

Par conséquent, il nous a semblé important de ne considérer qu’un sous-ensemble des critères d’évaluation de Scherer. Pour guider notre choix de sélection, nous avons décidé de nous limiter au SECs auxquels Scherer a associé une réaction faciale (Scherer & Ellgring, 2007; Scherer, 2009). En effet, toutes les composantes d’évaluation du CPM ne génèrent pas forcément une réponse musculaire. Ils peuvent donner lieu à d’autres réactions physiologiques.

Table 1. Prediction of sequential appearance of action units (AUs) in response to appraisal check results for five major emotions

Order	Appraisal domain	Check	Subcheck	AUs	Fear	Anger	Joy	Sadness	Disgust
1	Relevance	Novelty	Sudden	1 + 2, 5	**	*	*	—	*
2			Unfamiliar, unpredictable	4 + 7	**	*	*	—	*
3		Intrinsic pleasantness	Pleasant	12 + 25 + 38	—	—	**	—	—
3			Unpleasant	9, 10, 15, 39	*	—	—	—	*
4	Implications	Expectation	Discrepant	4 + 7	*	**	—	*	*
5		Goal attainment	Conducive	6 + 12 + 25	—	—	**	—	—
5	Coping potential	Power/control	Obstructive	17 + 23, 17 + 24	*	**	—	**	—
6			Low	20, 26, 27	**	—	—	**	—
6			High	23 + 25, 17 + 23, 6 + 17 + 24	—	**	*	—	*
7	Norm compatibility	External standards	Violated	10	—	*	—	—	—
7		Internal standards	Violated	14	—	—	—	—	—

FIGURE 9.2: Prédiction d’apparition des actions unitaires (AUs) suite à l’évaluation de SEC de CPM pour 5 émotions de base (Scherer, 2007)

Nous nous sommes basés sur les travaux de (Scherer & Ellgring, 2007) (figure 9.2) pour sélectionner un ensemble de ”subcheck” sur lesquelles nous allons

nous baser pour simuler l'évaluation de l'événement déclencheur de l'émotion. Notre système fonctionne donc suivant 7 SEC : nouveauté, plaisir intrinsèque, probabilité des conséquences, rapport au but, contrôle, rapport aux normes externes et le rapport aux normes internes. À ces 7 SECs sont associés 13 subchecks comme l'illustre la figure 9.3

Pertinence	Nouveauté	Soudain Non familier Non prédictible
	Plaisir intrinsèque	Plaisant Non plaisant
Conséquence	Probabilité des conséquences	En désaccord avec mes buts
	Rapport au but	Propice Obstructif
Capacité de maîtrise	Contrôle	Pas de contrôle Contrôle
Rapport aux normes	Normes externes	Enfreintes
	Normes internes	Enfreintes

FIGURE 9.3: Ensemble des SEC retenus pour l'évaluation émotionnelle

Une fois les niveaux d'évaluation sélectionnés, se pose alors la question de comment leur attribuer une valeur. L'idée principale de notre approche est de calculer la valeur des SECs sélectionnés suivant les différences individuelles qui caractérisent une identité. Deux principaux modules sont mis en œuvre dans ce but.

- Un ensemble de caractéristiques cognitives particulièrement pertinentes pour la fidélité
- Un profil émotionnel

9.4.2 Caractéristiques cognitives

Le modèle CPM suggère que l'événement et ses conséquences sont évalués avec un ensemble de critères sur plusieurs niveaux de traitement (la composante d'évaluation) dénommé "SEC".

Dans un premier temps il faudrait mettre en place un ensemble de caractéristiques cognitives définies par Scherer qui caractérise un profil cognitif (Figure 9.1). En effet, nous devons mettre en place un système logique et une base de connaissance qui doit prendre en compte les différentes caractéristiques, que nous avons dégagés dans le chapitre précédent, qui constitue une identité.

Ce système doit par exemple prendre en compte la mémoire du référent, son attention, ses buts à long terme, sa motivation et des self-concepts.

9.4.3 Profil émotionnel

Un profil émotionnel est constitué d'un "profil de personnalité" tel que les traits du Big Five ou/et des "traits émotionnels" relatifs aux référents. Ce profil sera extrait de questionnaires psychométriques remplis par le référent.

Notre idée consiste à injecter les résultats de ces tests pour d'une part influencer l'évaluation cognitive des différents SECs, mais aussi influencer l'intensité des AUs résultantes. (Revelle & Scherer, 2009)

D'habitude les tests psychométriques (tel que les questionnaires de personnalité, de profil émotionnel ou d'intelligence émotionnelle) sont utilisés du point de vue de l'utilisateur pour évaluer l'impact de la perception d'un stimulus sur une certaine catégorie de sujets. À notre connaissance, il n'y a pas eu de travaux qui ont essayé d'injecter les résultats de ces questionnaires dans un modèle computationnel pour générer un comportement personnalisé.

9.4.4 Synthèse

Comme illustré dans la figure-9.4, les valeurs de SECs sont donc évaluées de manière séquentielle. Chaque "Subcheck" peut prendre la valeur "open", "low" ou "high" comme cela a été définie dans (Scherer, 2001). L'évaluation de ces SEC est donc résultante à la fois des caractéristiques cognitives relatives à un individu et de son profil émotionnel. De surcroît, un profil expressif, composé d'un ensemble de mouvements idiosyncrasiques traduisant une signature dynamique du référent (Lander et al., 1999 ; O'Toole et al., 2002), agit directement sur l'expression faciale générée. Finalement des expressions unitaires sont activées. Ces expressions sont cumulatives tout au long du processus d'évaluation.

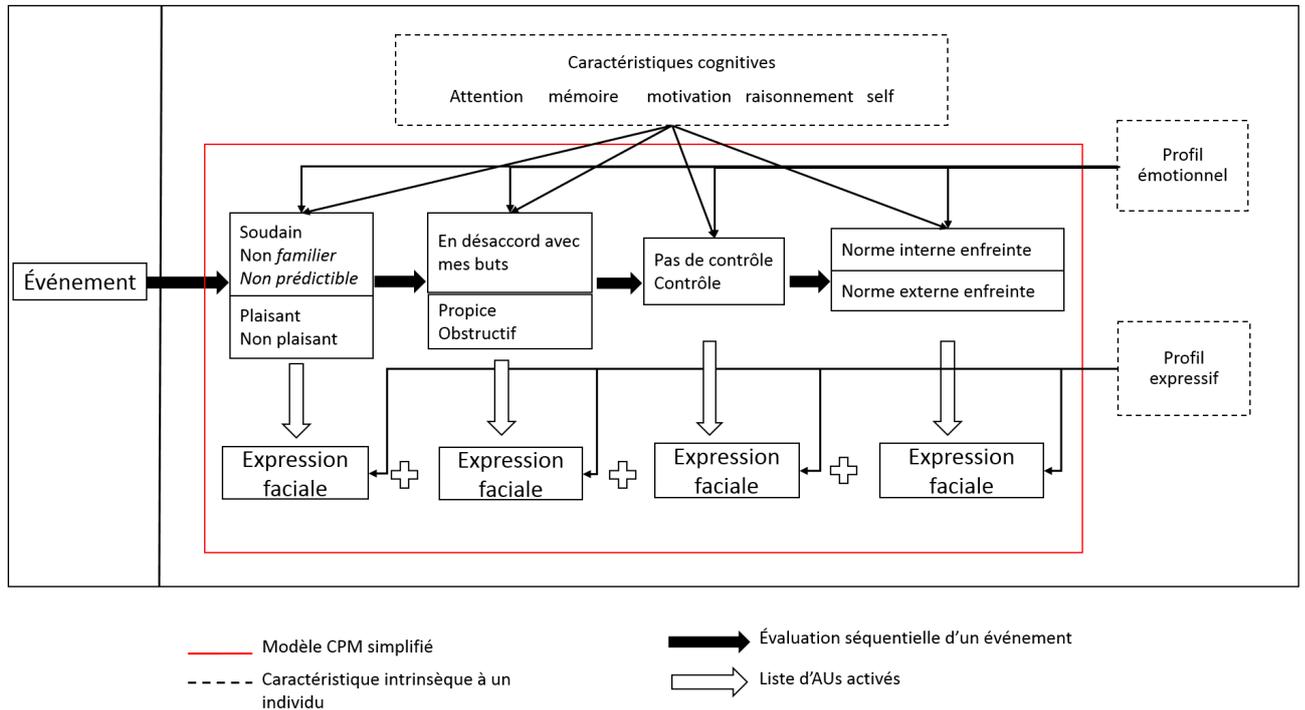


FIGURE 9.4: Modèle computationnel pour une expressivité émotionnelle fidèle

9.5 Cas d'étude : Protocole d'élicitation GE-MEP

Dans un premier abord, il fallait donner un contexte d'application afin d'adapter et d'évaluer une architecture de génération d'expression faciale individualisée. Nous avons opté pour une application simple où le clone réagit en temps réel à un texte qui s'affiche. Pour cela, nous avons opté pour les scripts de GE-MEP (Bänziger & Scherer, 2010).

En effet, nous avons décidé de nous baser sur le même scénario que ceux utilisés lors de l'élicitation des émotions au cours des sessions de capture de mouvement. Cela nous permettra d'avoir une base de comparaison puisque nous aurons l'évaluation expressive réelle du référent. Notre but est que le clone réagisse à ces scripts en se basant sur notre adaptation de l'architecture CPM. Nous pouvons envisager d'effectuer une évaluation perceptive de la fidélité de l'expression générée par le modèle computationnel en comparaison avec les données réelles de capture de mouvements.

9.5.1 Annotation manuelle des scripts GEMEP

Nous avons vu que le module des caractéristiques cognitives se base sur un système logique et une base de connaissance pour attribuer la bonne valeur aux critères de l'évaluation (SEC) et que ce profil englobe des notions telles que la mémoire, les buts à long terme et des self-concepts. Or, dans nos travaux actuels, nous n'avons pas l'ambition de modéliser un module cognitif aussi complexe qui serait capable d'évaluer les valeurs de SEC pour tous les événements possibles. Cela ferait appel à des notions avancées de l'intelligence artificielle qui sont éloignées de nos objectifs de recherches.

Notre objectif demeure l'étude du poids de l'information des expressions faciales émotionnelles dans l'évaluation de la fidélité d'un clone virtuel. L'évaluation cognitive de l'événement sera donc dans un premier temps simulée grâce à une annotation manuelle des scripts de GEMEP. En effet, notre approche consiste à faire annoter manuellement, suivant les 7 SEC sélectionnés de CPM, les différents événements des scripts de GEMEP. Cette annotation pourrait par exemple être effectuée par trois personnes qui sont familières au référent et par le référent lui-même pour ensuite moyenner les résultats. Un exemple d'annotation potentiel est illustré par la figure-9.5. Cette figure illustre l'évolution des valeurs des sous-check entre la première phrase du script qui est neutre et l'événement que présente la phrase 2.

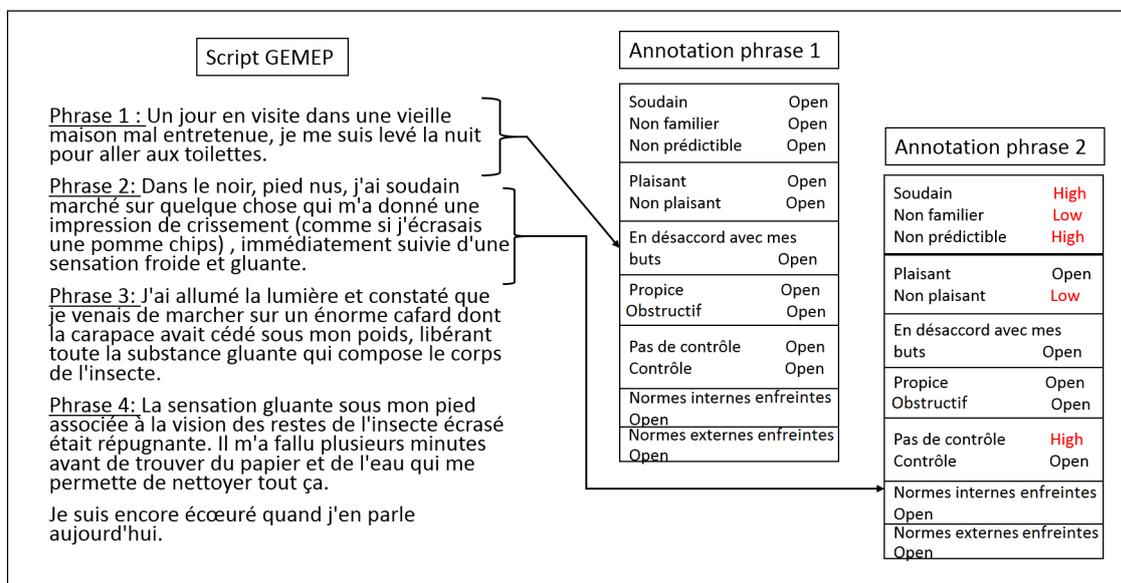


FIGURE 9.5: Annotation manuelle d'un script du corpus GEMEP pour un même individu

9.5.2 Déroulement d'un exemple

Nous prévoyons de faire passer des questionnaires psychométriques au référent que ce soit des tests de personnalité par exemple le BIG FIVE (Caprara, Barbaranelli, Borgogni, & Perugini, 1993) ou de trait émotionnel (Zenasni et al., 2003). Ces tests permettraient de définir un profil émotionnel du référent. Par la suite, grâce à la combinaison de ce profil émotionnel et à l'annotation manuelle des scripts nous pouvons générer comme définie dans la figure 9.4 une animation séquentielle relative à chaque phrase du script GEMEP. L'idée est de faire évaluer par des utilisateurs l'animation faciale engendrée en temps réel par l'activation de ces SEC sur deux clones virtuels. Cela nous permettra d'évaluer si cette couche émotionnelle impacte l'identification et le jugement de la fidélité du clone.

Nous pouvons également envisager d'animer un clone virtuel avec un modèle émotionnel d'un autre référent pour explorer l'effet de la non-congruence que cela pourrait générer.

Dans un protocole similaire à celui défini pour l'évaluation de la fidélité visuelle, nous pourrions également évaluer l'impact de la familiarité des observateurs sur le jugement de la fidélité et l'appréciation d'un clone virtuel animé par ce modèle computationnel des émotions personnalisées.

9.6 Conclusion et perspectives

Ce chapitre a pour but de présenter les prémisses d'une adaptation d'un modèle computationnel de l'émotion personnalisée. Cette architecture vise à produire des réactions émotionnelles similaires à celle que pourrait avoir le référent dans le monde réel. En effet, par le développement de cette section, notre ambition est de poser des perspectives de recherches et d'application pour la fidélité comportementale relative à un clone virtuel. Les études en psychologie ont mis en avant la nécessité de modéliser une cognition dans un clone pour renforcer l'identification. Pourtant l'approche actuelle de l'Intelligence Artificielle et de l'informatique affective tend à simuler un comportement générique ou relatif à un certain groupe.

Notre choix s'est porté sur le modèle CPM (Component process model) de (Scherer, 2001) car ce dernier, par ses composantes séquentielles, permet de prendre en compte les différences inter-individuelles et donc de personnaliser l'expression à chaque étape de l'émergence de l'émotion.

Nous pensons que notre problématique n'est pas restreinte à l'étude de clones virtuels. L'identité pourrait aussi être prise en compte dans la conception de l'humain virtuel plus anonyme afin d'améliorer le réalisme des comportements. En effet, une question récurrente en Intelligence Artificielle est "Sur quoi se

base notre jugement de l'intelligence d'un humain virtuel". Beaucoup de variables peuvent influencer l'évaluation de l'intelligence. Parmi ces variables nous pouvons citer l'autonomie, le réalisme du rendu, le réalisme de l'animation, le comportement non verbal et l'interactivité. Par nos travaux nous avons voulu suggérer qu'en plus de toutes ces variables, un autre aspect axé sur l'identité pourrait être pris en compte lors de l'évaluation du comportement de l'humain virtuel.

De plus, au vu du succès actuel des techniques de scan 3D et des dispositifs de capture de mouvement, la fidélité comportementale aura de plus en plus d'importance dans les futures applications impliquant des agents virtuels identifiables.

Cependant, plusieurs défis restent à résoudre pour mettre en œuvre cette architecture émotionnelle personnalisée. Il faudra par exemple justifier le lien entre le profil émotionnel et les différents SEC de CPM.

De plus, des études basées sur le modèle CPM pour la génération d'expressions faciales ont noté que certaines AUs rentrent en conflit lors du processus séquentiel de l'évaluation. Par composant, le résultat de l'animation séquentielle cumulative pourrait constituer une source de confusion (Malatesta et al., 2006).

Chapitre 10

Conclusions et perspectives

Sommaire

10.1 Introduction	130
10.2 Démarche méthodologique	131
10.3 Contributions	131
10.3.1 Contributions théoriques	131
10.3.2 Contributions techniques	133
10.3.3 Contribution expérimentale	133
10.4 Limites de nos travaux	134
10.5 Perspectives	135
10.5.1 Profils expressifs	135
10.5.2 Fidélité comportementale	135
10.5.3 Perspectives expérimental	136

10.1 Introduction

L'objectif principal de cette thèse était de modéliser, implémenter, et évaluer la fidélité d'un clone virtuel expressif. Un clone virtuel est la représentation en trois dimensions du visage d'un être humain existant. Cette spécificité nous a amenés à considérer de nouvelles questions de recherches sur la perception de visages virtuels puisque des variables telles que la familiarité avec le référent pourraient influencer notre perception et évaluation de sa représentation virtuelle.

Notre problématique principale est donc de comprendre comment les utilisateurs perçoivent et interprètent des visages familiers générés par ordinateurs. Cette problématique tire sa motivation de deux constatations :

- Les visages virtuels actuels sont de plus en plus développés, les avancées en informatique graphique ainsi que les dernières techniques de scan permettent des rendus photo-réalistes (von der Pahlen et al., 2014 ; O. Alexander et al., 2010). Ces rendus sont tellement avancés qu'ils arrivent à défier notre perception à distinguer le fait qu'ils soient réels ou virtuels.
- La familiarité vis-à-vis d'une identité est un phénomène primordial dans nos interactions sociales. Cet aspect a été largement exploré par les psychologues pour les visages réels sur la base de vidéo ou de photos. Leurs résultats démontrent clairement que notre comportement diffère face à des visages familiers.

En dépit de la maturité que connaît l'informatique graphique en terme de création et de rendu de visages virtuels, très peu d'études ont étudié un visage virtuel doté d'une identité. En effet, bien que les clones virtuels soient de plus en plus répandus dans diverses interfaces que se soit dans le monde de la recherche ou l'univers du divertissement, très peu d'études se sont intéressées à évaluer la manière avec laquelle on appréhende un visage virtuel représentant une identité. Très peu d'études évaluent la qualité des résultats obtenus. Les clones virtuels ont l'avantage de proposer un support de comparaison. Pourtant, cette opportunité n'est pas bien exploitée dans le domaine de recherche sur la perception d'humains virtuels. Il est donc devenu de plus en plus important pour les chercheurs en informatique graphique de comprendre et appréhender ces travaux.

Afin de tenter d'amener des réponses à ces questions, nous avons pris comme optique d'étudier les différentes composantes de la fidélité d'un visage virtuel. En effet, nous considérons que l'évaluation d'un clone virtuel revient à faire la comparaison entre le visage généré par ordinateur et le référent. La fidélité nous a paru donc comme la métrique naturelle pour l'évaluation d'un clone virtuel.

10.2 Démarche méthodologique

Notre démarche avait donc pour fil conducteur l'étude de la fidélité. Nous avons par conséquent opté pour une démarche incrémentale qui explore les différentes dimensions de la fidélité.

Nous n'avons pas rencontré d'étude qui abordait la fidélité de visage virtuel expressif. Par conséquent, dans une première étape nous avons étudié des disciplines annexes aux recherches sur les humains virtuels qui abordaient la notion de fidélité. Nous en avons conclu que la fidélité ne possède pas de définition unique et unanime. Nous avons également constaté qu'il manquait à la littérature un cadre formel qui énumère ses différentes composantes et ses différentes variables. Nous nous sommes donc intéressés à définir la fidélité d'un visage virtuel ce qui a abouti à la proposition d'un cadre conceptuel pour l'étude de la fidélité.

Par la suite nous avons entamé l'étude des différentes dimensions de la fidélité en commençant par la fidélité visuelle. Nous avons en effet exploré les différentes techniques qui nous permettraient de traduire d'un point de vue informatique graphique les éléments définis dans le cadre conceptuel. Nous avons par la suite proposé un processus permettant la modélisation de clones virtuels à bas coût. L'étape suivante consista à étudier les poids respectifs des composantes de la fidélité visuelle et l'effet de la familiarité à travers une évaluation perceptive.

Finalement nous avons commencé à mettre en place les perspectives de la recherche sur la fidélité comportementale à travers une proposition d'une architecture d'un modèle computationnel guidant une expressivité émotionnelle similaire à celle d'un humain virtuel.

10.3 Contributions

10.3.1 Contributions théoriques

Cadre conceptuel pour l'étude de la fidélité. La première contribution de nos travaux de thèse fut la prestation d'un cadre conceptuel pour l'étude de la fidélité. À travers la définition de ce cadre, nous avons proposé une définition de la fidélité adéquate à l'étude de visages virtuels en 3D et nous avons procédé à l'identification de ses principales dimensions. Nous avons dégagé les différentes variables intrinsèques aux participants ou dépendantes de l'environnement qui influencent l'évaluation de la fidélité d'un clone virtuel (figure 3.1). Les futures études expérimentales perceptives devront prendre en compte la complexité de

ce cadre pour mettre en place des protocoles d'expérimentation qui limiteront les biais lors du jugement de la fidélité d'un clone virtuel avec son référent.

Étude de la fidélité comportementale. Dans l'optique de déterminer les éléments cognitifs représentatifs d'un référent, nous avons étudié la notion d'identité. Nous avons découvert que l'identité est un concept hétéroclite qui englobe un nombre important de notions tels que les self-concepts, la personnalité, les croyances et les buts, mais aussi la construction sociale de l'individu. L'identité possède diverses dimensions qui s'influencent mutuellement. En effet, l'identité peut être personnelle, ethnique, sociale, mais aussi émotionnelle. Nous nous sommes par conséquent focalisés sur l'identité émotionnelle à travers l'étude des différences interindividuelle dans l'émergence et l'expression des émotions.

Par la suite nous nous sommes intéressés à regarder les approches théoriques de l'émotion ainsi qu'aux modèles computationnels.

Dans le monde de l'intelligence artificielle, c'est l'approche d'évaluation cognitive qui est la plus traduite en modèle informatique. De plus, une des principales hypothèses sur lesquelles se basent ces théories est que des personnes différentes peuvent réagir de manières distinctes à un même stimulus, et ce en fonction de leur évaluation de cet événement. Cela nous a permis de conclure que l'approche cognitive des émotions serait la plus adéquate pour la simulation fidèle des processus émotionnels.

Pour aborder l'approche cognitive des émotions, nous avons décidé de nous inspirer du modèle CPM et de proposer un modèle dynamique de l'émotion qui prend en compte des différences inter-individuelles assurant ainsi la fidélité par rapport au référent. En effet, notre approche consiste à considérer des questionnaires de personnalité ou de traits émotionnels et de les inclure dans le modèle CPM pour moduler la valeur associée aux différents SECs tout au long du processus d'évaluation.

Nous avons opté pour le modèle CPM, car ce dernier permet de considérer différents éléments cognitifs tels que les self-concepts, les croyances et d'autres différences interindividuelles tout au long du processus responsable de la genèse de l'émotion.

Pour spécifier notre module, nous avons décidé dans une première étape de proposer une expérience d'évaluation qui simule le déroulement de l'évaluation des différents SEC par une annotation manuelle des scripts du corpus GEMEP (Bänziger & Scherer, 2010). Nous prévoyons dans les prochains mois d'implémenter ce système et d'effectuer les tests perceptifs sur deux différents clones. Ce qui nous permettrait d'explorer la fidélité comportementale à travers l'expressivité émotionnelle faciale. Serions-nous capables de détecter une

identité sur la base de ces réactions émotionnelle face à un script ?

10.3.2 Contributions techniques

Dans la perspective de réaliser le rendu et l’animation en temps réel des différentes composantes de la fidélité visuelle définies dans notre cadre conceptuel, nous avons proposé un processus permettant la capture, l’édition et le rendu d’un clone virtuel dans le moteur de jeu Unity3D. Au cours de nos travaux, nous avons dû modéliser un clone virtuel d’une personne de notre entourage afin d’avoir la possibilité d’explorer la notion de familiarité. Nous avons par conséquent proposé une méthode accessible (figure 5.6), nous permettant l’édition d’un clone virtuel animé en temps réel à bas coût et ne nécessitant pas l’utilisation de scanner 3D professionnel.

Notre technique est basée d’une part sur un scan ”Kinect” pour la capture de la géométrie et de techniques de capture de mouvement (à l’aide du dispositif Optitrack) pour la capture de la résolution temporelle.

Le rendu et l’animation sont effectués en temps réel dans le moteur de jeu Unity3D. Différentes études sur l’évaluation de visages de clones peuvent être envisagées sur la base de ces travaux.

En comparaison aux dernières avancées réalisées dans l’édition d’humain virtuel notre technique présente certes un niveau de réalisme bien inférieur à celui des travaux récents tels que ceux de (von der Pahlen et al., 2014). Cependant, notre technique propose une solution accessible qui permet l’édition de clones virtuels familiers (faisant partie de notre entourage). Cela à bas coût puisqu’elle ne nécessite pas l’utilisation de dispositifs de capture onéreux.

10.3.3 Contribution expérimentale

Nous avons mené une étude perceptive sur la fidélité visuelle d’un clone virtuel. À partir de notre cadre conceptuel de l’étude de la fidélité et les différents travaux sur la perception du visage virtuel, nous avons dégagé quelques variables qui nous ont apparus comme les plus pertinentes pour l’étude de la fidélité.

En nous basant sur ces variables, nous avons mis en place un protocole pour l’évaluation de la fidélité visuelle. Nous avons développé différents degrés de réalisme pour étudier l’interaction entre la fidélité expressive et la fidélité d’apparence.

Nous avons constaté lors de notre étude de l’état de l’art que la communauté de psychologues octroie à l’expressivité un rôle dans le processus d’identification et par conséquent dans le jugement de fidélité. Cependant ces recherches soutiennent, pour la plupart, que l’information dynamique joue un rôle significatif quand l’information d’apparence est partielle ou altérée. Cependant le rôle de l’expressivité pour l’identification de visage virtuel a été très peu exploré et

demeure même incertain. (Bennetts et al., 2013).

Les résultats que nous avons obtenus suggèrent que la perception de clone virtuel diffère bien de la perception de visage virtuel anonyme, en effet nous avons constaté que de nouvelles variables telles que la familiarité avec le référent influent non seulement sur la perception de la fidélité, mais aussi sur le degré d'appréciation du clone virtuel.

Nous avons constaté qu'une partie de nos résultats confirme et sont en accord avec certains résultats qu'on retrouve dans la littérature psychologique.

En effet, nous avons confirmé l'hypothèse que l'expressivité renforce un jugement de la fidélité. Les stimuli en condition dynamique jouissent de taux de fidélité bien supérieurs à une expression statique. Ces résultats sont dans la continuité des études psychologiques (Lander et al., 1999).

En modifiant le degré d'abstraction de l'information d'apparence, nous avons également remarqué que l'impact de l'expressivité diffère selon la qualité de l'information d'apparence mise en place.

Nous avons confirmé que plus l'information d'apparence est altérée et incomplète plus l'information dynamique jouait un rôle important. De plus nos résultats suggèrent que l'information dynamique seule peut véhiculer un sentiment de fidélité. Par conséquent, les signatures dynamiques idiosyncrasiques du référent jouent un rôle important lors d'une tâche d'évaluation de visages virtuels. Ces résultats sont cohérents avec l'hypothèse de *l'information supplémentaire* vue précédemment dans l'état de l'art (O'Toole et al., 2002).

Par contre en ce qui concerne la primauté de l'information d'apparence sur l'information dynamique nous avons noté que l'information d'apparence a été mise en second plan dans certaines tâches. Deux niveaux de rendus différents ont été perçus comme similaires, car ils déployaient la même animation. Nos résultats suggèrent que nous devrions donner la même importance aux indices dynamiques que ceux de l'apparence pour renforcer le sentiment de fidélité chez des observateurs.

Nous avons également constaté que l'affichage des niveaux équivalents de réalisme d'expressivité et le rendu est essentiel pour la transmission de l'impression de fidélité. Un décalage du niveau réalisme d'une modalité par rapport à l'autre nuirait au sentiment de "likeability" qui s'est avéré être fortement corrélé au jugement de fidélité. Autrement dit, si nous déprécions un rendu il nous semble moins fidèle.

10.4 Limites de nos travaux

Nos travaux se restreignent à des expressions faciales émotionnelles. Nos travaux se sont uniquement articulés autour de l'animation faciale. Nous ne considérons pas les expressions corporelles qui constitue elle aussi une source d'évaluation de la fidélité. Cette thèse n'aborde pas non plus les problématiques de la synthèse

audiovisuelle de la parole ou la problématique de la direction du regard. Les travaux sur lesquels nous avons mené nos recherches sont complexes et sont influencés par plusieurs différences interindividuelles. En effet, la littérature soutient que de nombreuses caractéristiques personnelles influent considérablement la perception du réalisme et de la familiarité telle que : l'âge, le sexe, l'expérience avec les médias virtuels, le fait d'être en « exogroupe » ou « endogroupe » avec le référent. Tous ces aspects n'ont pas été explorés lors de nos différentes études.

10.5 Perspectives

10.5.1 Profils expressifs

Au cours de nos travaux nous nous sommes contentés de faire varier la variable de la dynamique entre les deux conditions limite : statique et rejoue des données de capture de mouvement. Nous pourrions considérer une échelle plus fine et faire varier les différentes composantes de l'animation afin d'étudier l'impact du réalisme expressif sur le jugement de la fidélité et de l'appréciation du clone virtuel. Dans la même optique sur la variabilité de l'expressivité. Dans des recherches futures, nous visons à établir un profil d'expression qui décrit la signature de mouvement d'un référent. Nous aurions à faire varier les composantes de ce profil pour identifier les principaux indices qui assurent la fidélité expressive. Pour cela des premières analyses sur la base de corpus vidéos d'expression émotionnelle ont été réalisées (sur la base du corpus GEMEP (Bänziger & Scherer, 2010)). L'idée est de délaisser ce qui est en commun à toutes les expressions faciales d'une même émotion et d'extraire les variables qui seraient spécifiques à un sujet en particulier qui constituerait son profil expressif.

10.5.2 Fidélité comportementale

Dans le cadre de la formalisation de l'étude de la fidélité comportementale, la prochaine étape consistera à implémenter le modèle computationnel personnalisé de l'émotion (animation sur la base d'annotations manuelles). Il faudra, par la suite, effectuer une étude perceptive permettant de comparer la perception de la fidélité de l'animation générée par le modèle computationnel en comparaison avec le rejoue des données de capture de mouvement sur la base de l'évaluation des mêmes scripts du corpus de GEMEP.

Dans une seconde étape, pour permettre une interaction plus riche et à plus long terme, l'évaluation de chacun des SEC doit être générée à l'aide d'un système logique et une base de connaissances. L'intégration d'un modèle du type belief-desire-intention (BDI) permettrait par exemple de modéliser l'état

interne de l'agent virtuel et apporter des informations importantes qui renforceraient l'individualité de la réaction émotionnelle et par conséquent sa fidélité. Le modèle BDI permettrait donc de remplacer l'annotation manuelle que nous avons effectuée sur les scripts du corpus GEMEP.

La deuxième phase de l'étude de la fidélité comportementale concerne l'analyse de l'impact et l'intérêt d'un clone virtuel dans un contexte d'interaction. Cette architecture pourrait être mise en œuvre dans un moteur de jeu tel que Unity3D ou encore dans la plateforme MARC (Courgeon, 2011) qui intègre un module d'animation faciale basé sur CPM.

Quelques études ont étudié l'impact de la ressemblance de l'avatar virtuel sur notre comportement lors de l'interaction avec un clone virtuel. Par exemple une étude de (Van der Land, 2014) ont étudié l'impact de la similarité de l'avatar sur la réalisation de tâches en équipe. Une autre étude (Vugt, Bailenson, Hoorn, & Konijn, 2010) a exploré l'interaction avec un agent virtuel dont le visage a été morphée à différents degrés avec une représentation en 3D du sujet. Leurs résultats suggèrent que nous avons tendance à faire plus confiance aux avatars qui nous ressemblent. Il serait par conséquent intéressant d'étudier les apports et les avantages d'interagir avec un clone virtuel plutôt qu'avec un agent ne possédant pas une identité. Une étude dans un contexte d'interaction doit être menée pour voir s'il existe une modification du comportement de l'utilisateur ou de ses performances lors d'une tâche interactive avec un clone virtuel qui lui est familier.

10.5.3 Perspectives expérimentales

Le troisième bloc de notre paradigme pour l'étude de la fidélité énumère plusieurs variables extérieures qui influencent le jugement de la fidélité. Au cours de nos travaux, nous nous sommes focalisés essentiellement sur la variable de la familiarité et celle du réalisme. Pourtant d'autres différences inter-individuelles devraient également être prises en considération. Par exemple l'âge, le sexe, l'origine ethnique et l'expérience préalable avec des médias virtuels sont des variables potentiellement importantes pour l'évaluation de la fidélité. De nombreuses études antérieures ont prouvé qu'elles avaient un impact important sur l'évaluation de visages virtuels (Fan et al., 2012; Sinha et al., 2006). De plus, nous pourrions également considérer les conditions environnementales telles que différents points de vue ou des changements d'éclairage.

Nous prévoyons également de considérer d'autres clones virtuels pour évaluer la fidélité dans un contexte dépendant d'une tâche (par exemple, une tâche de reconnaissance avec des identités différentes en fonction de plusieurs clones virtuels). Cela aura également pour intérêt de faire varier les particularités

du référent. En effet, nous devons également faire varier le genre, l'âge et les caractéristiques du référent.

De plus, une des perspectives des plus intéressantes que nous permettra cette multiplication de stimuli est d'envisager la perception de soi en faisant évaluer le clone virtuel par le référent lui-même à la manière des différentes études sur le self (Bailenson, Beall, Blascovich, Raimundo, & Weisbuch, 2001 ; Bailenson & Segovia, 2010).

Références

- Alexander, A. L., Brunyé, T., Sidman, J., & Weil, S. A. (2005). From gaming to training : A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games. *DARWARS Training Impact Group*, 5, 1–14.
- Alexander, O., Fyffe, G., Busch, J., Yu, X., Ichikari, R., Jones, A., ... others (2013). Digital ira : Creating a real-time photoreal digital actor. In *Acm siggraph 2013 posters* (p. 1).
- Alexander, O., Rogers, M., Lambeth, W., Chiang, J.-Y., Ma, W.-C., Wang, C.-C., & Debevec, P. (2010). The digital emily project : Achieving a photorealistic digital actor. *Computer Graphics and Applications, IEEE*, 30(4), 20–31.
- Allbeck, J., & Badler, N. (2002). Toward representing agent behaviors modified by personality and emotion. *Embodied Conversational Agents at AAMAS*, 2, 15–19.
- Allen, J. A., Hays, R. T., & Buffardi, L. C. (1986). Maintenance training simulator fidelity and individual differences in transfer of training. *Human Factors : The Journal of the Human Factors and Ergonomics Society*, 28(5), 497–509.
- Allport, G. W. (1961). Pattern and growth in personality.
- Amor, B. B., Ardabilian, M., & Chen, L. (2005). 3d face modeling based on structured-light assisted stereo sensor. In *Image analysis and processing-iciap 2005* (pp. 842–849). Springer.
- Arellano, D., Varona, J., Perales, F. J., Bee, N., Janowski, K., & André, E. (2011). Influence of head orientation in perception of personality traits in virtual agents. In *The 10th international conference on autonomous agents and multiagent systems-volume 3* (pp. 1093–1094).
- Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 289–304.
- Arnold, M. B. (1960). Emotion and personality.
- Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization.

- Academy of management review*, 14(1), 20–39.
- Averill, J. R. (1999). Individual differences in emotional creativity : Structure and correlates. *Journal of personality*, 67(2), 331–371.
- Bailenson, J. N. (2012). Doppelgangers-a new form of self? *PSYCHOLOGIST*, 25(1), 36–38.
- Bailenson, J. N., Beall, A. C., Blascovich, J., Raimundo, M., & Weisbuch, M. (2001). Intelligent agents who wear your face : Users' reactions to the virtual self. In *Intelligent virtual agents* (pp. 86–99).
- Bailenson, J. N., Blascovich, J., & Guadagno, R. E. (2008). Self-representations in immersive virtual environments1. *Journal of Applied Social Psychology*, 38(11), 2673–2690.
- Bailenson, J. N., & Segovia, K. Y. (2010). Virtual doppelgangers : Psychological effects of avatars who ignore their owners. In *Online worlds : Convergence of the real and the virtual* (pp. 175–186). Springer.
- Bailenson, J. N., & Yee, N. (2005). Digital chameleons automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological science*, 16(10), 814–819.
- Bänziger, T., & Scherer, K. R. (2010). Introducing the geneva multimodal emotion portrayal (gemep) corpus. *Blueprint for affective computing : A sourcebook*, 271–294.
- Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2009). My robotic doppelgänger-a critical look at the uncanny valley. In *Robot and human interactive communication, 2009. ro-man 2009. the 18th ieee international symposium on* (pp. 269–276).
- Bartz, J. F. D. C. D., & StraBer, C. W. H. B. W. (2006). Measuring the discernability of virtual objects in conventional and stylized augmented reality. In *12th eurographics symposium on virtual environments, lisbon, portugal, may 8th-10th, 2006* (p. 53).
- Bennetts, R. J., Kim, J., Burke, D., Brooks, K. R., Lucey, S., Saragih, J., & Robbins, R. A. (2013). The movement advantage in famous and unfamiliar faces : a comparison of point-light displays and shape-normalised avatar stimuli. *Perception*, 42(9), 950–970.
- Berisha, F., Johnston, A., & McOwan, P. W. (2010). Identifying regions that carry the best information about global facial configurations. *Journal of vision*, 10(11), 27.
- Bickmore, T., & Giorgino, T. (2006). Health dialog systems for patients and consumers. *Journal of biomedical informatics*, 39(5), 556–571.
- Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2), 293–327.
- Braje, W. L., Kersten, D., Tarr, M. J., & Troje, N. F. (1998). Illumination effects in face recognition. *Psychobiology*, 26(4), 371–380.

- Brockhaus, M. (2013). *Human visual processing of natural facial motion analyzed by fmri* (Thèse de doctorat non publiée). Universität Tübingen.
- Calder, A., Rhodes, G., Johnson, M., & Haxby, J. (2011). *Oxford handbook of face perception*. Oxford University Press.
- Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The “big five questionnaire” : A new questionnaire to assess the five factor model. *Personality and individual differences*, 15(3), 281–288.
- Cassell, J. (2000). *Embodied conversational agents*. MIT press.
- Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & cognition*, 26(4), 780–790.
- Cook, R., Johnston, A., & Heyes, C. (2012). Self-recognition of avatar motion : how do i know it’s me? *Proceedings of the Royal Society B : Biological Sciences*, 279(1729), 669–674.
- Courgeon, M. (2011). *Marc : modèles informatiques des émotions et de leurs expressions faciales pour l’interaction homme-machine affective temps réel* (Thèse de doctorat non publiée). Université Paris Sud-Paris XI.
- Damian, I., Endrass, B., Huber, P., Bee, N., & André, E. (2011). Individualized agent interactions. In *Motion in games* (pp. 15–26). Springer.
- Debevec, P. (2012). The light stages and their applications to photoreal digital actors. *SIGGRAPH Asia Technical Briefs*.
- Debevec, P., Hawkins, T., Tchou, C., Duiker, H.-P., Sarokin, W., & Sagar, M. (2000). Acquiring the reflectance field of a human face. In *Proceedings of the 27th annual conference on computer graphics and interactive techniques* (pp. 145–156).
- Deng, Z., Chiang, P.-Y., Fox, P., & Neumann, U. (2006). Animating blendshape faces by cross-mapping motion capture data. In *Proceedings of the 2006 symposium on interactive 3d graphics and games* (pp. 43–48).
- d’Eon, E., Luebke, D., & Enderton, E. (2007). Efficient rendering of human skin. In *Proceedings of the 18th eurographics conference on rendering techniques* (pp. 147–157).
- Di Loreto, I., Hervouet, F., et al. (2012). Identity : the contribution of contemporary philosophy to agent design.
- Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures inter-and intranational differences. *Journal of personality and social psychology*, 81(5), 869.
- Ekman, P., & Friesen, W. V. (1977). Facial action coding system.
- Ekman, P., & Friesen, W. V. (2003). *Unmasking the face : A guide to recognizing emotions from facial clues*. Ishk.
- El Rhalibi, A., Carter, C., Cooper, S., & Merabti, M. (2011). Highly realistic mpeg-4 compliant facial animation with charisma. In *Computer communications and networks (icccn), 2011 proceedings of 20th international conference on* (pp. 1–6).

- Fan, S., Ng, T.-T., Herberg, J. S., Koenig, B. L., & Xin, S. (2012). Real or fake? : human judgments about photographs and computer-generated images of faces. In *Siggraph asia 2012 technical briefs* (p. 17).
- Fan, S., Wang, R., Ng, T.-T., Tan, C. Y.-C., Herberg, J. S., & Koenig, B. L. (2014). Human perception of visual realism for photo and computer-generated face images. *ACM Transactions on Applied Perception (TAP)*, 11(2), 7.
- Fiske, S. T., & Taylor, S. E. (2013). *Social cognition : From brains to culture*. Sage.
- Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition and emotion*, 1(2), 115–143.
- Fyffe, G., Jones, A., Alexander, O., Ichikari, R., & Debevec, P. (2014). Driving high-resolution facial scans with video performance capture. *ACM Transactions on Graphics (TOG)*, 34(1), 8.
- Gebhard, P. (2005). Alma : a layered model of affect. In *Proceedings of the fourth international joint conference on autonomous agents and multiagent systems* (pp. 29–36).
- Giard, F., & Guitton, M. J. (2010). Beauty or realism : The dimensions of skin from cognitive sciences to computer graphics. *Computers in Human Behavior*, 26(6), 1748–1752.
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, 45(1), 32–41.
- Goffman, E. (2002). The presentation of self in everyday life [1959]. *Contemporary sociological theory*, 46–61.
- Goffman, E. (2009). *Stigma : Notes on the management of spoiled identity*. Simon and Schuster.
- Goffman, E. (2012). The presentation of self in everyday life [1959]. *Contemporary sociological theory*, 46–61.
- Gooch, B., Reinhard, E., & Gooch, A. (2004). Human facial illustrations : Creation and psychophysical evaluation. *ACM Transactions on Graphics (TOG)*, 23(1), 27–44.
- Green, R. D., MacDorman, K. F., Ho, C.-C., & Vasudevan, S. (2008). Sensitivity to the proportions of faces that vary in human likeness. *Computers in Human Behavior*, 24(5), 2456–2474.
- Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K., & Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human-Computer Studies*, 67(10), 842–849.
- Guadagno, R. E., Blascovich, J., Bailenson, J. N., & Mccall, C. (2007). Virtual humans and persuasion : The effects of agency and behavioral realism. *Media Psychology*, 10(1), 1–22.
- Gunes, H., Schuller, B., Pantic, M., & Cowie, R. (2011). Emotion representation, analysis and synthesis in continuous space : A survey. In *Automatic*

- face and gesture recognition and workshops (fg 2011), 2011 ieee international conference on* (pp. 827–834).
- Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current opinion in neurobiology*, *14*(2), 233–238.
- Hernandez, M., Choi, J., & Medioni, G. (2012). Laser scan quality 3-d face modeling using a low-cost depth camera. In *Signal processing conference (eusipco), 2012 proceedings of the 20th european* (pp. 1995–1999).
- Hole, G. J., George, P. A., Eaves, K., Rasek, A., et al. (2002). Effects of geometric distortions on face-recognition performance. *Perception-London*, *31*(10), 1221–1240.
- Hudlicka, E. (2011). Guidelines for designing computational models of emotions. *International Journal of Synthetic Emotions (IJSE)*, *2*(1), 26–79.
- Izard, C. E. (2013). *Human emotions*. Springer Science & Business Media.
- Jimenez, J. (2012). *Practical real-time strategies for photorealistic skin rendering and antialiasing* (Thèse de doctorat non publiée). Jimenez.
- Jimenez, J., Sundstedt, V., & Gutierrez, D. (2009). Screen-space perceptual rendering of human skin. *ACM Transactions on Applied Perception (TAP)*, *6*(4), 23.
- Kaba, F. (2013). Hyper realistic characters and the existence of the uncanny valley in animation films. *International Review of Social Sciences and Humanities*, *4*(3), 188–195.
- Kang, S.-H., Gratch, J., Wang, N., & Watt, J. H. (2008). Does the contingency of agents' nonverbal feedback affect users' social anxiety? In *Proceedings of the 7th international joint conference on autonomous agents and multiagent systems-volume 1* (pp. 120–127).
- Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., & Piepol, D. (2007). Building interactive virtual humans for training environments. In *Proceedings of i/itsec* (Vol. 174).
- King, L. A., & Emmons, R. A. (1990). Conflict over emotional expression : Psychological and physical correlates. *Journal of personality and social psychology*, *58*(5), 864.
- Knappmeyer, B., Thornton, I. M., & Bühlhoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision research*, *43*(18), 1921–1936.
- Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual cognition*, *4*(3), 265–273.
- Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual differences in dispositional expressiveness : development and validation of the emotional expressivity scale. *Journal of personality and social psychology*, *66*(5), 934.
- Lander, K., & Bruce, V. (2000). Recognizing famous faces : Exploring the benefits of facial motion. *Ecological Psychology*, *12*(4), 259–272.

- Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory & cognition*, 27(6), 974–985.
- Lander, K., & Chuang, L. (2005). Why are moving faces easier to recognize? *Visual Cognition*, 12(3), 429–442.
- Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference characteristic : A review. *Journal of Research in personality*, 21(1), 1–39.
- Levita, D. J., & Ruenke, H. (1965). *The concept of identity*. Mouton.
- Li, H., Weise, T., & Pauly, M. (2010). Example-based facial rigging. In *Acm transactions on graphics (tog)* (Vol. 29, p. 32).
- Lisetti, C., Amini, R., & Yasavur, U. (2015). Now all together : Overview of virtual health assistants emulating face-to-face health interview experience. *KI-Künstliche Intelligenz*, 1–12.
- Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). The computer expression recognition toolbox (cert). In *Automatic face & gesture recognition and workshops (fg 2011), 2011 ieee international conference on* (pp. 298–305).
- Liu, Z., Zhang, Z., Cohen, M. F., & Jacobs, C. E. (2004, octobre 19). *Rapid computer modeling of faces for animation*. Google Patents. (US Patent 6,807,290)
- MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? uncanny responses to computer generated faces. *Computers in human behavior*, 25(3), 695–710.
- Malatesta, L., Raouzaïou, A., Karpouzis, K., & Kollias, S. (2009). Towards modeling embodied conversational agent character profiles using appraisal theory predictions in expression synthesis. *Applied intelligence*, 30(1), 58–64.
- Marinier, R. P., Laird, J. E., & Lewis, R. L. (2009). A computational unification of cognitive behavior and emotion. *Cognitive Systems Research*, 10(1), 48–69.
- Markus, H., & Nurius, P. (1986). Possible selves. *American psychologist*, 41(9), 954.
- Marsella, S. (2010). Modeling emotion and its expression in virtual humans. In *User modeling, adaptation, and personalization* (pp. 1–2). Springer.
- Marsella, S., Gratch, J., & Petta, P. (2010). Computational models of emotion. *A Blueprint for Affective Computing-A sourcebook and manual*, 21–46.
- Marsella, S. C., & Gratch, J. (2009). Ema : A process model of appraisal dynamics. *Cognitive Systems Research*, 10(1), 70–90.
- McAdams, D. P., & Pals, J. L. (2006). A new big five : fundamental principles for an integrative science of personality. *American Psychologist*, 61(3), 204.
- McDonnell, R., Breidt, M., & Bülthoff, H. H. (2012). Render me real? : investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics (TOG)*, 31(4), 91.

- Mehlmann, G., & André, E. (2015). Games are better than books : In-situ comparison of an interactive job interview game with conventional training. *Artificial Intelligence in Education : 17th International Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings*, 9112, 84.
- Mehrabian, A. (1996). Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology*, 48(2), 86–92.
- Metallo, A., Rossi, V., Blundell, J., Waibel, G., Graham, P., Fyffe, G., ... Debevec, P. (2015). Scanning and printing a 3d portrait of president barack obama. In *Siggraph 2015 : Studio* (p. 19).
- Mori, M. (1970). Bukimi no tani [the uncanny valley]. *energy*, 7 (4) 33-35.(translated by karl f. macdorman and takashi minato in 2005 within appendix b for the paper androids as an experimental apparatus : Why is there an uncanny and can we exploit it? In *Proceedings of the cogsci-2005 workshop : Toward social mechanisms of android science* (pp. 106–118).
- Murakami, Y., Sugimoto, Y., & Ishida, T. (2005). Modeling human behavior for virtual training systems. In *Proceedings of the national conference on artificial intelligence* (Vol. 20, p. 127).
- Newell, A. (1994). *Unified theories of cognition*. Harvard University Press.
- Nguyen, H. (2007). *Gpu gems 3*. Addison-Wesley Professional.
- Ochs, M., Niewiadomski, R., Brunet, P., & Pelachaud, C. (2012). Smiling virtual agent in social context. *Cognitive processing*, 13(2), 519–532.
- Ortony, A. (1990). *The cognitive structure of emotions*. Cambridge university press.
- O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces : A psychological and neural synthesis. *Trends in cognitive sciences*, 6(6), 261–266.
- Paleari, M., Grizard, A., & Lisetti, C. L. (2007). Adapting psychologically grounded facial emotional expressions to different anthropomorphic embodiment platforms. In *Flairs conference* (pp. 565–570).
- Pandzic, I. S., & Forchheimer, R. (2002). Mpeg-4 facial animation. *The standard, implementation and applications*. Chichester, England : John Wiley&Sons.
- Panksepp, J. (1998). *Affective neuroscience : The foundations of human and animal emotions*. Oxford university press.
- Parke, F. I. (1972). Computer generated animation of faces. In *Proceedings of the acm annual conference-volume 1* (pp. 451–457).
- Parke, F. I., & Waters, K. (2008). *Computer facial animation*. CRC Press.
- Pasquariello, S., & Pelachaud, C. (2002). Greta : A simple facial animation engine. In *Soft computing and industry* (pp. 511–525). Springer.
- Pike, G. E., Kemp, R. I., Towell, N. A., & Phillips, K. C. (1997). Recognizing moving faces : The relative contribution of motion and perspective view information. *Visual Cognition*, 4(4), 409–438.

- Platt, S. M., & Badler, N. I. (1981). Animating facial expressions. In *Acm siggraph computer graphics* (Vol. 15, pp. 245–252).
- Posdamer, J., & Altschuler, M. (1982). Surface measurement by space-encoded projected beam systems. *Computer graphics and image processing*, 18(1), 1–17.
- Puklavage, C., Pirela, A., Gonzalez, A. J., & Georgiopoulos, M. (2010). Imitating personalized expressions in an.
- Ramamurthy, B., Lewis, B., & Duchowski, A. T. (2012). Eye tracking to enhance facial recognition algorithms. In *30th acm conference on human factors in computing systems*.
- Rapsak, S. Z. (2003). Face memory and its disorders. *Current neurology and neuroscience reports*, 3(6), 494–501.
- Revelle, W., & Scherer, K. R. (2009). Personality and emotion. *Oxford companion to emotion and the affective sciences*, 304–306.
- Ring, L., Utami, D., & Bickmore, T. (2014). The right agent for the job? In *Intelligent virtual agents* (pp. 374–384).
- Rizzo, A. A., Neumann, U., Enciso, R., Fidaleo, D., & Noh, J. (2001). Performance-driven facial animation : basic research on human judgments of emotional state in facial avatars. *CyberPsychology & Behavior*, 4(4), 471–487.
- Robinette, K. M., Daanen, H., & Paquet, E. (1999). The caesar project : a 3-d surface anthropometry survey. In *3-d digital imaging and modeling, 1999. proceedings. second international conference on* (pp. 380–386).
- Rodríguez, L.-F., & Ramos, F. (2014). Development of computational models of emotions for autonomous agents : a review. *Cognitive Computation*, 6(3), 351–375.
- Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research*, 47(2), 157–165.
- Sadrô, J., Jarudi, I., & Sinha, P. (2003). The role of eyebrows in face recognition. *Perception*, 32(3), 285–293.
- Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C., Palfai, T. P., et al. (1995). Emotional attention, clarity, and repair : Exploring emotional intelligence using the trait meta-mood scale. *Emotion, disclosure, and health*, 125, 154.
- Santella, A., & DeCarlo, D. (2004). Visual interest and npr : an evaluation and manifesto. In *Proceedings of the 3rd international symposium on non-photorealistic animation and rendering* (pp. 71–150).
- 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. (2009). The dynamic architecture of emotion : Evidence for the component process model. *Cognition and emotion*, 23(7), 1307–1351.

- Scherer, K. R., & Ekman, P. (1984). On the nature and function of emotion : A component process approach. *Approaches to emotion*, 2293, 317.
- Scherer, K. R., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion*, 7(1), 113.
- Seyama, J., & Nagayama, R. S. (2007). The uncanny valley : Effect of realism on the impression of artificial human faces. *Presence : Teleoperators and Virtual Environments*, 16(4), 337–351.
- Sifakis, E., Neverov, I., & Fedkiw, R. (2005). Automatic determination of facial muscle activations from sparse motion capture marker data. In *Acm transactions on graphics (tog)* (Vol. 24, pp. 417–425).
- Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans : Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, 94(11), 1948–1962.
- Smith, G. (2011). Self, self-concept, and identity. *Handbook of Self and Identity*, 69.
- Stevenage, S. V. (1995). Can caricatures really produce distinctiveness effects? *British Journal of Psychology*, 86(1), 127–146.
- Struck, F., Bohn, C.-A., Schmidt, S., & Helzle, V. (2004). Realistic shading of human skin in real time. In *Proceedings of the 3rd international conference on computer graphics, virtual reality, visualisation and interaction in africa* (pp. 93–97).
- Sun, Q., Tang, Y., Hu, P., & Peng, J. (2012). Kinect-based automatic 3d high-resolution face modeling. In *Image analysis and signal processing (iasp), 2012 international conference on* (pp. 1–4).
- Tajfel, H. (1981). *Human groups and social categories : Studies in social psychology*. CUP Archive.
- Tanaka, J. W., & Gordon, I. (2011). Features, configuration and holistic face processing. *The Oxford handbook of face perception*, 177–194.
- Tong, J., Zhou, J., Liu, L., Pan, Z., & Yan, H. (2012). Scanning 3d full human bodies using kinects. *Visualization and Computer Graphics, IEEE Transactions on*, 18(4), 643–650.
- Van der Land, S. A. F. F. H. M. . V. d. H. B., S.F. (2014). Does avatar appearance matter? how team visual similarity and member-avatar similarity influence virtual team performance. *Human Communication Research*.
- von der Pahlen, J., Jimenez, J., Danvoye, E., Debevec, P., Fyffe, G., & Alexander, O. (2014). Digital ira and beyond : creating real-time photoreal digital actors. In *Acm siggraph 2014 courses* (p. 1).
- Vugt, H. C. V., Bailenson, J. N., Hoorn, J. F., & Konijn, E. A. (2010). Effects of facial similarity on user responses to embodied agents. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 17(2), 7.
- Wallis, G., & Bühlhoff, H. H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences*, 98(8),

- 4800–4804.
- Wallraven, C., Breidt, M., Cunningham, D. W., & Bühlhoff, H. H. (2005). Psychophysical evaluation of animated facial expressions. In *Proceedings of the 2nd symposium on applied perception in graphics and visualization* (pp. 17–24).
- Wallraven, C., Bühlhoff, H. H., Cunningham, D. W., Fischer, J., & Bartz, D. (2007). Evaluation of real-world and computer-generated stylized facial expressions. *ACM Transactions on Applied Perception (TAP)*, 4(3), 16.
- Waters, K. (1987). A muscle model for animation three-dimensional facial expression. In *Acm siggraph computer graphics* (Vol. 21, pp. 17–24).
- Weiss, A., Hirshberg, D., & Black, M. J. (2011). Home 3d body scans from noisy image and range data. In *Computer vision (iccv), 2011 ieee international conference on* (pp. 1951–1958).
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception-London*, 31(8), 995–1004.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16(6), 747–759.
- Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition a critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, 16(2), 116–142.
- Zaiane, O. R. (2002). Building a recommender agent for e-learning systems. In *Computers in education, 2002. proceedings. international conference on* (pp. 55–59).
- Zenasni, F., Lubart, T., Pahlavan, F., Jacob, S., Jacquet, A.-Y., & Lemoine, C. (2003). Traits émotionnels, intelligence émotionnelle : Intérêt de ces concepts et étude de leurs interrelations. *PSN*, 1(4), 19–27.
- Zhang, S. (2010). Recent progresses on real-time 3d shape measurement using digital fringe projection techniques. *Optics and lasers in engineering*, 48(2), 149–158.

Titre : Modélisation et Evaluation de la Fidélité d'un Clone Virtuel

Mots clés : Clone virtuel, Fidélité, Visage 3D, Informatique affective, Agent virtuel

Résumé : Les recherches et applications en humains virtuels ont connu un intérêt croissant pour les clones virtuels (des agents virtuels ayant un aspect familier ou du moins reconnaissable). Les clones virtuels sont de plus en plus répandus dans des interfaces homme-machine et dans les plus grands succès de l'industrie audio-visuelle. Les dernières techniques de rendu 3D et les dernières avancées des scans 3D ont permis la création de visages virtuels photo-réalistes modélisant des personnes réelles existantes. De plus, la tendance actuelle pour modéliser des humains virtuels est de se baser sur des techniques d'acquisition de données réelles (issues de scans et de sessions de capture de mouvement). L'identification des visages est primordiale lors de nos interactions sociales. Ainsi, notre comportement change suite à l'identification de la personne avec laquelle nous interagissons.

L'objectif de cette thèse consiste à explorer la réunion de ces deux axes de recherche en examinant le processus de perception de la fidélité d'un visage virtuel, clone d'une personne réelle

Nos travaux répondent à plusieurs questions de recherche : Quels sont les éléments qui nous permettent d'évaluer la ressemblance du clone virtuel avec son référent?

Parmi les multiples possibilités de techniques de rendu, d'animation et d'acquisition de données qu'offre l'informatique graphique, quelle est la meilleure combinaison pour assurer le plus haut degré de fidélité perçue ?

L'apparence visuelle n'est cependant qu'une des composantes qui interviennent dans la reconnaissance de personnes familières. Nos travaux s'intéressent également aux autres composantes qui comprennent ainsi l'expressivité mais aussi le traitement des connaissances que nous avons sur cette personne.

Nos contributions apportent des éléments de réponse à ces questions à plusieurs niveaux. Nous avons défini un cadre conceptuel identifiant les principaux concepts pertinents pour l'étude de la fidélité d'un visage virtuel. Nous avons aussi étudié l'aspect visuel de la fidélité à travers l'exploration de différentes techniques de rendu. Nous avons étudié dans une autre étape l'impact de la familiarité dans le jugement de la fidélité. Finalement, nous avons proposé les premières briques d'un modèle informatique individuel basé sur une approche cognitive des émotions qui permettrait de guider l'animation expressive du clone virtuel.

Title : Modeling and Evaluating the Fidelity of Virtual Clones

Keywords : Virtual Clone, Fidelity, 3D faces, Affective computing, Virtual agent

Abstract: Research and applications in virtual humans have experienced a growing interest in so-called virtual clones (agents with a familiar aspect or at least recognizable). Virtual clones are therefore increasingly used in human-machine interfaces and in the audiovisual industry.

Creating photorealistic and animated human-like face of a real person is now possible thanks to recent advances in Computer Graphics and 3D scan systems. Recent rendering techniques are challenging our ability to distinguish between computer generated faces and real human faces. Besides, the current trend to model virtual humans is to involve real data collected using scans and motion capture systems.

Face identification plays a crucial role in our daily social interactions. Indeed, our behavior changes according to the identification of the person with whom we interact. Moreover, several studies in Psychology and Neurosciences have observed that our cognitive processing of familiar faces is different from the cognitive processing of unfamiliar faces.

Studies about the perception and interaction with virtual clones are therefore required to better understand

how we should design and evaluate this kind of technology.

The main goal of this thesis is to explore this line of research. Our work rises several research questions: What are the features of the virtual clone that enable us to evaluate the resemblance between a virtual clone and its referent? Among several possibilities of rendering, animation and data acquisition techniques offered by Computer Graphics, what is the best combination of techniques to ensure the highest level of perceived fidelity? However, visual appearance is not the only component that is involved in recognizing familiar people. The other components include facial expressiveness but also the possible knowledge that we have about the referent.

Our contributions provide answers to these questions at several levels. We define a conceptual framework identifying the key concepts which are relevant for the study of the fidelity of a virtual face. We explore different rendering techniques. We describe an experimental study about the impact of familiarity in the judgment of fidelity. Finally, we propose a preliminary individual computational model based on a cognitive approach of emotions that could drive the animation of the virtual clone.



