

496

République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieur et de la recherche
scientifique

Université de 8 Mai 1945 - Guelma -
Faculté des Mathématiques, d'Informatique et des
Sciences de la matière
Département d'Informatique



Mémoire de Fin d'études Master
Filière : Informatique
Option : Ingénierie des Médias

45/884

Thème :

**Le transfert d'animation faciale à partir de
vidéo réelle**

Présenté par :
-Atamnia Hadjer
-Mohamedatni Leila

Encadré Par : Mme. Bordjiba
Yamina

Juin 2015

le transfert d'animation faciale à
partir de vidéo réelle

*A mes parents, mes frères, mes sœurs ,
mes amis et mes enseignants...*

Remerciement

Nous exprimons nos profonds remerciements à Madame bordjiba yamina, notre encadreur à l'université de Guelma 8 mai 1945, pour son suivi permanent, sa disponibilité et ses conseils qu'elle nous a apporté lors de notre travail. Nous remercions l'ensemble des membres du jury qui nous ont fait l'honneur d'accepter de lire et de juger ce mémoire.

Merci à nos amis et à nos proches pour leur soutien et leur encouragements et pour nous avoir supporté pendant ces quelques années. Nous remercions également nos parents de nous avoir toujours soutenu pendant toutes nos années d'études et pour nous avoir toujours encouragé à aller le plus loin possible. Nos sincères remerciements vont aussi à nos frères et à nos sœurs

Résumé

Dans le domaine de l'animation par ordinateur, l'humain virtuel a été le sujet de nombreuses recherches. De par son intérêt dans le processus de communication, le visage est un domaine particulièrement important. L'interactivité disponible dans le cadre d'applications en temps réel offre de vastes domaines d'applications à des telles technologies. Dans ce mémoire, nous présentons l'animation faciale qui est l'un des points clés dans le réalisme des scènes 3D qui mettent en scène des personnages virtuel, et on s'intéresse au visage et les nombreux muscles qui le composent, permettent de générer une multitude d'expressions. Le problème majeur de l'utilisation d'un système de paramétrisation MPEG4 réside dans la complexité de conception du lien entre ces paramètres et la topologie d'un modèle spécifique. Il existe différentes méthodes de déformation qui permettent de faire ce lien pour ceci, notre méthode effectue la déformation de maillage en utilisant les fonctions à base radiale (Radial Basis Function, RBF). Après avoir validé cette technique pour l'animation de visage, nous l'avons utilisée pour le transfert. Le transfert d'animation faciale est un processus assez complexe et difficile, nous souhaitons réaliser un transfert d'animation d'un visage 2D vers un visage 3D.

les mots clés : animation faciale ,Radial Basis Function,systeme de paramétrisation ,scènes 3D,expressions .

Abstract

In the field of computer animation, the virtual human has been the subject of much research. Through its interest in the communication process, the face is a particularly important area. The interactivity available through real-time applications offers vast areas of applications for such technologies. In this thesis, we present the facial animation is one of the key points in the realism of 3D scenes that feature virtual characters, and we look to the face and the many muscles that make up the generate a multitude of expressions. The major problem with the use of a parameterization MPEG4 system is the design complexity of the link between these parameters and the topology of a specific model. There are different methods of deformation that can make that connection for this, our method performs the mesh deformation using the radial basis functions (Radial Basis Function RBF). After validating this technique for facial animation, we used it for transfer. The facial animation transfer is quite a complex and difficult process, we want to make an animated face transfer of a 2D to a 3D face.

keywords :facial Animation,Radial Basis Function ,system of parameterization, expressions, 3D scene

TABLE DES MATIÈRES

Remerciement	iii
Résumé	iv
Abstract	v
Table des matières	viii
Table des figures	x
Table des tableaux	xi
Introduction Générale	1
1 l'animation faciale	3
1.1 Introduction	4
1.2 Historique	4
1.3 Les Caractéristiques faciales	5
1.3.1 la bouche	5
1.3.2 Les yeux et les sourcils	5
1.4 Expression Faciale Vs émotion	6
1.4.1 L'expression faciale	6
1.4.2 L'émotion	6
1.5 Les contraintes	7
1.6 L'animation faciale non temps réel	7
1.7 L'animation faciale temps réel	8
1.8 Les domaines d'applications	8
1.9 Conclusion	9
2 les techniques d'animation faciale	10
2.1 Introduction	11
2.2 Les processus de l'animation faciale	11
2.3 Le système de codage	12

TABLE DES FIGURES

1.1	Parke 1973 : Un des premiers modèles d'animation faciale	5
1.2	Générateurs de l'expression faciale et de l'émotion	6
2.1	Les 84 FDP définis par le standard MPEG-4. Les points pleins sont ceux influencés par les paramètres d'animation FAP. Les points creux sont fixes et sont utilisés pour le calibrage des visages. Les FDP sont groupés par zone (9 pour le nez et 3 pour les yeux par exemple).	13
2.2	Organisation des différentes techniques d'animation faciale	15
2.3	Exemple d'interpolation sur un maillage 3D	16
2.4	vue d'ensemble du processus	20
3.1	Les étapes de l'interpolation par RBF.	24
3.2	L'entrée : un maillage source et ses expressions sous forme de maillages, un maillage cible.	24
3.3	Initialisation des points de contrôle source et Cible.	25
3.4	de la transformation source vers cible.	25
3.5	du maillage cible selon les points de contrôle obtenus pour chaque expression.	25
3.6	Système d'expression cloning de Noh et al[41]	27
3.7	Vue d'ensemble du processus de clonage d'animations[42]	28
4.1	l'architecture générale de notre application	32
4.2	définition du maillage 3D	33
4.3	les points de contrôles selon FaceSDK	33
4.4	l'environnement Eclipse	37
4.5	interface FaceGen	37
4.6	visage cible : état neutre	39
4.7	l'attente de chargement de vidéo à partir d'une barre de progression	40
4.8	chargement de la vidéo préenregistré	41
4.9	classifieurs en cascade de haar	42
4.10	detection de visage	42
4.11	visage source neutre	43
4.12	visage source : avec expression	44

4.13 Extraits d'animation des visages source et cible.	45
4.14 test sur les deux visages source et cible	45

LISTE DES TABLEAUX

4.1 matériel utilisé	36
4.2 le choix des points de contrôle	43

Introduction Générale

1.1 Introduction

Animer signifie amener à la vie. L'action d'animer s'applique à tout changement pouvant avoir un effet visuel. Les variations des caractéristiques telles que la position, la couleur, la structure d'une entité dans le temps sont donc des animations de celle-ci. Une variation, dans le temps, du point de vue d'une scène produit également des effets visuels qui sont considèrent comme animation. Cette partie se propose de faire un état de l'art dont les premiers travaux remontent au tout début de l'informatique graphique. Dans ce chapitre, nous allons présenter des généralités sur le visage qui est une partie de l'anatomie qui comporte beaucoup de subtilités et dont les mouvements sont plus souvent des déformations que des déplacements rigides. En fait, le visage transmet plus d'émotions que l'ensemble du reste du corps. Il est donc justifié de considéré son animation de façon particulière. Tout comme pour l'animation en général, l'animation faciale consiste à produire des variations dans le temps de configurations du modèle de la face. Cela conduit à l'utilisation de certaines méthodes de l'animation telles que la modélisation de l'entité à animer, l'interpolation de positions intermédiaires à partir de positions clé, et la variation dans le temps de paramètres de contrôle des postures du modèle.

1.2 Historique

L'expression du visage humain a fait l'objet de la recherche scientifique depuis plus de cent ans. L'étude des mouvements du visage et des expressions a commencé à partir d'un point de vue biologique. Les premiers travaux sur ordinateur avec la représentation du visage a été fait dans les années 1970. Les premières animations faciale ont été créé par Parke en 1972. En 1973, Gillenson a développé un système interactif à assembler en ligne d'édition tirée d'images faciales. Et en 1974, Parke a développé un modèle en trois dimensions du visage paramétré.

Depuis 1990, l'animation faciale 3D est en plein essor, avec de nombreux films utilisant ce type de techniques et des réalisations marquantes comme " Toy's Story " de Pixar et plus récemment des exploits de réalisme avec " Bingo " de Chris Landreth.[1]

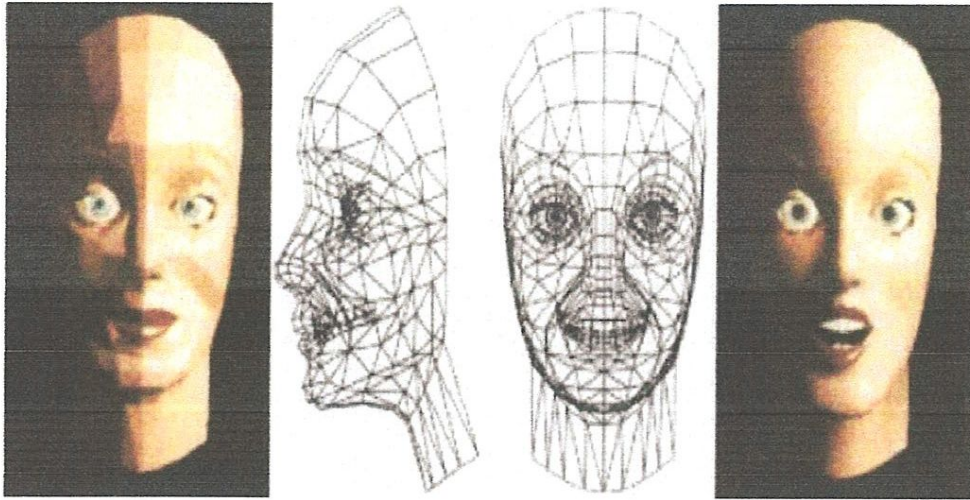


FIGURE 1.1 – Parke 1973 : Un des premiers modèles d'animation faciale

1.3 Les Caractéristiques faciales

Un modèle de visage est la réunion de plusieurs éléments et détails. La description brève des principaux éléments de description du visage, ainsi que leurs rôles respectifs dans la génération des expressions ou de la parole.

1.3.1 la bouche

La bouche est une des caractéristiques faciales les plus expressives, notamment grâce aux nombreux muscles l'entourant qui permettent des mouvements variés. Constituée de deux lèvres, inférieure et supérieure indépendante l'une de l'autre, elle permet à la fois d'exprimer la parole et les émotions. Notons que la lèvre inférieure est solidaire des mouvements de la mâchoire.[2]

1.3.2 Les yeux et les sourcils

Le mouvement des yeux est vraiment est très important pour transmettre des émotions et créer des expressions. Situés chacun dans une orbite du squelette, ils sont entourés de muscle permettant un positionnement précis de l'axe de vision et aidant à la détermination du rapport nécessaire entre les deux yeux pour la vision binoculaire. Mentionnons qu'une des règles habituellement prise en compte pour la fabrication d'un personnage est d'assurer la convergence des directions de regard pointées par les deux yeux. Du point de vue de la modélisation, les yeux sont représentés schématiquement par un globe, dont une partie est visible et l'autre non. Sur ce globe oculaire se trouvent l'iris et la pupille. Pour certains dessins animés simplifiés, les yeux des personnages sont réduits à une boule plus ou moins sphérique. Pour ajouter plus de vie au visage, il est possible de jouer sur la dilatation de la

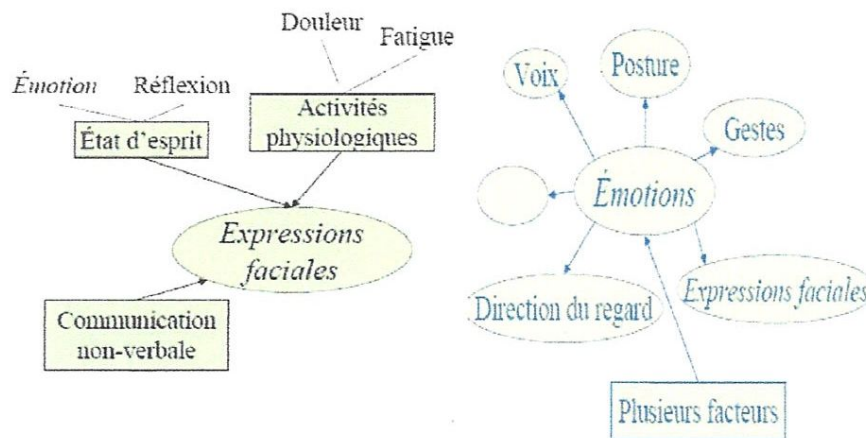


FIGURE 1.2 – Générateurs de l'expression faciale et de l'émotion

pupille, la coloration de l'iris, ou bien sur les réflexions de la lumière. Notons enfin que les sourcils peuvent accompagner les mouvements des yeux pour accentuer ainsi l'expression. [3]

1.4 Expression Faciale Vs émotion

1.4.1 L'expression faciale

Est une mimique faciale chargée de sens. Le sens peut être l'expression d'une émotion, un indice sémantique ou une intonation dans la Langue des Signes. L'interprétation d'un ensemble de mouvements musculaires en expression est dépendante du contexte d'application.

1.4.2 L'émotion

Expressions et émotions sont très liées et parfois confondues, l'émotion est un des générateurs des expressions faciales. L'émotion se traduit via de nombreux canaux comme la position du corps, la voix et les expressions faciales. Une émotion implique généralement une expression faciale correspondante, mais l'inverse n'est pas vrai : il est possible de mimer une expression représentant une émotion sans pour autant ressentir cette émotion. Alors que les expressions dépendent des individus et des cultures, on distingue généralement un nombre limité d'émotions universellement reconnues. [4]

Aspect universel des émotions Ekman et Friesen [5] ont établi qu'il existe un nombre limité d'expressions reconnues par tous, indépendamment de la culture. Ces expressions innées correspondent aux sept émotions suivantes : la neutralité, la joie, la tristesse, la surprise, la peur, la colère et le dégoût.

1.5 Les contraintes

Bien que des systèmes performants existent [6][7] et que l'amélioration du matériel permette de plus en plus de réalisme, l'animation faciale se heurte à deux problèmes majeurs :

- La complexité du visage humain : Celui-ci se compose de nombreuses couches en interaction, contrôlées par un ensemble de nombreux muscles enchevêtrés.
- La forte sensibilité du cerveau humain à l'expression faciale : Celui-ci interprète inconsciemment certains signes et expressions du visage, parfois difficiles à identifier.

L'animation faciale peut être décomposée en deux grandes familles répondant à des applications et des contraintes radicalement différentes : les animations faciales non temps réel et celles temps réel[8][9].

1.6 L'animation faciale non temps réel

Aujourd'hui, des infographistes sont capables, avec des outils informatiques tels que les logiciels de synthèse d'image (3DS Max, Maya), des moteurs de rendu complexes permettant d'obtenir des images photoréalistes, et les effets spéciaux, de produire des animations faciales sur des modèles de type dessins animés ou sur des humains virtuels très réalistes, comme nous pouvons le voir dans les films d'animation tel que Toy Story.

L'approche la plus simple et la plus utilisée pour animer des modèles 3D consiste à simuler les déformations du visage en manipulant les modèles 3D point par point, image par image. Ce travail, très complexe, nécessite beaucoup de temps et de grandes qualités infographiques pour produire des animations réalistes. C'est cette méthode, apparentée à celle du dessin animé, qui a été longtemps utilisée dans le domaine du cinéma. Cette technique, hormis son inconvénient en terme de ressources nécessaires autant humaines que matérielles, est fort bien adaptée au domaine du cinéma. En effet, le but d'un film n'est pas d'obtenir de l'interactivité sur son déroulement. Une fois les animations produites, celles-ci sont utilisées directement, de façon linéaire. Néanmoins, en manipulant les modèles et les images point par point, cette technique permet énormément de libertés graphiques, laissant toutes les libertés imaginables à l'animateur. Cette technique a été améliorée en utilisant des calculs d'interpolation³, afin de réduire la charge de travail nécessaire pour concevoir les différentes animations. D'autres techniques ont été développées en utilisant un système de paramétrage simple et souvent adapté au cas par cas : le but consiste à faire correspondre à un paramètre une expression ou une déformation faciale.

* Les modèles utilisés dans certains films d'animation possèdent plusieurs centaines de milliers de polygones, plusieurs couches de textures, et des centaines de paramètres permettant d'animer chaque partie du visage indépendamment.

Un travail important, souvent manuel, reste toujours à accomplir : définir comment chaque modèle va se déformer suivant un paramètre donné. La construction de ces informations d'influence peut prendre plusieurs jours à plusieurs semaines de conception suivant la complexité du modèle, le nombre de paramètres de contrôle et la finesse des animations désirées[9].

1.7 L'animation faciale temps réel

N'est plus nécessairement linéaire, mais peut être influencée et/ou modifiée par l'utilisateur ou par des interventions extérieures durant son exécution. L'emploi d'une telle technique d'animation impose d'avantage de contraintes que les animations non temps réel, autant sur la complexité du modèle utilisable, que sur les techniques de déformations employées, la production en nombre de modèle animable, la capacité à reproduire des images photo-réalistes utilisant les jeux de lumières et d'ombres.

Exemple : Les principaux domaines de l'application des visages parlants en temps réel vont des interfaces utilisant des personnages virtuels, que ce soit dans le cadre d'applications stand-alone (i.e. fonctionnant de façon autonome).

Le fait d'animer un humain virtuel en temps réel, de lui faire reproduire des expressions et de parler, implique beaucoup plus de contraintes techniques que les animations non temps réel. Les principales contraintes du temps réel sont :

1. **Rapidité des calculs :** dans le cas d'un film, plusieurs heures peuvent être nécessaires pour faire les calculs des déformations, du rendu d'une image... Dans le cadre d'applications en temps réel, il faut être capable, idéalement, de déformer et d'afficher le modèle et la scène environ 25 fois par seconde.
1. Les ressources matérielles mises en œuvre, autant du point de vue des calculs nécessaires que de la capacité mémoire, influencent la complexité des modèles déformables utilisables. Chaque plate-forme technologique impose des contraintes différentes (puissance de calcul, mémoire, capacité graphique, etc.)

● **les avantages d'animation faciale temps réel :** Le fait de pouvoir animer et produire des animations en temps réel apporte les avantages principaux suivants :

- interaction : possibilité d'intervenir pendant le déroulement des animations.
- les animations peuvent être conçues à la volée suivant l'environnement, les désirs des utilisateurs.
- réduction des informations à transmettre et à stocker : un clone est ainsi capable de reproduire les expressions d'un interlocuteur distant en ne transmettant qu'un nombre restreint de paramètres par rapport à une vidéo (systèmes de paramétrisation)[10].

1.8 Les domaines d'applications

L'animation faciale peut être utilisée dans de nombreux domaines, allant des interfaces basées dialogue pour l'interaction homme-machine, d'un « speaker » virtuel pour les informations, aux loisirs numériques et aux divertissements. Nous présentons ici quelques domaines d'utilisation majeurs de l'animation faciale en trois dimensions. En effet, suivant le type d'application désirée, les contraintes matérielles et humaines impliquent des choix technologiques radicalement différents

- Cinéma et télévision (Le film Tony de Peltrie).
- Jeux vidéo et mondes virtuels.

- Le domaine médical, pour la simulation d'opérations chirurgicales ou des cas cliniques.
- La communication à distance, comme via le Web par exemple.

1.9 Conclusion

Plus généralement, l'animation faciale est un point délicat à cause notamment de la complexité de sa création. Dans ce chapitre, nous avons présenté une description des principaux éléments de visage que l'on souhaite modéliser et représenter ; les caractéristiques faciales les animations faciales en temps réel et non temps réel ect. . . . Cette description montre que la problématique de modélisation et d'animation faciale est un problème complexe, nécessitant l'élaboration et la mise en œuvre de techniques souvent sophistiquées pour faciliter, améliorer, rendre plus accessible et plus rapide le travail du graphiste et/ou de l'animateur.

Dans le chapitre suivant, nous analysons comment ces aspects sont traités par les systèmes existants.

CHAPITRE 2

LES TECHNIQUES D'ANIMATION FACIALE

Sommaire

2.1	Introduction	11
2.2	Les processus de l'animation faciale	11
2.3	Le système de codage	12
2.4	Les techniques d'animation faciale	14
2.5	Le système d'animation faciale idéal	21
2.6	Conclusion	21

2.1 Introduction

La création de l'animation en elle-même continue d'occuper les chercheurs, pour proposer des systèmes toujours plus performants et qui proposent des effets de plus en plus fins et subtils. Certaines techniques qui ont été proposées sont dédiées à l'adaptation morphologique qui permet de changer la morphologie du visage. Nous concentrons en premier lieu, sur la description de processus de l'animation faciale après les différents systèmes de codage pour paramétrer un visage ensuite les diverses techniques d'animation faciale. Ainsi, certaines techniques travaillent sur la géométrie du visage, d'autres travaillent dans l'espace d'image, d'autres n'animent que certaines parties du visage (lèvres, sourcils,...).

2.2 Les processus de l'animation faciale

Le but général de l'animation faciale est d'arriver à reproduire sur un modèle géométrique (2D ou 3D) des expressions pour retranscrire des émotions et/ou des animations correspondant au son et à la parole. Mais chaque technique possède des avantages et des inconvénients suivant le cadre de l'utilisation que nous souhaitons faire de ces animations produites par ordinateur.[11] Le système le plus simple consiste à produire des images-clefs représentant les déformations globales du modèle avec les différentes expressions, déformations pour chaque séquence. Cette technique a l'avantage de laisser une grande liberté aux animateurs, tant sur la conception du modèle lui-même, que sur les animations qu'il souhaite reproduire. De plus, l'interpolation utilisée pour définir les images se trouvant entre les images-clefs nécessite peu de calculs. Les différentes étapes d'un tel système se présentent de la façon suivante :

- Définir un modèle géométrique (maillage, texture, ...).
 - Modifier le modèle géométrique afin de définir les différentes expressions nécessaires durant la séquence d'animation.
- Interpoler au cours du temps.

Cette technique est adaptée à la création de séquences de courte durée. Par contre, elle ne convient pas forcément à la création de séquences en temps réel (en fonction de la complexité des modèles utilisés) permettant l'interactivité avec un modèle virtuel.

Une autre approche consiste non pas à déformer la topologie du modèle directement, mais à utiliser un système de paramètres.

Le fait d'utiliser un système de paramétrisation permet de faire abstraction de la topologie du modèle utilisé et ainsi de concevoir des animations indépendantes de celui-ci. Il devient aussi possible de concevoir des expressions ou des animations de façon interactive en produisant les paramètres par différentes approches comme à base de texte, d'audio ou de séquences vidéo.

Le problème majeur de l'utilisation d'un système de paramétrisation réside dans la complexité de conception du lien entre ces paramètres et la topologie d'un modèle

spécifique. Il existe différentes méthodes de déformation qui permettent de faire ce lien. Certaines sont plus ou moins adaptées à tel ou tel système de paramétrisation[12].

● D'une façon générale, un système d'animation faciale basé sur un système de paramétrisation comporte les étapes suivantes :

1. Définir un modèle géométrique correspondant au sexe du modèle (homme/femme) et au contexte de l'application (modèle réaliste / cartoon ...).
2. Définir une structure d'animation pour ce modèle permettant de faire le lien entre le système de paramétrisation et le modèle géométrique.
3. Définir des expressions en fonction de ces paramètres (phonèmes, expressions).
4. Utiliser ces expressions ou phonèmes, ou une combinaison de ceux-ci comme images-clefs et définir différentes interpolations ou fonctions pour les mélanger.
5. Animer le modèle géométrique en fonction de ces données de paramétrisation en utilisant un moteur de déformation[13].

2.3 Le système de codage

Cette partie est consacrée à différents systèmes de codage d'expressions ou d'animations faciales. Bien que ces systèmes ne soient pas utilisés systématiquement par chacune des méthodes existantes d'animation faciale, il est important de les connaître. Ici, nous ne nous posons pas la question de savoir comment le visage virtuel est déformé, mais nous présentons des méthodes permettant de décrire une expression ou une animation faciale [14].

Le MPEG 4 peut être choisie comme système de paramétrisation pour les raisons suivantes :

- * **Utilisation intuitive** : la conception d'expressions pour un animateur est facilitée par la manipulation des points caractéristiques du visage. Il n'est pas nécessaire d'avoir des connaissances relatives à l'anatomie et aux muscles du visage.
- * **Totale indépendance** : nous avons une totale indépendance entre la partie conception d'animations à partir de diverses méthodes (par interpolation), et la partie correspondante au moteur d'animation faciale utilisé.
- * **Réutilisation des données d'animation** : le déplacement des points de contrôle (FAPs) est décrit dans un système d'unité basé sur les proportions du visage (FAPU1). Par l'intermédiaire de cette normalisation des déplacements, tous les modèles de visage compatibles avec la paramétrisation MPEG 4, peuvent réutiliser les mêmes séquences d'animation, sans aucune modification ou adaptation au modèle.

2.3.1 Le standard MPEG-4

MPEG-4 est une norme adaptée aux déformations à base de points de contrôle. L'utilisation de ce modèle est courante dans les applications distantes puisqu'une faible quantité d'information suffit à définir des déformations sur un visage. Cette

norme est proposée par le groupe MPEG visant à réduire les flux d'information nécessaires dans les fichiers et les communications multimédia.[15] MPEG-4 est utilisable avec des scènes 3D et compte parmi ses fonctionnalités, une méthodologie pour déformer des visages à base de 84 points de contrôle (FDP). Le standard définit entre autres :

- 84 points caractéristiques (Facial Definition Parameters, FDP).
- 6 unités basées sur des mesures anthropométriques du visage (Facial Animation Parameter Units, FAPU).
- 68 paramètres d'animation (Facial Animation Parameters, FAP).

Les FAP : sont les paramètres d'animation qui décrivent comment le visage se déforme, aussi bien pour produire des visèmes que pour générer des mouvements de la tête et des expressions.

Les FDP : Les FDP sont des points caractéristiques du visage. Ils jouent deux rôles principaux :

1. est de permettre la mesure des caractéristiques d'un visage ainsi que le calibrage d'un modèle générique sur un modèle défini par des données obtenues à partir d'un visage réel.
2. est directement lié à l'animation puisque les FAP décrivent les déplacements des FDP.

Les FAPU : Les FAPU sont des unités de normalisation basées sur des distances et des angles entre certains points caractéristiques du visage qui permettent d'interpréter les valeurs des FAP en fonction de la morphologie du visage animé. Les distances, déplacements ou autres paramètres sont alors décrits de façon relative plutôt que de façon absolue[16].

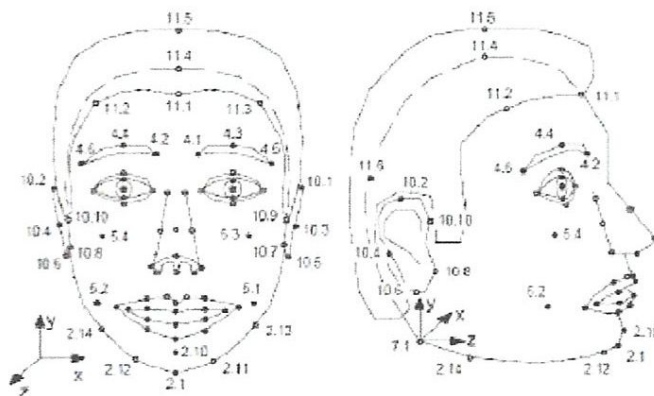


FIGURE 2.1 – Les 84 FDP définis par le standard MPEG-4. Les points pleins sont ceux influencés par les paramètres d'animation FAP. Les points creux sont fixes et sont utilisés pour le calibrage des visages. Les FDP sont groupés par zone (9 pour le nez et 3 pour les yeux par exemple).

2.3.2 Facial Action Coding System (FACS)

Le système FACS (Facial Action Coding System) développé par Ekman et Friesen est un système de description exhaustif des mouvements faciaux. Il s'agit d'associer un code à chaque activation musculaire du visage qui peut être distinguée visuellement. Ces éléments atomiques sont appelés Action Units ou AU. Le manuel du codeur FACS contient ainsi la description visuelle des changements du visage lors de l'occurrence de chaque AU ou chaque combinaison d'AUs.

De plus, chaque AU peut être affichée avec une amplitude différente. Les auteurs ont retenu un maximum de 5 amplitudes pour chaque AU. Une Action Unit ne correspond pas nécessairement à un muscle facial isolé [17].

2.4 Les techniques d'animation faciale

Il existe plusieurs approches en animation faciale. L'approche **par image-clef** correspond aux techniques de base de l'interpolation.

La **paramétrisation** implique plusieurs approches : l'animation basée sur la physique des muscles ou les pseudo-muscles. Dans le domaine des pseudo-muscles, il ya beaucoup d'approches différentes : **animation par région faciale**, **par insertion des muscles** ou par utilisation de points de contrôle.

- Le but de toutes ces techniques reste le même : manipuler la surface virtuelle représentant le visage au cours du temps de façon à produire une séquence animée. Toutes ces techniques impliquent directement ou indirectement la manipulation des points du maillage formant le visage et/ou de points de contrôle liés au visage [18] .

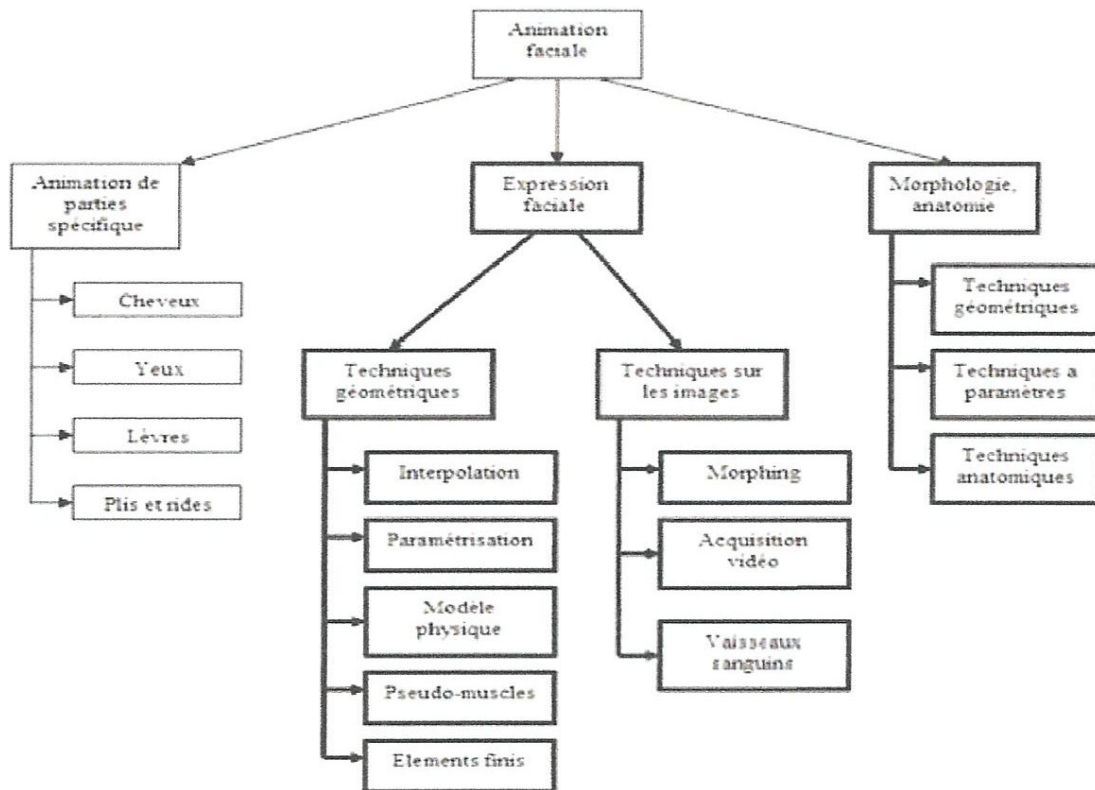


FIGURE 2.2 – Organisation des différentes techniques d’animation faciale

2.4.1 Méthode par image clé

Les positions-clés ont été une des premières approches choisies pour animer le visage, et fonctionnent par interpolation linéaire entre différents maillages. Le principe de cette méthode est le suivant : on dispose d’une base de données de visages (représentant chaque état du visage à un temps donné), obtenue par acquisition 3D ou travail d’artistes. Tous ces visages ont la propriété d’avoir exactement la même topologie (on connaît une correspondance point à point entre deux visages de la base de données). Grâce à cette correspondance entre points, on peut aisément calculer une interpolation (linéaire par exemple) entre deux visages[19].

2.4.2 techniques géométriques

L’interpolation

L’interpolation est certainement la technique de déformation la plus simple à mettre en place. Elle correspond au concept d’images clés utilisées dans les logiciels pour animer des personnages ou autres objets 3D. Il s’agit de modéliser ces images clés (ici des maillages pour les différentes positions extrêmes des visages).

Le choix de la fonction d’interpolation (qui définit les états intermédiaires à partir

des états extrêmes) est très important car c'est de lui que dépendra la qualité et les performances du résultat. Plusieurs fonctions ont déjà été utilisées pour l'animation faciale :

- Interpolation linéaire : La plus simple, mais aussi la plus courante [20].
- Interpolation bilinéaire : Généralement utilisée dans le cadre du "morphing"[21].
- Interpolation sinusoïdale [22].

L'interpolation est très souvent utilisée pour animer des paramètres du visage. Dans cette configuration, l'interpolation ne fonctionne pas seule, mais est un moyen de contrôler plus aisément un modèle de visage sous-jacent. Ce modèle sous-jacent peut être de nature diverse : on peut entre autre, animer les paramètres des muscles d'un modèle musculaire, ou d'une manière générale tout paramètre continu d'un modèle[23].

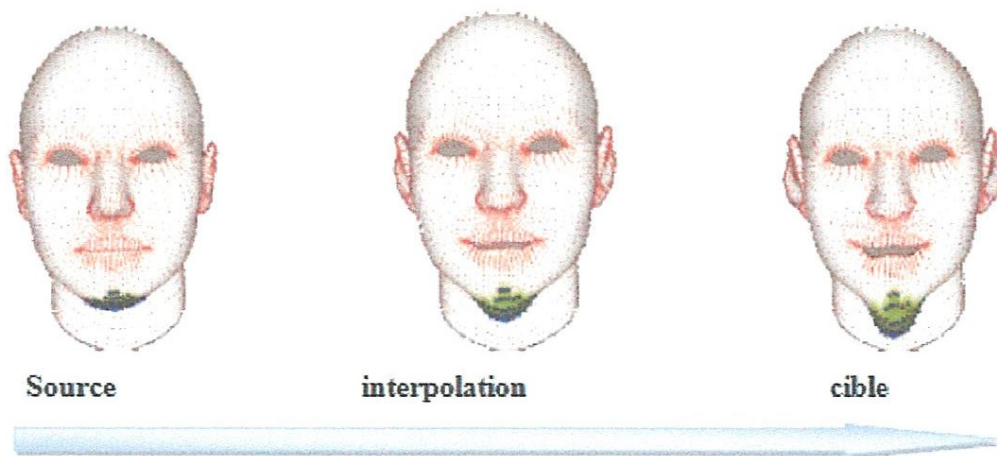


FIGURE 2.3 – Exemple d'interpolation sur un maillage 3D

les avantages et limitations

L'utilisation de positions-clés est [24] :

- * Rapide, surtout si l'interpolation choisie est linéaire
- ** Simple et facilement combinable à d'autres méthodes.

En revanche, la techniques des positions-clés (et d'interpolation en général) présente de nombreux inconvénients :

- Les données de base sont volumineuses et potentiellement fastidieuses à mettre en place.
- Les degrés de liberté des mouvements sont limités par le nombre de positions-clés disponibles ; ainsi les visages ne sont pas assez flexibles pour exprimer une grande gamme d'émotions.
- Les positions-clés obtenues pour un visage sont difficilement adaptables à un autre, et il faut recréer la base de données de visages pour chaque modèle.
- Le résultat obtenu peut devenir incohérent quand on interpole entre plusieurs positions avec des valeurs importantes.

En résumé, l'interpolation est la méthode d'animation faciale la plus ancienne, mais c'est aussi un outil bien maîtrisé et indispensable dans la plupart des techniques d'animation faciale et en particulier dans les paramétrisations.

2.4.3 Animation par paramétrisation

Les techniques d'animation par paramétrisation sont les plus anciennes et permettent non seulement de faire de l'animation mais aussi de la conformation (déformation d'un modèle générique pour qu'il se conforme à la personne désirée). Ici on cherche à découper le visage en régions contrôlées indépendamment par un jeu de paramètres adéquats (par exemple : « déplacer le coin de la lèvre vers le haut »). Les méthodes sont en général purement géométriques et les paramètres sont issus de mesures acquises par capture vidéo[25].

2.4.4 capture de mouvement

Les méthodes de capture de mouvement consistent à recueillir des données d'animation du visage à partir d'une personne réelle pour les utiliser sur un visage synthétique. Travailler sur des données réelles dans l'animation permet de limiter les problèmes liés à la création manuelle et orientée artiste. La fluidité et la dynamique des déformations sont alors mieux respectées visuellement. Les systèmes de capture reposent sur des bases communes qui consistent à détecter et/ou suivre des points de contrôle ou des contours dans les images d'une vidéo, issues d'une ou plusieurs caméras. Une fois ces données obtenues, il faut les corriger si nécessaire, puis les transformer en données utilisables par un système de paramétrisation (par exemple, déterminer les forces appliquées par les muscles d'un visage contrôlé par simulation à partir des mouvements de points de contrôle capturés)[26].

2.4.5 Les méthodes basées pseudo-muscles

Les méthodes basées pseudo-muscles offrent la possibilité de déformer plus facilement un visage qu'en déplaçant les sommets d'un maillage un par un. Le principe est identique dans la plupart des méthodes : une série de maillages clés est créée et des interpolations entre ceux-ci permettent de générer l'animation. Un maillage clé est défini comme étant une modification géométrique d'un maillage initial (la connectivité reste la même). On parle de « pseudo-muscles » car il n'y a pas de modélisation des muscles, contrairement aux méthodes de simulations physiques. On essaie simplement de reproduire visuellement les phénomènes produits par les muscles réels [27][28].

2.4.6 modélisation physique

Pour obtenir des résultats réalistes et physiquement corrects, les paramétrisations visent à simuler les mécanismes de déformations tels qu'ils existent réellement sur le visage humain. Cette partie aurait pu être associée à la paramétrisation par déformation de surface. Mais étant donné les différences notables entre un système visant à produire des déformations visuellement réalistes et un autre visant à modéliser

la biomécanique du visage, il nous a semble plus indiquer de différencier clairement ces deux familles. 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 [29][30][31].

les avantages et limitations : La simulation physique de muscles pour l'animation faciale a le mérite de donner de bons résultats. Mais elles sont en général trop coûteuses (en terme de temps de calcul). On préfère souvent à cette simulation physique des modèles heuristiques : les pseudo-muscles. Les modèles pseudo-musculaires sont plus rapides que la simulation physique de muscles réels et essaient d'en mimer les effets. Toutefois, un pseudo-muscle n'est qu'une approximation d'un muscle réel.

2.4.7 Les déformations de formes libres

Déformations de formes libres Une autre méthode largement répandue pour l'animation d'objets déformables est la technique des déformations de forme libre, plus connue sous le nom de FFD, Free Form Deformations en anglais. Elle a été proposée pour la première fois en infographie par Sederberg et Parry [32] en 1986. Le principe de la technique est simple : l'objet que l'on souhaite déformer est plongé dans la grille de contrôle 3D d'un volume spline tridimensionnel de façon à ce que la position de chaque point de l'objet dans la grille puisse être définie à l'aide d'une paramétrisation unique. Ainsi, pour chaque point de l'objet, ses coordonnées paramétriques sont déterminées au repos et stockées. Ensuite, la grille est déformée, et puisque les points gardent les mêmes coordonnées, l'objet se trouve déformé lui aussi. Cette grille de contrôle définit un volume de Bezier et sa déformation est contrôlée par les points de contrôle traditionnels. L'intérêt d'interpoler les positions des points de l'objet déformé par des fonctions tri-paramétriques est de pouvoir générer des déformations lisses, la continuité de la surface étant préservée. Cette technique permet ainsi des déformations globales de l'objet, la continuité de la surface déformée n'étant cependant pas assurée dans le cas d'une déformation locale (si seulement une partie de l'objet est plongée dans la grille de contrôle). Griessmair et Purgathofer [33] reprennent le même principe mais en utilisant des fonctions de base B-splines tri-paramétriques.

- La force des FFD est qu'elles peuvent être appliquées à différentes représentations graphiques : points, polygones, splines, surfaces paramétriques et surfaces implicites.
- Cependant, le gros inconvénient des premières approches proposées est que la grille de contrôle est parallélépipède pour simplifier la paramétrisation des points.

2.4.8 Déformations dans les sous-espaces du squelette

Le skinning ou Skeleton Subspace Deformation [34] est une technique d'animation pour laquelle les sommets de la surface d'un objet sont déplacés en réponse au mou-

vement d'un squelette interne. Le squelette est composé d'os individuels connectés aux articulations et est recouvert d'un maillage appelé la peau. Les articulations sont bougées, soit à la main par l'animateur, soit par cinématique inverse, pour produire le mouvement, les os entraînant les sommets de la peau. Le plus simple serait de faire subir aux sommets de la peau les mêmes transformations rigides qu'au squelette. Le problème dans ce cas se situe au niveau des articulations, quand un point est influencé par plusieurs os.

- Le skinning résout ce problème en affectant une somme pondérée des transformations des os à la position de chaque sommet, selon l'influence qu'il subit de tel ou tel os. Chaque terme de la somme est calculé en utilisant un repère associé à l'os animé correspondant. La position finale d'un point de la surface déformable d'un objet articulé se trouve donc dans le sous-espace défini par les transformations rigides de ce point par un certain nombre de repères.
- D'autre part, en comparaison avec l'interpolation de formes clefs, cette technique ne permet pas les manipulations directes. En effet, pour chaque point de la surface, il y a un poids par repère du squelette qui l'affecte. Il faut donc éditer le maillage des poids wib pour corriger une déformation manuellement, un travail fastidieux qui ne permet pas toujours d'obtenir le résultat voulu.

2.4.9 Fonction à base radiale

Le principe est de déformer le maillage à partir d'un faible nombre de points de contrôle. La technique définit un élément de déformation géométrique comme étant un point de contrôle, une zone d'influence, un point d'ancrage et une fonction RBF. Lors de la phase d'animation, nous mettons à jour le RBF sur toute la zone d'influence et nous obtenons des déformations lisses sur l'ensemble du maillage.

L'utilisation de RBF comme fonction d'interpolation, garantit une déformation lisse sur les régions. Ce concept se rapproche du système de paramétrisation MPEG 4 qui traite l'information uniquement en terme de points de contrôle, et non pas en terme de déformations musculaires. Les fonctions RBF les plus courantes sont la multiquadratique, la Thin-Plate Spline (TPS), la gaussienne et la CS-RBF. L'utilisation de la RBF gaussienne peut aboutir à de mauvais résultats pour de grands déplacements [35]. L'ensemble des autres fonctions fournissent de bons résultats. La méthode des RBF est similaire au système proposé par [36]. Celui-ci utilise les points de contrôle pour animer le visage, et les RBF permettent une représentation plus compacte et directe des paramètres d'animation.

- Les RBF sont indépendantes de la topologie du maillage.
- La possibilité de concevoir des animations en manipulant uniquement des points de contrôle.
- Les points de contrôle influencent directement les régions proches et fournissent une manipulation intuitive des déformations.

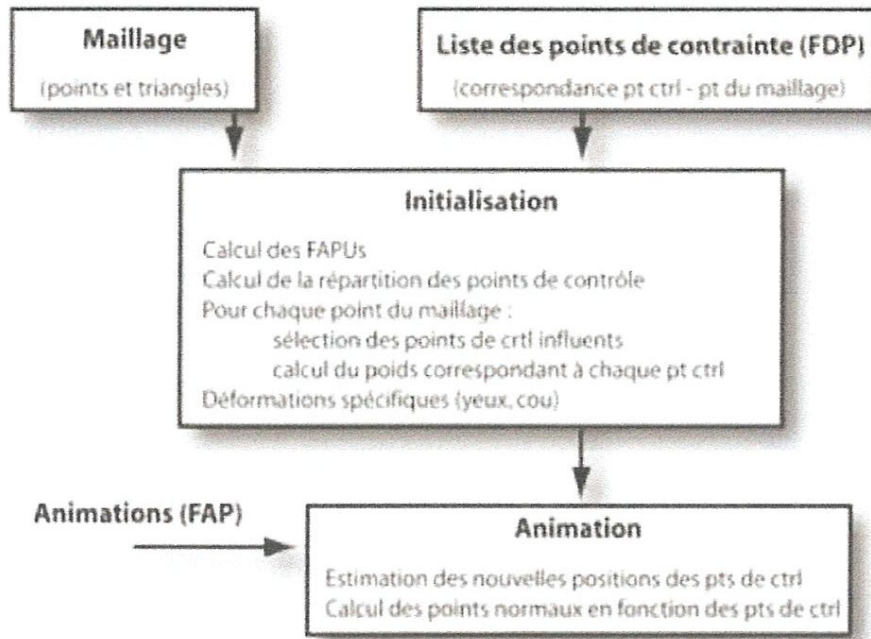


FIGURE 2.4 – vue d'ensemble du processus

L'algorithme de base

Nous décrivons ci-dessous les procédures de base utilisées par Noh pour concevoir un élément déformable :

- 1 Spécifier un point de contrôle sur le visage et sa zone d'influence. Le point sélectionné n'a pas besoin de coïncider avec un point du maillage.
- 2 A partir d'un point sélectionné sur une image 2D, une conversion en un point 3D sur le modèle est faite [37].
- 3 Trouver le point du maillage le plus proche du point 3D. Ce point devient la racine pour la recherche dans les nœuds de l'arbre représentant le maillage.
- 4 Rechercher dans l'arbre des points du maillage, permettant de déterminer l'ensemble des points avec une distance métrique.
- 5 Chaque feuille de l'arbre de recherche devient un point d'ancrage, et avec le point de contrôle spécifique, permet d'initialiser le système RBF associé avec le GDE "geometry deformation elements").

- Le principal inconvénient des réseaux de fonctions de base radiale est d'avoir un temps de calcul qui augmente considérablement en fonction du nombre de points dans le maillage, Une méthode de localisation des déformations permet d'améliorer ce problème.

Les techniques d'analyse/synthèse permettent de capturer les mouvements faciaux afin de les rejouer. Ainsi, le naturel du mouvement est préservé, puisqu'on ne fait

que rejouer une séquence préenregistrée. Cependant, l'animation faciale est difficile à capturer car elle nécessite un nombre important de points de contrôle. Il faut alors trouver un compromis entre précision et le temps de calcul.

2.5 Le système d'animation faciale idéal

Un système d'animation faciale idéal devrait répondre aux contraintes suivantes :

- **Simplicité d'utilisation** : Un système d'animation faciale doit être facile à utiliser et simple à mettre en œuvre.
 - De consacrer un minimum de temps de conception pour rendre le modèle animé, voir entièrement automatique.
 - De permettre une conception interactive des différentes déformations par un animateur.
 - D'obtenir des résultats des déformations réalistes.
 - De pouvoir contrôler précisément certaines déformations.
- **Intégration** : L'utilisation d'animation faciale doit être facile, rapide, fonctionner en temps réel et applicable sur n'importe quel système.
- **Qualité visuelle** : Le but d'atteindre le maximum de ces objectifs pour obtenir un système d'animation facial idéal.

2.6 Conclusion

Dans la littérature, il existe différentes techniques permettant la modélisation et l'animation d'un visage. Nous avons présenté dans ce chapitre un état de l'art contenant les techniques considérées comme majeures dans ce domaine. Les différentes techniques de déformation ont été évoluées et nous avons vu émerger des systèmes de paramétrisation pour gérer les déformations du visage qui permettent d'animer un visage virtuel. Le chapitre qui suit se focalise sur le transfert d'animation faciale.

CHAPITRE 3

LE TRANSFERT D'ANIMATION FACIALE

Sommaire

3.1	Introduction	23
3.2	Le transfert d'animation	23
3.3	Les RBF pour l'interpolation de surfaces	23
3.4	La zone d'influence	26
3.5	Expression cloning	26
3.6	Motion cloning	27
3.7	La comparaison entre les deux méthodes	29
3.8	Conclusion	29

3.1 Introduction

De nombreux axes de recherche ont été développés dans l'animation faciale : des méthodes simples, directement inspirées des techniques du dessin animé, jusqu'au développement de systèmes complexes (tel que l'influence des os et des tissus mous sur la forme u visage, la simulation physique des muscles et leur influence sur la surface de la peau). Dans ce chapitre, nous détaillerons les différentes méthodes de transfert d'animation faciale à partir d'un visage réel vers un visage virtuel. Ces méthodes permettant de construire des déformations d'un modèle soit par RBF ou avec les deux techniques concernant le clonage d'expression.

3.2 Le transfert d'animation

Le transfert automatique des mouvements de visage depuis un modèle existant source à un modèle cible peut enregistrer significativement des spécifications et animations soigneuse d'un modèle réel à un modèle du visage. Il existe plusieurs méthodes de transfert d'animation dans la littérature, parmi les quelles nous citons au-dessous le clonage et le transfert avec RBF.

3.3 Les RBF pour l'interpolation de surfaces

Les RBF sont largement utilisées en informatique graphique pour l'approximation ou l'interpolation d'images et de surfaces. Arad et al. Présentent leurs applications aux images 2D pour modifier les expressions d'un personnage d'une image en la déformant [38].

En pratique, il s'agit de calculer une transformation de l'espace caractérisée par les déplacements d'un nombre limité de points que nous appellerons points de contrôle (PdC). Cette transformation peut ensuite être appliquée à chaque point de l'espace et fournit les nouvelles coordonnées de ce point en adéquation avec les déplacements des PdC. Il s'agit en fait d'une interpolation entre l'espace des PdC dans leur position initiale et dans leur position modifiée.

Le calcul des RBF s'effectue en deux étapes. La première consiste à trouver les pondérations associées à chaque point de contrôle (une pour chaque axe), seules les positions des PdC avant et après déplacement sont utilisées. Ensuite la transformation est appliquée à tous les autres points de la surface pour obtenir leurs coordonnées après l'interpolation [39].

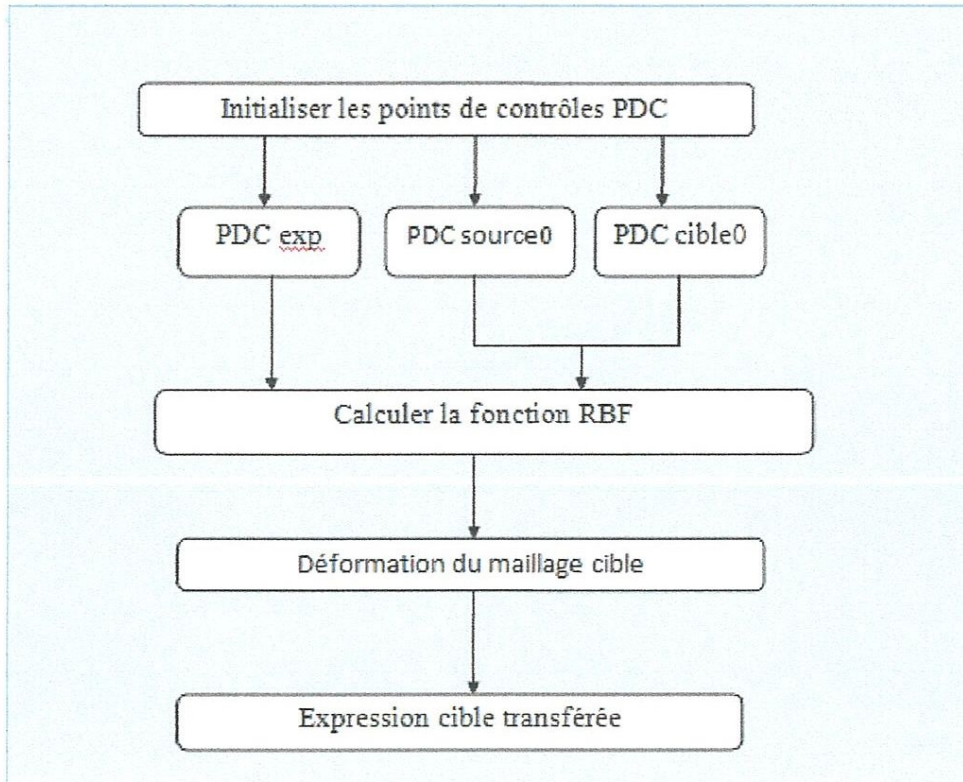


FIGURE 3.1 – Les étapes de l'interpolation par RBF.

Les points de contrôles Une fois les coordonnées 2D des points de contrôle sources obtenues, nous calculons une nouvelle interpolation par RBF entre l'espace défini par les points repères du visage source et l'espace défini par les points repères du visage cible. En appliquant cette transformation aux points de contrôle sources, nous obtenons les points de contrôle cibles en 2D.

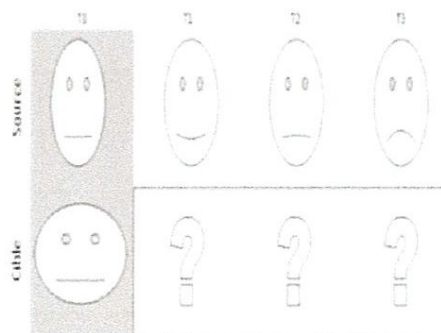


FIGURE 3.2 – L'entrée : un maillage source et ses expressions sous forme de maillages, un maillage cible.

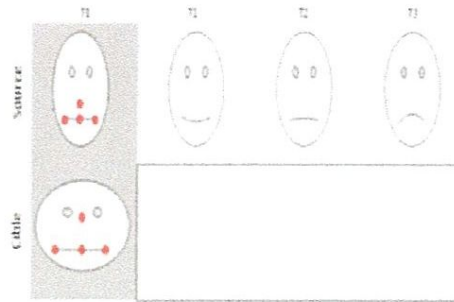


FIGURE 3.3 – Initialisation des points de contrôle source et Cible.

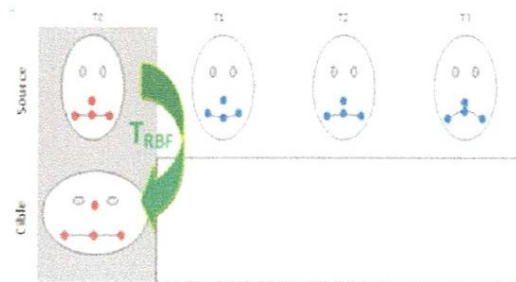


FIGURE 3.4 – de la transformation source vers cible.

- La correspondance entre deux maillages nous permet de calculer la transformation nécessaire c'est-à-dire l'interpolation à base de RBF qui déforme le maillage cible « comme » le maillage source. Le résultat de cette interpolation correspond à un maillage possédant la topologie du maillage source et une géométrie se rapprochant de celle du maillage cible. Il nous reste maintenant à déformer l'intégralité du maillage cible pour lui donner l'expression désirée[40].

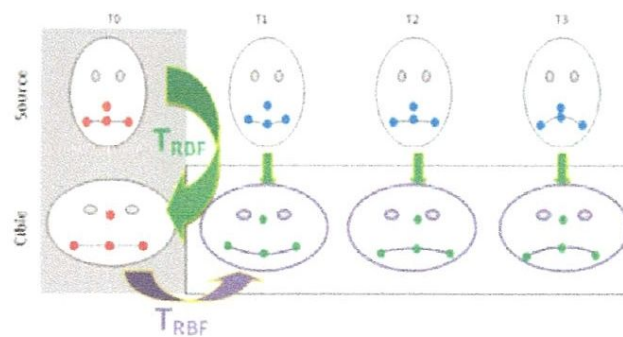


FIGURE 3.5 – du maillage cible selon les points de contrôle obtenus pour chaque expression.

3.4 La zone d'influence

Lorsque l'on souhaite définir des régions d'influence, il est nécessaire d'utiliser une métrique. Il existe plusieurs approches qui prennent en compte la diversité de la topologie d'un maillage. Par exemple, la simple distance euclidienne entre les points permet de définir une région contenue dans une sphère de rayon r . Nous pouvons étendre cette approche à différentes formes géométriques comme les ellipses ou les rectangles. Pour définir l'influence de certains points caractéristiques. L'ensemble des points de contrôle influençant un point ordinaire du maillage, et les facteurs correspondant entre le point ordinaire et ses points de contrôle, sont ainsi estimés à partir des informations suivantes :

1. La prise en compte des distances entre les points de contrôle.
 2. Les distances entre le point ordinaire du maillage et les points de contrôle les plus proches.
 3. L'influence relative des points de contrôle autour d'un point donné.
- dans cette partie nous avons présenté les principales étapes nécessaire a notre travail on traitant l'aspect mathématique de déroulement de l'interpolation par RDF.

3.5 Expression cloning

L'approche proposée par Noh et al[41]. consiste à produire des animations faciales par réutilisation de données. Des animations faciales de haute qualité sont créés à partir de n'importe quelle méthode (à la main, utilisation de modèle musculaire, etc), l'expression cloning (EC) réutilise les vecteurs de déplacement des points du maillage du modèle source pour appliquer les animations similaires sur un nouveau modèle . Si les animations du modèle source sont expressives et de bonne qualité, les animations du modèle cible devraient aussi disposer des mêmes qualités.

- **La première étape** consiste à déterminer la correspondance entre les points de la surface des modèles source et cible. Aucune contrainte n'est imposée quant au nombre de points du maillage et la topologie des modèles utilisés. A partir d'un ensemble de paramètres sélectionnés à la main sur les 2 modèles, Noh calcule une correspondance pour tous les points des maillages. La correspondance initiale nécessite la sélection manuelle d'une dizaine de points afin de calibrer les deux modèles, par la suite une recherche automatique (règles heuristiques) des autres points caractéristiques est appliquée. Sans recherche automatique, les expérimentations ont montré qu'une vingtaine de points étaient nécessaires pour obtenir une bonne correspondance.
- **La seconde étape** consiste à transférer les vecteurs de déplacement du modèle source vers le modèle cible. L'amplitude et la direction des déformations sont ajustées afin de correspondre à la surface du modèle cible. Utilisant le calcul des correspondances de la première étape, le transfert de déformations est effectué par interpolations linéaires utilisant les coordonnées barycentriques.

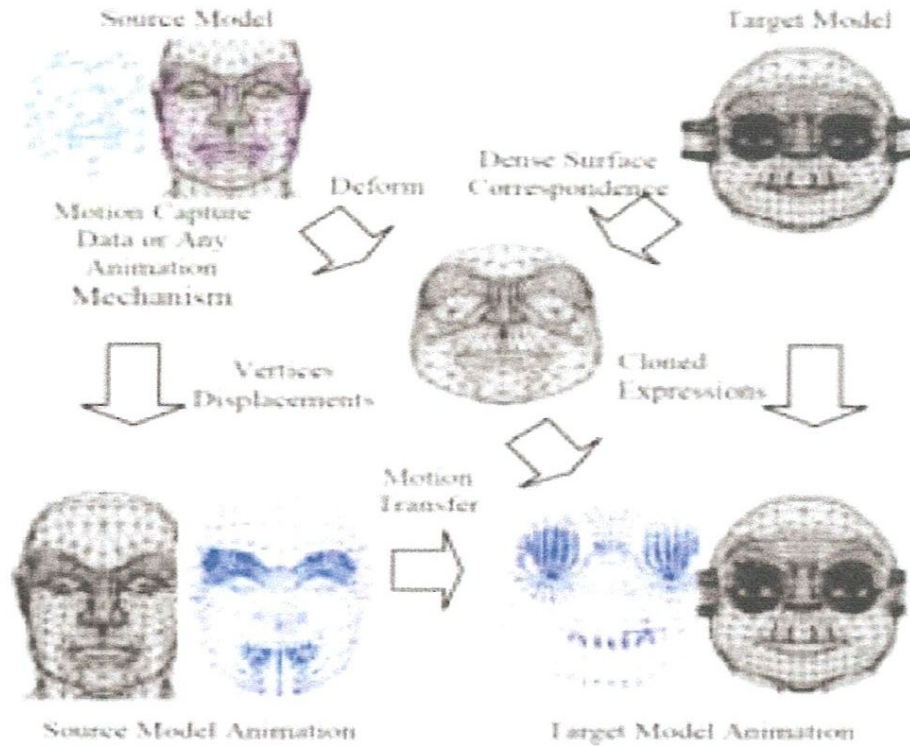


FIGURE 3.6 – Système d'expression cloning de Noh et al[41]

3.6 Motion cloning

Le processus de clonage proposé par Pandzic [42]. Les entrées du processus sont les modèles source et cible. Le modèle source est fourni au système en position neutre ainsi que dans des positions contenant différentes déformations que l'on souhaite copier. Le modèle cible existe uniquement en position neutre. Le but est d'obtenir pour le modèle cible, les déformations équivalentes du modèle source. Le processus se déroule de la façon suivante :

- **Normalisation du modèle** : Le but de la normalisation est de transformer le modèle dans un nouvel espace 3D, dans lequel tous les modèles disposent des mêmes proportions. Ce transfert a pour conséquence de normaliser les déplacements pour l'ensemble des modèles. Pandzic utilise la paramétrisation MPEG 4 (FAPU et FDP) comme paramétrisation et normalisation de ses modèles. Il est nécessaire de définir préalablement les points caractéristiques sur chaque modèle, et utiliser ces informations pour normaliser les modèles.
- **Calcul des déplacements faciaux** : Les déplacements faciaux sont définis comme la différence de position des points du maillage entre l'état neutre et une expression. Ceux-ci sont exprimés par un tableau de vecteurs où chaque vecteur correspond à un point du maillage du modèle. Ces vecteurs sont calculés dans l'espace normalisé (FAPU).
- **Alignement des modèles source et cible** : Pandzic utilise une correspon-

- dance spatiale basée sur les points caractéristiques des modèles pour transférer les régions de déformations d'un modèle source sur un modèle cible.
- **Copie des déplacements** : Cette étape consiste simplement à transférer les régions d'un modèle sur l'autre en fonction de la normalisation, des régions définies précédemment et de l'alignement des modèles.
 - **Corrections par anti-aliasing** : Les régions étant différentes entre le modèle source et cible, certains points du maillage n'ont pas les mêmes déplacements sur le modèle cible. Un processus de lissage des déformations sur le modèle cible permet de réduire les cassures dans les régions clonées.
 - **des Traitements spéciaux** Les lèvres disposent de régions de déformation difficiles à cloner. Ceci est dû au fait que les lèvres sont jointes dans la position neutre. Pandzic applique un traitement spécifique par projection des régions dans un espace 2D. Les autres parties du modèle comme les yeux, les dents ou la langue, sont traitées par déplacement ou rotation des parties spécifiques de la structure du modèle.

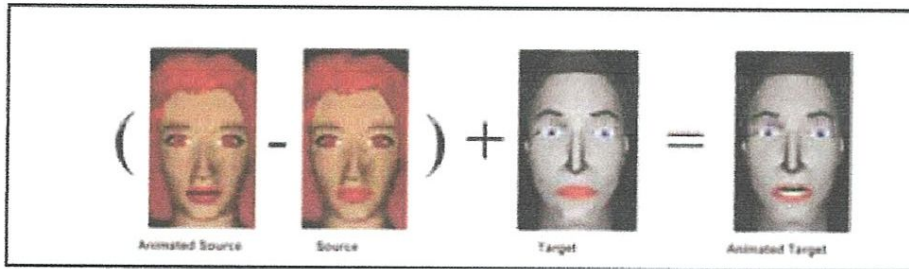


FIGURE 3.7 – Vue d'ensemble du processus de clonage d'animations[42]

L'yun et al [43] présentent une approche pour le clonage d'expression à partir d'un modèle source pour un modèle cible en préservant les caractéristiques du modèle cible. Leur approche consiste en 3 étapes :

1. La construction des modèles clefs.
2. La paramétrisation.
3. Le mélange des expressions.

Pour la construction des modèles clefs, ils utilisent une décomposition en modèles clefs verbaux et émotionnels. Les modèles verbaux correspondent aux expressions produites par l'animation concernant la parole, alors que les modèles émotionnels sont utilisés pour reproduire des expressions. Ainsi, à partir d'une simple paramétrisation ils peuvent définir le modèle cible dans l'espace des paramètres. Pour définir les fonctions de poids pour la paramétrisation du modèle cible, ils utilisent une interpolation des données multi-dimensionnelles. Durant l'animation, le modèle cible est généré par le mélange des modèles cibles clefs en utilisant les poids provenant du modèle source.

- Premièrement, elle nécessite la préparation d'un grand nombre de modèles clefs source et cible par un animateur.
- Une autre limitation provient du fait que cette méthode ne peut pas correctement copier une expression quand elle est trop loin de la base de modèles clefs source

3.7 La comparaison entre les deux méthodes

Les deux approches présentées ci-dessus sont basées sur le même concept. Bien que la comparaison soit difficile, car les modèles sont différents, Pandzic [42] propose de comparer ces deux approches de transfert :

- La méthode Motion Cloning préserve la compatibilité des animations MPEG 4 contrairement à l'Expression Cloning (due au fait du manque de normalisation).
- Le Motion Cloning traite les déplacements liés aux yeux, les dents et la langue.
- L'Expression Cloning utilise des fonctions radiales pour aligner les points en 3D. Le Motion Cloning aligne les points en 2D par simple interpolation.
- L'Expression Cloning ajuste les vecteurs de déplacement localement et corrige leur direction. Le Motion Cloning utilise la normalisation des FAPUs.
- L'Expression Cloning propose des règles heuristiques pour identifier la correspondance entre deux modèles. Motion Cloning peut utiliser ces règles mais nécessite au moins la définition des points caractéristiques MPEG 4 (FDP). D'autres points de correspondance peuvent être ajoutés pour affiner la correspondance.

les deux méthodes présentées ci-dessus utilisent le même concept de clonage, et fournissent des résultats proches. Elles semblent équivalentes sauf dans le cadre d'une utilisation de la paramétrisation MPEG 4 favorisant les Motion Cloning par conservation de la compatibilité avec la normalisation.

3.8 Conclusion

le transfert est un axe important dans l'animation faciale, différentes méthodes existent déjà, nous les avons brièvement présenté dans ce chapitre.

CHAPITRE 4

CONCEPTION ET IMPLÉMENTATION

Sommaire

4.1	Introduction	31
4.2	Objectif	31
4.3	Architecture générale de transfert	31
4.4	Implémentation	35
4.5	Résultats et discussion	44
4.6	Conclusion	46

4.1 Introduction

L'animation faciale est l'un des points clés dans le réalisme des scènes 3D qui mettent en scène des personnages virtuels. Actuellement, pour générer une animation faciale en 3D, deux questions principales doivent être posées :

- Comment animer le visage ?
- Par quel outil nous pouvons transférer ce visage animé ?

Pour répondre à ces questions, il est nécessaire de connaître le contexte d'utilisation de l'animation faciale : en temps réel ou hors-ligne ; ressources disponibles (temps, ressources matérielles, ressources humaines, ressources financières) etc. Pour répondre au mieux à ces besoins, il est également important de prendre en compte les nombreuses difficultés qu'implique la création d'une animation faciale réaliste.

Nous nous intéressons dans ce mémoire au transfert d'expressions faciales d'un visage réel vers un visage virtuel. Dans ce chapitre nous détaillons les différentes étapes de la conception de notre travail ainsi que l'implémentation de notre application.

4.2 Objectif

Le principal objectif de notre travail est de transférer des animations faciales à partir d'une vidéo réelle à un visage virtuel :

Nos données d'entrées sont :

- une vidéo d'un visage source avec différentes expressions, qui débute par une expression neutre.
- un modèle de visage tridimensionnel en état neutre.

Comme sortie, nous obtenons une vidéo d'un visage cible (3D) avec les expressions du visage source transférées.

4.3 Architecture générale de transfert

Notre travail passe par les étapes suivantes :

- Chargement de la vidéo source.
- Fragmentation de la vidéo.
- Détection de visage et des points de contrôle dans les frames.
- Détection des points de contrôle dans le visage 3D.
- Calcul de la fonction à base radiale (calcul des poids)
- Déformation du maillage cible par application de RBF calculée.

Le schéma suivant illustre les étapes de notre conception :

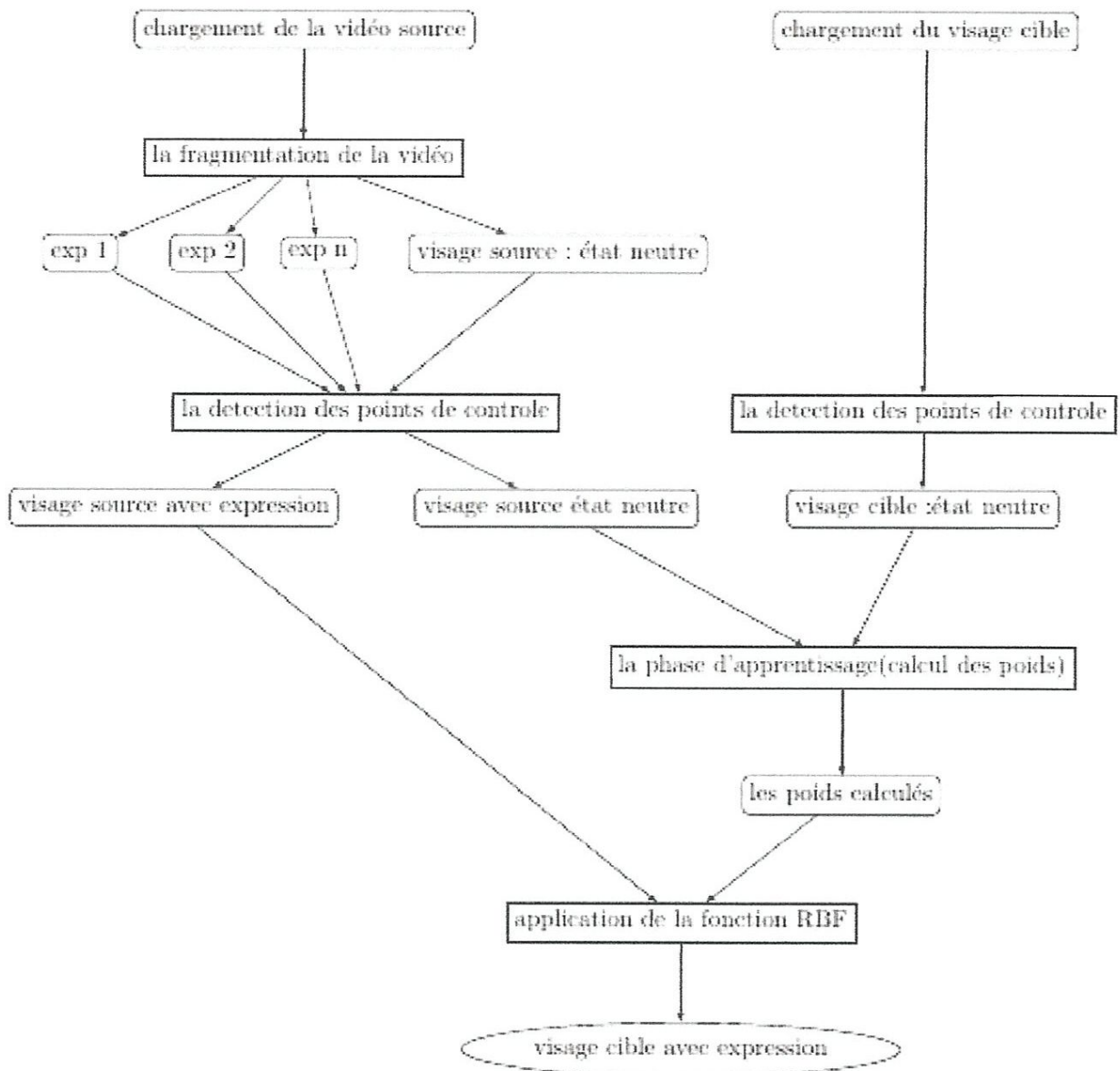


FIGURE 4.1 – l'architecture générale de notre application

4.3.1 Chargement des entrées

- * chargement de la vidéo source en segmentant cette vidéo en plusieurs frames .
- * chargement du visage cible :pour générer un visage 3D,on doit respecter certaines contraintes pour que le visage puisse être animé correctement comme les ouvertures au niveau des yeux et de la bouche par exemple.Il faut associer au maillage 3D des textures pour rendre le visage plus réaliste.

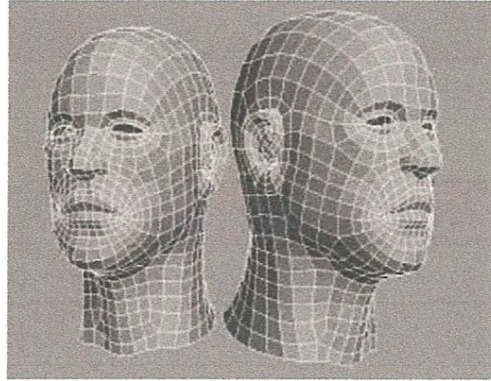


FIGURE 4.2 – définition du maillage 3D

4.3.2 Détection des points de contrôle

Le processus d'initialisation des points de contrôles s'applique sur le visage source et le visage cible dans l'état neutre, ainsi que sur chaque expression du visage source

Luxand FaceSDK détecte 66 points caractéristiques du visage. Ces points caractéristiques du visage peuvent être consultés par leurs noms dans le tableau FSDK Features. Nous avons choisis seulement 44 points de contrôle pour toutes les frames source fragmentés .

analyse

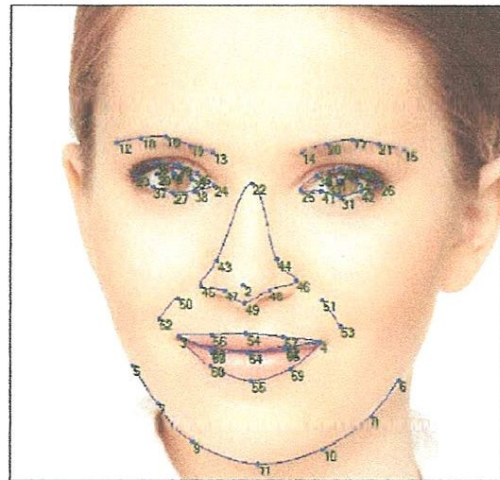


FIGURE 4.3 – les points de contrôles selon FaceSDK

Pour la Détection des points de contrôle sur le visage cible à l'état neutre nous déterminons manuellement sur le maillage 3D.

4.3.3 calcul de la fonction à base radiale

le calcul RBF s'effectue en deux phases :

phase d'apprentissage (calcul des poids) le poids qu'on doit calculer ,permet de définir une approximation entre un point de l'espace du visage source dans l'espace du visage cible.Chaque visage est défini par les points caractéristiques liés à l'animation faciale.

Ce processus représente la phase d'initialisation de l'interpolation par RBF calculée sur les points de de contrôle des deux visages dans la position neutre.Cette étape n'est réalisé qu'une seule fois pour les deux visages.

Nous calculons trois RBF, une pour chaque axe \vec{x} , \vec{y} et \vec{z} .

Soit \vec{S} l'ensemble des sommets du maillage \vec{s}_i et \vec{t}_i correspondent respectivement aux PdC numéros i avant et après déplacement. $\vec{a} \in \vec{S}$ Correspond à un sommet quelconque du maillage.

- L'interpolation va s'effectuer en plusieurs étapes. Tout d'abord, une phase d'apprentissage est effectuée, elle est basée sur le système d'équations :

$$F(\vec{s}_j) = \vec{t}_j = \sum_{i=1}^N .wi.h(\|\vec{s}_j - \vec{s}_i\|) \quad (4.1)$$

avec

$$h.(\|\vec{s}_j - \vec{s}_i\|) = \sqrt{\|\vec{s}_j - \vec{s}_i\|^2 + scj^2} \quad (4.2)$$

$$scj = \text{mini} \neq j(\|\vec{s}_j - \vec{s}_i\|) \quad (4.3)$$

Cette dernière représentant la distance minimale entre \vec{s}_j et les autres PdC. N correspond au nombre de PdC et wi les poids à calculer.Soit H la matrice telle que :

$$H_{i,j} = h.(\|\vec{s}_j - \vec{s}_i\|) \quad (4.4)$$

trouver les poids wi correspond à résoudre l'équation :

$$\vec{W} = H^{-1} . \vec{t} \quad (4.5)$$

avec :

$$\vec{t} \in \{\vec{tx}, \vec{ty}, \vec{tz}\}$$

et

$$\vec{tx} = \{tx_0, tx_1, \dots tx_n\}$$

txi représentant la coordonnée sur l'axe \vec{x} du PdC \vec{t}_i , après déplacement. En fait, la seule différence de calcul entre les trois axes se situe au niveau de cette multiplication matricielle qu'il faut donc effectuer trois fois. Un terme de régularisation est ajouté, il sert à diminuer l'importance du bruit éventuel dans les calculs. L'équation à résoudre devient donc :

$$\vec{W} = (H + \lambda I)^{-1} \cdot \vec{t} \quad (4.6)$$

- Par expérimentation et pour les maillages que nous avons utilisés pour nos tests, une valeur $\lambda = 0.001$:
- offre de bons résultats. Pour effectuer le calcul matriciel, il faut inverser la matrice $H + I \cdot \lambda$.

pour cela nous avons utilisé la décomposition LU. Une fois les poids w_i trouvés, il est possible de connaître les nouvelles coordonnées interpolées de chaque point \vec{a} . Pour cela, il suffit de calculer l'équation :

$$F(\vec{a}) = \sum_{i=1}^N w_i \cdot h(\|\vec{a} - \vec{t}_i\|) \quad (4.7)$$

phase de transfert A l'issue de l'étape précédente, on a deux poids calculés en ignorant la troisième dimension qui nous permet la transformation nécessaire c'est-à-dire l'interpolation à base de RBF qui déforme le maillage cible comme le maillage source.

4.3.4 Déformation du maillage cible

Il nous reste maintenant à déformer l'intégralité du maillage cible avec les nouveaux points de contrôle pour lui donner l'expression désirée.

4.4 Implémentation

plus particulièrement en informatique, l'implémentation désigne la création d'une application à partir d'une conception elle doit répondre à des objectifs fixés préalablement.

4.4.1 Ressources, environnement de développement et logiciels

* Ressources matérielles

L'environnement de base qui a constitué le support de notre travail est le système d'exploitation Windows 7 qui offre les fonctionnalités pour obtenir d'excellentes

performances de façon aisée sur un ordinateur portable dont les caractéristiques sont décrites dans le tableaux suivant :

Micro ordinateur portable	Caractéristiques
Processeur	Intel(R)Core(TM)i5-2410M CPU@2.3GHZ 2.3GHZ
Capacité de disque dur	298 Go
Capacité du RAM	4.00 Go

TABLE 4.1 – matériel utilisé

* Environnement logiciel

java

Java est un langage de programmation orienté objet et un environnement d'exécution, développé par Sun Microsystems. Il fut présenté officiellement en 1995. Le Java était à la base un langage pour Internet, pour pouvoir rendre plus dynamiques les pages (tout comme le JavaScript aujourd'hui). Mais le Java a beaucoup évolué et est devenu un langage de programmation très puissant permettant de presque tout faire, je dis bien presque car nous verrons pourquoi il ne permet pas de tout faire. Java est aujourd'hui officiellement supporté par Sun, mais certaines entreprises comme IBM font beaucoup pour Java[44]

JAVA 3D est un excellent outil pour manipuler facilement des modèles géométriques 3D[45]

pourquoi java ?

- Portabilité excellente.
- JDK très riche.
- Applications plus sûres et stables nombreuses implémentations, JVM et compilateurs, libres ou non IDE de très bonne qualité et libres : Eclipse et Netbeans par exemple.
- langage orienté objet et aussi puissant.
- Nombreuses librairies tierces sollicite très grande productivité.

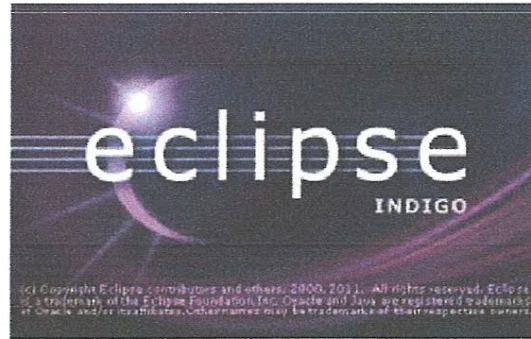


FIGURE 4.4 – l'environnement Eclipse

Eclipse

Eclipse est un environnement de développement intégré libre extensible, universel et polyvalent, permettant de créer des projets de développement mettant en œuvre n'importe quel langage de programmation. Eclipse IDE est principalement écrit en Java (à l'aide de la bibliothèque graphique SWT, d'IBM), et ce langage, grâce à des bibliothèques spécifiques, est également utilisé pour écrire des extensions.

FaceGen Modeller 3.1 :

Est un logiciel de modélisation de visage 3D pour Windows. Il permet de créer facilement des visages 3D réalistes à partir d'une ou plusieurs images ou de manière aléatoire, de modifier les visages dans plus de 150 façons, y compris la race, l'âge et le sexe et d'exporter les modèles de visages vers la plupart des formats de fichiers 3D joints à des animations de morphing.[46]

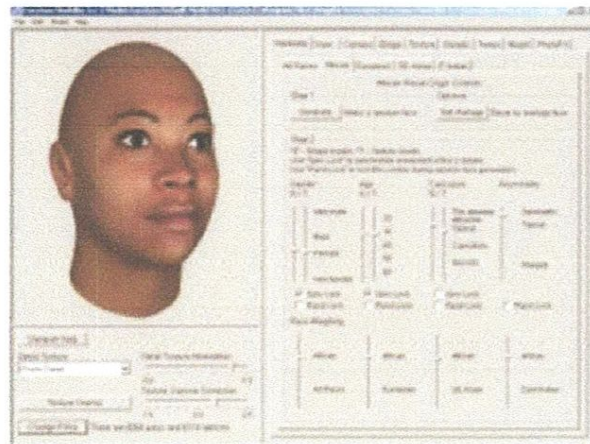


FIGURE 4.5 – interface FaceGen

La librairie OpenCV : «Open Source Computer Vision Library » C'est une bibliothèque qui offre des fonctions de traitement d'image par ordinateur. Toutes

ses fonctionnalités ont été programmées en langage C/C++ puis exportées vers une multitude de langages comme Python et JAVA. Afin d'utiliser les fonctions de cette bibliothèque, nous avons eu recours à son interface sous java dite «JavaCV» représentant une couche d'interfaçage entre les programmes JAVA et la librairie OpenCV.

le format ".obj" : Défini par la société WaveFront Technologies, un fichier OBJ contient la description d'une géométrie 3D. Les formes géométriques peuvent être définies par des polygones ou des surfaces lisses telles que des surfaces rationnelles et non rationnelles. Ces surfaces polygonales sont décrites par un ensemble de sommets (accompagné de coordonnées de texture et de normales en chaque sommet) et d'un ensemble de faces. Le format OBJ sera utilisé dans des cas particuliers, plus évolués mais aussi plus rares (en particulier pour des imprimantes 3D qui gèrent plusieurs couleurs).

le format ".off " Les fichiers de maillage que nous allons utiliser sont au format .off et ils se présentent de la façon suivante :

```
OFF 8 12 0
0.256 0.365 0.569
2.365 2.654 8.145
...
...
3 0 1 2
3 2 5 9 3 6 0 4 ...
...
o'u :
```

- Tout d'abord, il s'agit de fichiers ASCII.
- Les trois premiers caractères rappellent l'extension du fichier. Ici : OFF
- Le premier entier est le nombre de sommets du maillage. Ici : 8
- Le deuxième entier est le nombre de facettes du maillage. Ici : 12
- Le troisième entier est le nombre d'arêtes et bien souvent cette valeur est à 0 : elle n'a pas été mise à jour lors de la création du fichier.

• Ensuite vient la liste des coordonnées des sommets du maillage sous forme de flottants. Les trois premiers flottants sont les coordonnées x, y, z du premier sommet (numéro 0), les trois suivants sont les coordonnées x, y, z du deuxième sommet (numéro 1) et ainsi de suite.

• Une fois tous les sommets stockés, vient la liste des facettes sous la forme suivante : La première valeur est le nombre de sommets de la facette (pour nous, il n'y aura que des facettes triangulaires, donc cette valeur sera 3). Cette valeur est suivie du numéro des sommets composant la facette. Ici, la première facette triangulaire est composée du sommet numéro 0, du sommet 1 et du sommet 2. La deuxième facette triangulaire est composée des sommets numéro 2, 5 et 9, et ainsi de suite.

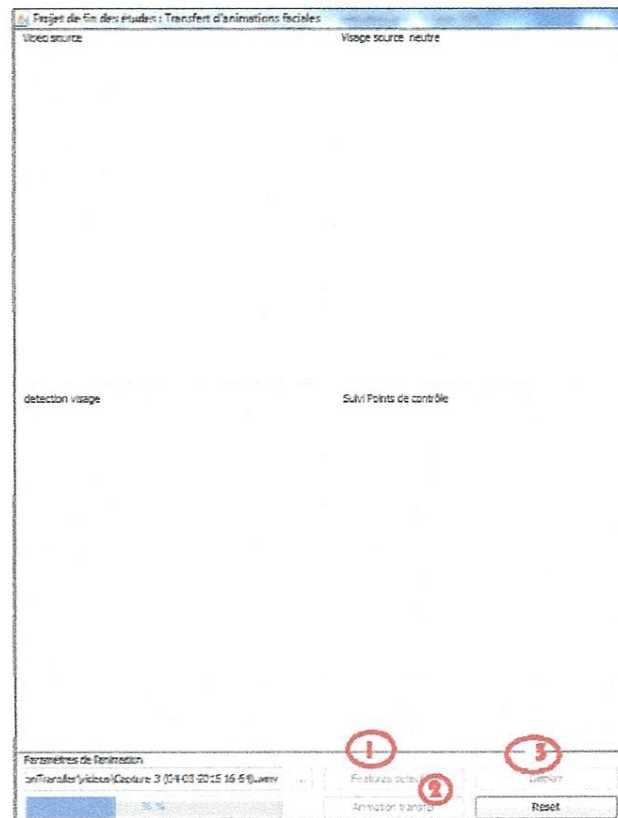


FIGURE 4.7 – l'attente de chargement de vidéo à partir d'une barre de progression

- 1 le premier bouton : l'extraction des points de contrôle pour chaque frame
- 2 le deuxième bouton : le transfert d'animation avec une barre de progression
- 3 le troisième bouton : lorsqu'on remarque le chargement de la barre est complet en peut afficher les résultats de transfert.

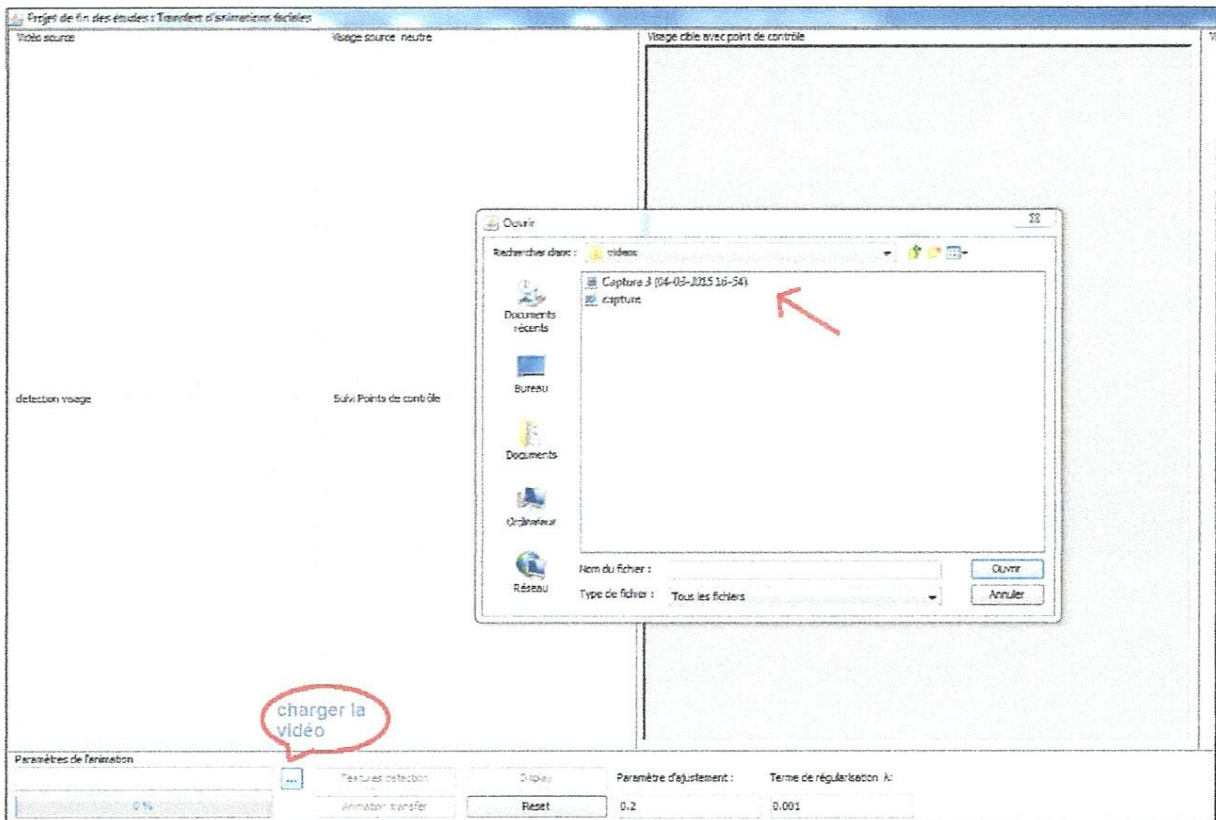


FIGURE 4.8 – chargement de la vidéo préenregistré

la Détection de visage : L'étape de détection permet de décider exclusivement sur l'existence ou non d'un visage dans une image. Nous avons donc besoin d'un détecteur (classifieur) de visages permettant de chercher les caractéristiques relatives à ce dernier. La bibliothèque OpenCV implémente la méthode de détection de visages et offre comme résultat une liste de fichiers XML dits classifieurs en cascade de haar. Nous avons donc pu éviter la phase d'apprentissage d'un détecteur en exploitant le classifieur évalué dans le fichier `haarcascade_frontal_face_alt2.xml` :

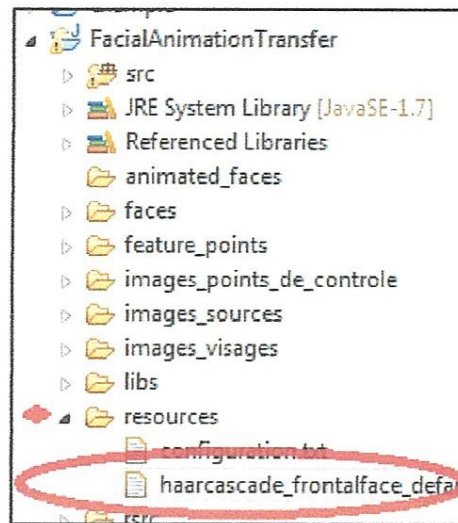


FIGURE 4.9 – classifieurs en cascade de haar

- destiné à détecter un visage en particulier lorsqu'il est présent en vue frontale. Nous avons aussi utilisé les fonctions de cette même bibliothèque pour détecter le visage dans chaque frame de la séquence vidéo. Les coordonnées du visage détecté ont été passées à une fonction chargée d'englober ce dernier dans un rectangle tout en traitant les séquences vidéo image par image.



FIGURE 4.10 – détection de visage

Nous affichons l'état neutre de l'image source avec les points de contrôles. Nous proposons d'utiliser des points de contrôle choisis selon FaceSDK.

les points caractéristiques du visage	les valeurs
-la bouche : 14 points	- 3, 4, 54, 55, 56, 58, 59, 57, 60, 61, 62, 63, 64, 65
les joues : 4 points	- 50, 51, 52, 53
- oeil gauche : 8 points	- 25, 26, 31, 32, 39, 40, 41, 42,
- oeil droite : 8	- 23, 24, 27, 28, 35, 36, 37, 38
- sourcil gauche : 5 points	12, 13, 17, 18, 19
- sourcil droite : 5 points	14, 15, 17, 20, 21

TABLE 4.2 – le choix des points de contrôle

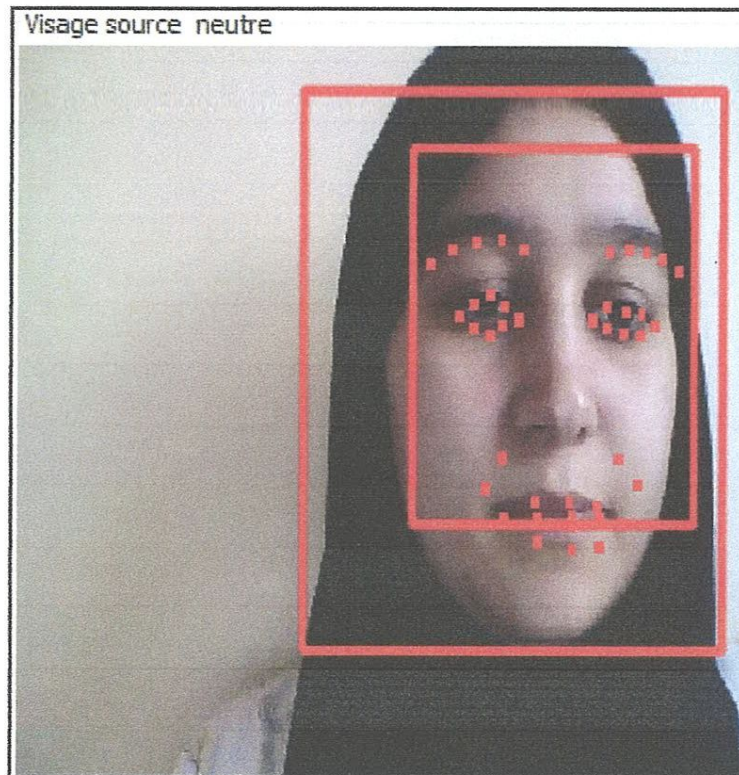


FIGURE 4.11 – visage source neutre

la figure suivante (FIGURE 4.9) montre l'entrée :visage source avec expression

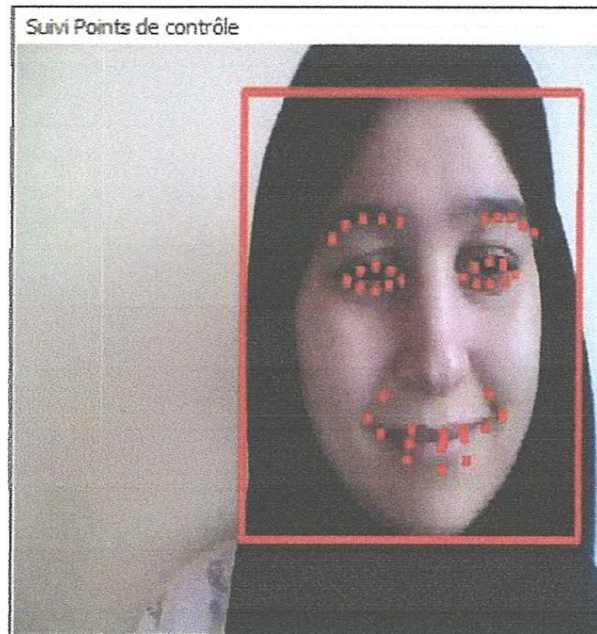


FIGURE 4.12 – visage source avec expression

- Pour créer une déformation clé de maillage cible, il suffit de déplacer ces points selon l'expression désirée puis de calculer l'interpolation par RBF entre l'ensemble des PdC initiaux et les PdC déplacés.

4.5 Résultats et discussion

Nous avons testé sur notre application la méthode de transfert qui peut être utilisée au sein d'un système de capture de mouvement, pour adapter les données capturées depuis un visage réel sur le visage virtuel.

La fonction à base radiale permet d'appliquer les vecteurs d'animation d'un maillage de visage à un autre. Ces vecteurs correspondent aux déplacements de chaque sommet du maillage initial pour chaque position clé de l'animation. Ensuite à partir de la correspondance des points de contrôle de la source et de la cible en déduisant le transfert.



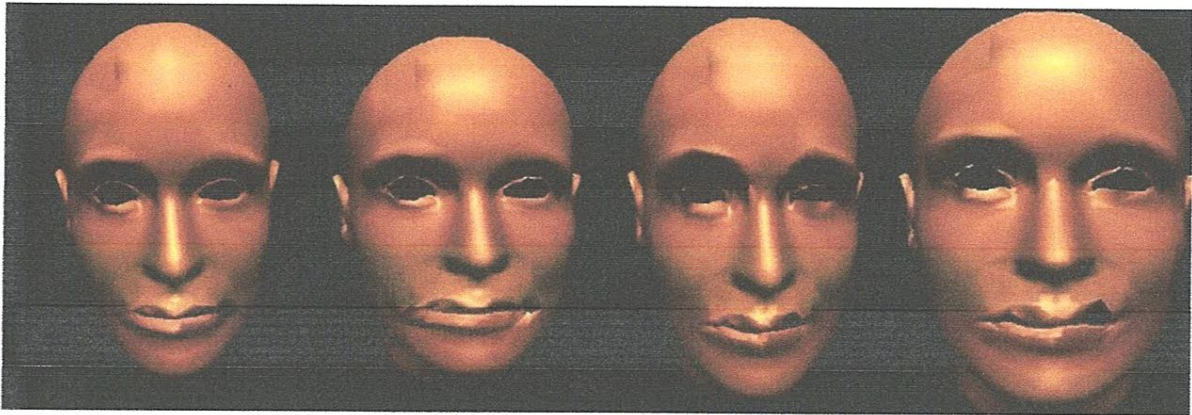


FIGURE 4.13 – Extraits d'animation des visages source et cible.

Tests Nous avons testé notre application :le transfert se fait sur le visage cible maillage 3d ,l'affichage de visage virtuel de format .off résulte un transfert des expression n'est pas clair et réaliste ne reflète pas exactement l'expression désirée comme sur le visage source.



FIGURE 4.14 – test sur les deux visages source et cible

4.6 Conclusion

Nous nous sommes intéressés dans ce dernier chapitre sur la procédure que nous avons suivie étape par étape pour la réalisation de notre application, ensuite nous avons également appliqué sur un environnement avec plusieurs bibliothèques compatibles. Au cours de ce travail, nous avons tenté d'apporter des solutions désirées. Malgré l'évolution rapide des techniques et la qualité dans le domaine de l'animation faciale, de nombreux points sont encore à améliorer comme le travail sur les détails fins dynamiques (texture et rides) ont permis d'améliorer le réalisme des scènes virtuelles.

Conclusion Générale

L'ANIMATION faciale est un des plus vieux domaines de l'infographie. C'est aussi un des moins évidents; c'est pourquoi il y a tant de pistes différentes qui ont été essayées, du dessin d'artiste à la synthèse d'images. Les animations 3D sont aujourd'hui incontournables dans le monde du multimédia, des films d'animations et de l'audiovisuel. Toutefois, les systèmes actuellement présents sur le marché ne permettent pas de répondre à la créativité toujours renouvelée de l'animation traditionnelle. Allier dans un même outil, possibilités artistiques et potentialités techniques et technologiques reste un important défi.

Nous avons présenté dans un premier temps le rôle de la modélisation de visage dans l'animation. Dans une deuxième partie : Le terme d'animation est très large et englobe en fait un nombre important de techniques s'appliquant à des niveaux diverses. Elles ont été détaillées ici en fonction d'une part de leur support (géométrie ou image), d'autre part de leur approche. La plupart des techniques récentes ne se limitent pas à une de ces pistes, mais sont des agencements de plusieurs d'entre elles. Les méthodes utilisées pour assurer le bon transfert arrivent à satisfaire les objectifs fixes sont basées sur des notions mathématiques. Les fonctions à base radiale sont essentielles dans chacune des étapes de notre approche, aussi bien au niveau de l'animation de visage, du transfert d'expression ou des déformations de maillage. Enfin dans la dernière partie, l'implémentation et le développement présentés en détail. Lors de la réalisation de ce travail, nous avons appris beaucoup de chose comme : la programmation en java, l'utilisation de différents logiciels comme le FaceGen et l'intégration de plusieurs bibliothèques.

Ainsi que ce mémoire ouvre quelques perspectives sur les futures travaux seront consacrés à résoudre les problèmes de transfert au niveau de la bouche et des yeux. Nous essaierons également de travailler sur les textures, pour appliquer les détails fins d'une animation (rides...) ou pour que nos méthodes de morphing prennent également en compte les textures de nos maillage sources et cibles.

BIBLIOGRAPHIE

- [1] Carol Leon-Yun Wang. Automating facial gestures and synthesized speech in human character animation. In Proceedings of the 1991 Western Computer Graphics Symposium, pages 39–40, April 1991.
- [2] Perrine Monjaux. Modélisation et animation interactive de visages virtuels de dessins animés. Computer Science. Université René Descartes - Paris V, 2007. France. ✓
- [3] Perrine Monjaux. Modélisation et animation interactive de visages virtuels de dessins animés. Computer Science. Université René Descartes - Paris V, 2007. French. ✗
- [4] Paul EKMAN and W. V. FRIESEN. Unmasking the Face. New Jersey :Pentice Hall, 1975.
- [5] F.I. Parke and K. Waters. Computer Facial Animation. A K Peters Wellesley, 1996. ISBN 1-56881-014-8.
- [6] A. Butfield, A new approach to rapid image morphing for lip motion synthesis, 2003, 26th Australasian computer science conference in research and practice in information Technology
- [7] J. Chai, J. Xiao, J. Hodgins, Vision-based Control of 3D Facial Animation, 2003, Eurographics/Siggraph Symposium on Computer Animation
- [8] F.I. Parke. Computer generated animation faces. In Proc. ACM annual conf.1972.
- [9] F.I. Parke. A Parametric Model for Human Faces. PhD thesis, University of Utah, Salt Lake City, Utah, 1974. UTEC-CSc-75-047.
- [10] J. Stokes and J. Ragan-Kelly. Final fantasy : The technology within. .
- ✗ [11] D. Terzopoulos, 1997. Siggraph : Panel on Facial Animation : Past, Present and Future.
- [12] Animation Faciale Temps Réel Multi Plates-formes , présentée à la Faculté des Sciences Economiques et Sociales de l'Université de Genève, Stéphane GARCHERY , Thèse no. 569 ,Genève, 2004
- [13] H. Kim, C. Joslin, T. Di Giacomo, S. Garchery and N. Magnenat-Thalmann, "Multi-resolution Meshes for Multiple Target, Single Content Adaptation within the MPEG-21 Framework", IEEE ICME Conference, June 2004.

- [14] Preda, M. Salomie, A. Prêteux, F. Lafruit, G. 2004. Virtual character definition and animation within the MPEG-4 standard. In M. Strintzis, N. Sarris (Ed.), 3D modeling and animation : Synthesis and analysis techniques for the human body, IRM Press, Hershey, PA, July 2004, pp. 27-69.
- [15] Preda, M. Prêteux, F. 2002. MPEG-4 Human Virtual Body Animation In M. Bourges-Sévenier (Ed.), MPEG-4 Jump-Start, Prentice Hall, Upper Saddle River, NJ, January 2002.
- [16] F. Scheepers, R. Parent, W. Carlson, S. May, Anatomy-based Modeling of the Human Musculature, 1997, Proceedings of Siggraph
- [17] T. Di Giacomo, C. Joslin, S. Garchery, and N. Magnenat-Thalmann. Adaptation of facial and body animation for mpeg-based architectures. In IEEE International Conference on CyberWorld (CW'03), December, 2003.
- [18] P. Kalra, A. Mangili, N. Magnenat-Thalmann, and D. Thalmann. Smile : Multi layered facial animation system. In IFIP WG 5.10, Tokyo, pages 189–198, 1991.
- [19] F. I. Parke, Computer Generated Animation of faces, Proc. ACM annual conf., 1972.
- [20] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, D. H. Salesin, Synthesizing Realistic Facial Expressions from Photographs, Siggraph proceedings, 1998, pp. 75-84
- [21] K. Waters, T. M. Levergood, Decface : An automatic Lip-synchronization Algorithm for Synthetic Faces, 1993, DEC. Cambridge Research Laboratory Technical ReportSeries
- [22] F. I. Parke, A Parametric Model for Human faces, Ph. D. Thesis, University of Utah, Salt lake City, Utah, 1974, UTEC-CSc-75-047
- [23] K. Arai, T. Kurihara, K. Anjyo, Bilinear Interpolation for Facial Expression and Metamorphosis in Real-time Animation, The Visual Computer, 1996, vol. 12 pp. 105-116.
- [24] Parke FI « A Parametric Model for Human Faces » PhD Thesis, University of Utah, Salt Lake City, Utah, UTEC-CSc-75-047, 1974
- [25] Parke, F.I. 1982. Parameterized Models for Facial Animation, In IEEE Computer Graphics and Applications, 2(9), pp. 61-68.
- [26] T. Kuratate, E. Vatikiotis-Bateson, and H. Yehia. Cross-subject face animation driven by facial motion mapping. Swets and Zeitlinger, 2003.
- [27] Nahas M, Hutric H, Rioux M, Domey J « Facial image syntheses using skin texture recording » Visual Computer, 6(6) :337–343, 1990.
- [28] Viad ML, Yahia II « Facial animation with wrinkles » Forsey D, Hegron G, editors, Proceedings of the Third Eurographics Workshop on Animation and Simulation, 1992.
- [29] K.Waters, A muscle model for animating three-dimensional facial expression. InMaureen C. Stone, editor, Computer Graphics (Siggraph proceedings, 1987) vol. 21 pp. 17-24

- [45] www.java3d.org.
- [46] www.facegen.com
- [47] www.luxand.com