

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn này là kết quả của quá trình tìm hiểu, nghiên cứu của bản thân dưới sự hướng dẫn của PGS.TS. Đỗ Năng Toàn cùng với quá trình học tập tại Trường Đại học Công nghệ - Đại học Quốc Gia Hà Nội. Trong luận văn có sử dụng một số mã nguồn mở và các tài liệu tham khảo đã được chỉ rõ. Nếu có gì sai phạm tôi xin hoàn toàn chịu trách nhiệm.

Người cam đoan

Tạ Thị Thảo

LỜI CẢM ƠN

Trong quá trình thực hiện luận văn, tôi đã nhận được nhiều sự hướng dẫn, giúp đỡ tận tình của các thầy cô, gia đình, bạn bè. Trước tiên, tôi xin được bày tỏ lòng biết ơn chân thành tới thầy giáo hướng dẫn, PGS.TS Đỗ Năng Toàn. Trong suốt quá trình làm luận văn, tôi đã nhận được sự giúp đỡ, động viên và đặc biệt sự hướng dẫn tận tình giúp tôi nắm rõ mục tiêu và định hướng nghiên cứu trong đề tài luận văn của mình.

Tôi xin được gửi lời cảm ơn đến tập thể các thầy cô giáo của Khoa Công nghệ thông tin – Trường Đại học Công nghệ đã trang bị cho tôi thêm kiến thức quý giá trong suốt thời gian học tập tại trường.

Tôi cũng xin được gửi lời cảm ơn đến tập thể Bộ môn Truyền thông đa phương tiện – Khoa Truyền thông đa phương tiện - Trường Đại học CNTT&TT – Đại học Thái Nguyên đã tạo điều kiện giúp đỡ tôi rất nhiều trong thời gian tôi tham gia khóa học.

Cuối cùng, với những tình cảm sâu sắc nhất, tôi xin chân thành gửi tới gia đình và bạn bè, những người đã luôn ở bên, động viên, chia sẻ với tôi về mọi mặt giúp tôi hoàn thành tốt khóa học của mình.

Thái Nguyên, ngày 26 tháng 10 năm 2016

MỤC LỤC

LỜI CAM ĐOAN.....	1
LỜI CẢM ƠN	2
DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT	5
DANH MỤC CÁC BẢNG BIỂU	6
DANH MỤC CÁC HÌNH VẼ.....	7
MỞ ĐẦU	9
CHƯƠNG 1 – TỔNG QUAN VỀ THỰC TẠI ẢO VÀ BÀI TOÁN BIỂU DIỄN TRẠNG THÁI KHUÔN MẶT 3D	11
1.1. Tổng quan về thực tại ảo	11
1.2. Đối tượng 3D.....	12
1.2.1. Giới thiệu về đối tượng 3D.....	12
1.2.2. Mô phỏng đối tượng 3D	13
1.3. Bài toán biểu diễn trạng thái khuôn mặt 3D.....	14
1.3.1. Giới thiệu	14
1.3.2. Cách thức thể hiện trạng thái biểu cảm trên khuôn mặt người.....	16
1.3.3. Mô hình hóa khuôn mặt 3D	19
1.3.4. Một số phương pháp biểu diễn trạng thái khuôn mặt 3D.....	26
CHƯƠNG 2 - MỘT SỐ KỸ THUẬT NỘI SUY TRONG BIỂU DIỄN TRẠNG THÁI KHUÔN MẶT 3D	35
2.1. Kỹ thuật nội suy tuyến tính.	35
2.1.1. Giới thiệu	35
2.1.2. Sử dụng trong biểu diễn trạng thái khuôn mặt 3D	37
2.1.3. Nhận xét.....	42
2.2. Kỹ thuật nội suy song tuyến.	43
2.2.1. Giới thiệu	43
2.2.2. Sử dụng trong biểu diễn trạng thái khuôn mặt 3D	44
2.2.3. Nhận xét.....	47
2.3. Kỹ thuật nội suy dựa trên hàm cơ sở bán kính.....	47
2.3.1. Giới thiệu	47
2.3.2. Sử dụng trong biểu diễn trạng thái khuôn mặt 3D	50

2.3.3. Nhận xét.....	52
CHƯƠNG 3 –THỰC NGHIỆM	54
3.1. Giới thiệu bài toán	54
3.2. Các công cụ kỹ thuật	55
3.3. Kết quả thực nghiệm	58
KẾT LUẬN	62
TÀI LIỆU THAM KHẢO.....	63

DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT

STT	Chữ viết tắt	Từ hoặc cụm từ
1	VR	Virtual Reality
2	AMA	Abstract Muscle Action
3	FACS	Facial Action Coding System
4	FFD	Freeform Deformations
5	AU	Action Unit
6	RBF	Radial Basic Funtion

DANH MỤC CÁC BẢNG BIỂU

Bảng 1.1. Các mô hình khuôn mặt 3D

Bảng 2.1. Biểu diễn hành động và các đơn vị của chúng.

DANH MỤC CÁC HÌNH VẼ

Hình 1.1. Các đặc tính của hệ thống thực tại ảo

Hình 1.2. Đối tượng 3D

Hình 1.3. Nhân viên tiếp tân ảo tại quận Brent

Hình 1.4. Trạng thái vui vẻ

Hình 1.5. Trạng thái buồn

Hình 1.6. Trạng thái tức giận

Hình 1.7. Khuôn mặt sợ hãi của một chú bé trong hoàn cảnh xa lạ

Hình 1.8. Trạng thái ngạc nhiên

Hình 1.9. Trạng thái căm ghét

Hình 1.10. Mô hình khuôn mặt được đề xuất bởi Parke vào năm 1974

Hình 1.11. Mô hình mặt CANDIDE

Hình 1.12. Mô hình mặt Morphable

Hình 1.13. Kết cấu phẳng trong lưới Cartesian

Hình 1.14. Mô hình được phát triển bởi Golovinskiy

Hình 1.15. Mô hình khuôn mặt của Ramirez

Hình 1.16. Biến dạng tự do

Hình 1.17. Sợi cơ và cơ trong mô hình của Platt và Badler

Hình 1.18. Mô hình cơ của Parke và Waters, 1996

Hình 1.19. Mô hình cậu bé của Kahler và các đồng nghiệp

Hình 2.1. Nội suy tuyến tính trên một tập dữ liệu

Hình 2.2. Nội suy cosin trên một tập dữ liệu

Hình 2.3. Phép nội suy tuyến tính được thực hiện trên các giá trị kết hợp cơ mặt

Hình 2.4. Minh họa phép nội suy song tuyến tính

Hình 2.5. Dữ liệu hình dạng hai trạng thái khuôn mặt

Hình 2.6. Nội suy song tuyến tính

Hình 2.7. Hàm cơ sở bán kính (RBF)

Hình 2.8. Các điểm điều khiển của mô hình khuôn mặt chung

Hình 2.9. Minh họa sự biến đổi của điểm điều khiển.

Hình 3.1. Mô hình khuôn mặt 3D được xây dựng trong Maya

Hình 3.2. Các trạng thái biểu cảm cơ bản của khuôn mặt

Hình 3.3. Mô hình khuôn mặt 3D dưới dạng lưới

Hình 3.4. Mô hình khuôn mặt 3D ở trạng thái ban đầu

Hình 3.5. Các trạng thái biểu cảm cơ bản của khuôn mặt

Hình 3.6. Quá trình hoạt hóa từ trạng thái khuôn mặt tự nhiên sang trạng thái vui vẻ

Hình 3.7. Hoạt hóa từ trạng thái vui vẻ sang trạng thái tức giận

MỞ ĐẦU

Ngày nay, sự phát triển mạnh mẽ của công nghệ thông tin và trí tuệ nhân tạo cùng với phần cứng máy tính, các thiết bị ngoại vi như máy ảnh, máy quét 3D, v.v. đã làm cho việc thu nhận hình ảnh, mô hình 3D, xây dựng đối tượng 3D ngày càng phát triển. Chúng ta có thể làm việc và giao tiếp với máy tính như với con người.

Trong lĩnh vực thực tại ảo, các nghiên cứu về mô phỏng con người cũng đang ngày càng được nghiên cứu chuyên sâu và đã có nhiều ứng dụng trong thực tế, phải kể đến các ứng dụng về người ảo như giáo viên ảo, bệnh nhân ảo phục vụ cho việc thực hành, thí nghiệm của các sinh viên trường y, phụ tá ảo được sử dụng tại các bưu điện ở Anh, hay các nhân vật trong các bộ phim viễn tưởng đẹp mắt như Avatar, Star Trek ... Người ảo có thể thay thế con người trong nhiều công việc để giảm chi phí, tăng hiệu quả công việc. Người ảo cũng có những đặc điểm căn bản của con người như khả năng giao tiếp, cảm nhận thế giới, suy nghĩ, biểu hiện cảm xúc .v.v. Trong các hệ thống mô phỏng con người này, việc biểu diễn biểu cảm khuôn mặt 3D cũng là một vấn đề hết sức quan trọng. Khi biểu diễn một cách chính xác có cảm xúc các trạng thái trên khuôn mặt 3D có thể xây dựng được các giáo viên ảo có thể thể hiện cảm xúc của mình khi người học mắc lỗi, khi nói về một nội dung mang tính hài hước, hay có thể xây dựng các bệnh nhân ảo có các biểu hiện cảm xúc khi các bác sĩ khám bệnh cho mình giống như các bệnh nhân trong thực tế, hoặc xây dựng được các nhân vật mà người ta tưởng tượng ra trong phim ảnh, hoặc các nhân vật game có những đặc tính giống như con người như vẻ mặt biểu cảm, khả năng thể hiện cảm xúc khi giao tiếp .v.v.

Ở Việt Nam, các nghiên cứu liên quan tới lĩnh vực này cũng đang được chú trọng trong những năm gần đây tại các đơn vị như Đại học Quốc Gia Hà Nội, Đại học Thái Nguyên, Viện công nghệ thông tin – Viện Hàn lâm Khoa học và Công nghệ Việt Nam, .v.v. Đặc biệt là hiện nay, một đề tài cấp bộ có tên là *“Phát triển các kỹ thuật biểu diễn cử chỉ, trạng thái khuôn mặt 3D dựa trên các kỹ thuật nội suy phục vụ cho bài toán xây dựng phát thanh viên ảo”* cũng được nghiên cứu và thực hiện tại Viện Hàn lâm khoa học và Công nghệ Việt Nam.

Xuất phát từ hoàn cảnh đó, luận văn lựa chọn đề tài “***Biểu diễn trạng thái khuôn mặt 3D dựa vào kỹ thuật nội suy***” nhằm nghiên cứu về việc biểu diễn trạng thái khuôn mặt 3D trong thực tại ảo và nghiên cứu về kỹ thuật nội suy trong việc biểu diễn trạng thái khuôn mặt 3D này.

Cấu trúc của luận văn bao gồm:

Chương 1: Tổng quan về thực tại ảo và bài toán biểu diễn trạng thái khuôn mặt 3D

Chương 2: Một số kỹ thuật nội suy trong biểu diễn trạng thái khuôn mặt 3D

Chương 3: Chương trình thực nghiệm

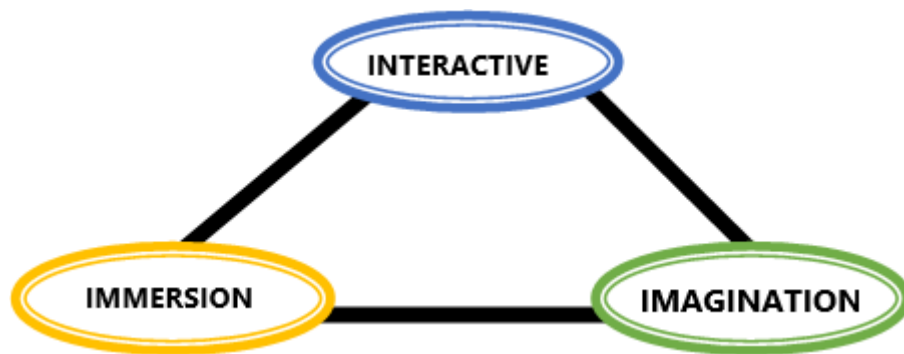
CHƯƠNG 1 – TỔNG QUAN VỀ THỰC TẠI ẢO VÀ BÀI TOÁN BIỂU DIỄN TRẠNG THÁI KHUÔN MẶT 3D

1.1. Tổng quan về thực tại ảo

VR- Virtual Reality – Thực tại ảo (hay thực tế ảo): là một hệ thống giao diện cấp cao giữa người dùng và máy tính. Hệ thống này mô phỏng các sự vật hiện tượng theo thời gian thực và tương tác với người sử dụng thông qua các kênh cảm giác như thị giác, thính giác, xúc giác, khứu giác và vị giác.

Hay nói một cách khác là, người sử dụng có thể di chuyển, quan sát, tương tác với một thế giới nhân tạo được xây dựng mô phỏng bằng máy tính như đối với thế giới thực. Người dùng có thể nhìn thấy các đối tượng đồ họa 3D nổi, điều khiển được các đối tượng trong thế giới nhân tạo đó và có thể sờ, nắm, cảm nhận các đối tượng đó giống như trong thế giới thực.

Một hệ thống thực tại ảo có 3 đặc tính chính: Tương tác (interactive) – Đắm chìm (Immersion) – Tưởng tượng (Imagination)



Hình 1.1. Các đặc tính của hệ thống thực tại ảo

+ Tính tương tác: hệ thống thực tại ảo có khả năng tạo ra một thế giới nhân tạo giống như một thế giới thật. Thế giới nhân tạo này không tĩnh mà lại phản ứng, thay đổi theo ý muốn của người sử dụng. Người sử dụng có thể tương tác với các mô hình 3D ảo trong thế giới ảo đó và thay đổi ngay lập tức theo ý muốn của mình.

+ Đắm chìm: Hệ thống thực tại ảo có khả năng thu hút người sử dụng vào trong một thế giới ảo giống như thật, người dùng sẽ có cảm giác trở thành một phần của các hành động trên màn hình khi họ đang trải nghiệm thực tại ảo.

Người dùng có thể nhìn thấy, sờ thấy và cảm thấy như các đối tượng có thật trong thực tế.

+ **Tưởng tượng:** Có hai khía cạnh của tính tưởng tượng trong một hệ thống thực tại ảo: Thứ nhất chính là khả năng di chuyển khắp nơi một cách độc lập như đang ở bên trong thế giới thật của người dùng. Thứ hai là những quy tắc về cách thức mà con người, sự vật, hiện tượng trong thế giới ảo tương tác với nhau theo một trật tự để trao đổi năng lượng và thông tin.

Một hệ thống thực tại ảo tổng quát sẽ bao gồm có 5 thành phần: Phần mềm, phần cứng, mạng liên kết, người dùng và các ứng dụng. Trong đó có phần mềm, phần cứng và các ứng dụng là quan trọng nhất.

Ngày nay, thực tại ảo đang có rất nhiều các ứng dụng được sử dụng vào thực tế trong tất cả các lĩnh vực như y học, giáo dục, tin học, hàng không, giải trí, quân sự, quốc phòng, thiết kế và trang trí nội thất, v.v. Đặc biệt là trong lĩnh vực giải trí, game hiện nay, người dùng đang bị thu hút bởi một trò chơi thực tại ảo mới ra đời sử dụng trên điện thoại di động đó là POKEMON GO. Điều đó cho thấy tiềm năng rất lớn của thực tại ảo ứng dụng trong tương lai.

1.2. Đối tượng 3D

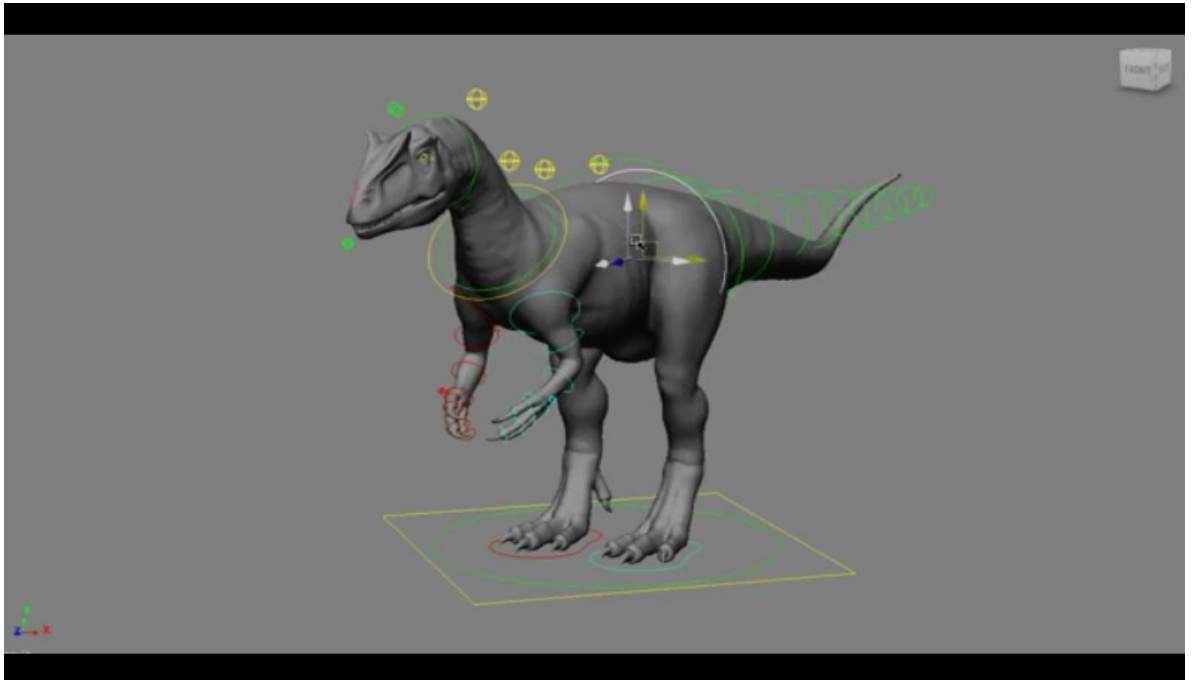
1.2.1. Giới thiệu về đối tượng 3D

Khái niệm đối tượng thường được dùng để chỉ các vật thể tồn tại xung quanh chúng ta, chúng có thể là các đối tượng có thực trong thế giới thực như con người, nhà cửa, cây cối, sinh vật, v.v. hoặc chúng cũng có thể là các đối tượng do con người chúng ta tưởng tượng ra được xây dựng và mô phỏng trong thế giới ảo.

Đối tượng 3D là đối tượng mô phỏng lại các đối tượng trong thế giới thực hoặc mô phỏng lại đối tượng mà chúng ta tưởng tượng ra, bằng máy tính. Đối tượng 3D là một phần trong thế giới ảo. Để xây dựng nên một không gian ảo, chúng ta phải có các đối tượng 3D sau đó kết hợp các đối tượng này với các thiết bị ngoại vi để giúp người sử dụng điều khiển và tương tác được với thế giới ảo giống như tương tác với các đối tượng trong thực tế. Một đối tượng 3D bao gồm có hai phần chính:

- Thứ nhất là một mô hình hình học biểu diễn đường nét, hình dáng của đối tượng và màu sắc (hay chất liệu) của đối tượng đó được gọi chung là mô hình 3D.

- Thứ hai là các thuộc tính, các ràng buộc và các mối quan hệ giữa các thành phần của đối tượng đó.



Hình 1.2. Đối tượng 3D

1.2.2. Mô phỏng đối tượng 3D

Có hai cách để thực hiện mô phỏng đối tượng 3D:

- Thứ nhất là mô phỏng các mô hình 3D nhờ các ngôn ngữ lập trình truyền thống như C, C++ .v.v. Cách này có thể thực hiện các mô phỏng phức tạp đòi hỏi sự chính xác cao, tuy nhiên nó cũng đòi hỏi máy tính phải có cấu hình mạnh về phần cứng, người lập trình phải có trình độ lập trình cao vì cách này gồm rất nhiều các thuật toán phức tạp, mất nhiều thời gian và cũng rất khó trong việc tạo ra những cảnh rộng lớn. Mặc dù vậy, nó lại là một lựa chọn duy nhất cho những ai muốn mô phỏng chính xác các sự vật, hiện tượng tự nhiên đúng với bản chất của nó. Ví dụ như mô phỏng nước, mô phỏng lửa, mô phỏng các hiện tượng vật lý, .v.v. Mô phỏng mô hình 3D nhờ các ngôn ngữ lập trình truyền thống thường chỉ thích hợp trong những trường hợp mô phỏng có quy mô nhỏ hoặc sử dụng trong quá trình học tập.

- Cách thứ hai là sử dụng các công cụ mô phỏng đã được xây dựng sẵn như 3DSmax, Maya, Autocad, .v.v. Cách này thường phù hợp với việc mô phỏng các mô hình không yêu cầu độ chính xác cao và nó không đòi hỏi người mô phỏng phải có trình độ lập trình cao, không tốn nhiều thời gian thực hiện. Tuy nhiên nó cũng có nhược điểm là yêu cầu cấu hình hệ thống phải mạnh để cài đặt và chạy chương trình, đặc biệt là khi xuất (Rendering) mô hình. Hiện nay, cách này rất được ưa chuộng và được sử dụng rất phổ biến trong các công việc như làm Game 3D, Web 3D, dựng các công trình kiến trúc, hoặc phim hoạt hình 3D, .v.v..

Khuôn mặt 3D và các trạng thái của khuôn mặt 3D cũng là các đối tượng 3D. Việc xây dựng mô hình 3D khuôn mặt có thể sử dụng các phần mềm dựng sẵn và việc mô phỏng lại sự thay đổi các trạng thái khuôn mặt có thể sử dụng các phần mềm có sẵn hoặc sử dụng các ngôn ngữ lập trình truyền thống.

1.3. Bài toán biểu diễn trạng thái khuôn mặt 3D

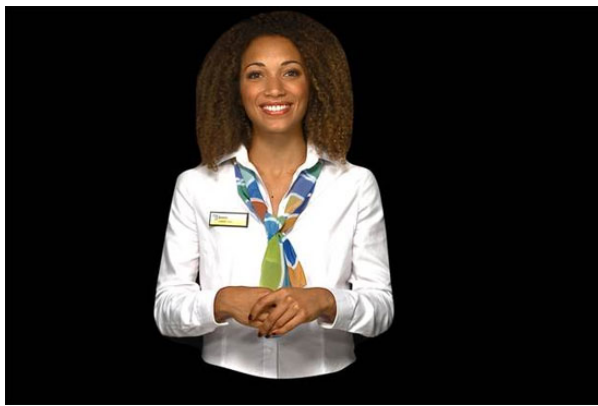
1.3.1. Giới thiệu

Ngày nay, với sự phát triển của các phần mềm thông minh, chúng ta có thể giao tiếp với máy tính như thể chúng ta giao tiếp với con người. Con người luôn tìm cách biến máy tính thành con người bằng cách trang bị cho nó giao diện giống con người thông qua giọng nói, khuôn mặt, cử chỉ, .v.v. thay vì sử dụng bàn phím, chuột hay những dòng chữ vô cảm. Người ảo là một trong những ứng dụng thể hiện mong muốn đó của con người.

Người ảo là một nhân vật có hình dáng, hành động tương tự như con người. Người ảo có thể giao tiếp với con người thông qua lời nói cử chỉ, có khả năng cảm nhận, cảm thụ thế giới, có suy nghĩ hành động, có thể thể hiện các trạng thái tâm lý giống như con người.

Hiện nay, trên thế giới đã có rất nhiều nhà nghiên cứu tiến hành xây dựng mô hình nhân vật ảo có khả năng thể hiện mô tả sự chuyển động, nói, phát âm, thể hiện cảm xúc, .v.v. như dự án bệnh viện ảo của Khoa kỹ thuật – khoa học máy tính và thông tin thuộc Đại học Florida (Mỹ). Khoa này đã xây dựng một nhân vật bệnh nhân ảo phục vụ cho việc giảng dạy sinh viên ngành y, các sinh viên có thể nói chuyện, phỏng vấn, sờ nắn bệnh nhân ảo để thực hành thăm khám, bệnh nhân ảo cũng có khả năng biểu lộ cảm xúc và trả lời các câu hỏi của

bác sĩ. Hay tại bưu điện Anh người ta đã xây dựng một nhân vật ảo được sử dụng để giúp những người khiếm thính mua tem và điền đúng các giấy tờ bằng cách dịch ngôn ngữ nói sang ngôn ngữ dấu hiệu của người anh. Hoặc mới đây chính quyền quận Brent của London (Anh) đã phóng một nhân viên trợ lý ảo 3D lên một màn hình như thể nhân viên này đang ngồi sau bàn tiếp tân tại văn phòng hội đồng quận Brent. Nhân viên ảo này sẽ được lập trình để trả lời một số câu hỏi về khai sinh, khai tử, đăng ký kết hôn và quyền công dân, cũng như chỉ dẫn lối đi cho khách viếng thăm. Việc sử dụng nhân viên ảo này sẽ làm cho chính quyền quận Brent tiết kiệm được rất nhiều chi phí, tiền bạc mà lại không ảnh hưởng đến dịch vụ công.



Hình 1.3. Nhân viên tiếp tân ảo tại quận Brent

Để có được một nhân vật ảo ứng dụng trong các lĩnh vực khác nhau như trên, các nhà nghiên cứu đã đặt ra rất nhiều các bài toán liên quan tới việc xây dựng các nhân vật ảo như: Mô phỏng trạng thái khuôn mặt, mô phỏng tư thế con người, mô phỏng cử chỉ, hành động con người, mô phỏng chuyển động đầu, mô phỏng phản hồi trong giao tiếp, .v.v. Trong đó, bài toán mô phỏng trạng thái khuôn mặt 3D là một bài toán hết sức quan trọng trong việc xây dựng các nhân vật ảo 3D, đồng thời cũng có nhiều ứng dụng trong các lĩnh vực như điện ảnh truyền hình, game, y học, giáo dục, .v.v. Đặc biệt là việc biểu diễn trạng thái khuôn mặt 3D có thể kết hợp với hệ chuyên gia và trí tuệ nhân tạo để tạo nên những ứng dụng như khám bệnh qua internet, phân tích hoạt ảnh mặt tìm tội phạm hay là một trong những nhu cầu tạo nên thế giới ảo thân thiện có tính trị liệu tâm lý cho mọi người, ..v.v.. Đây là các ứng dụng mang tính thực tế và hiệu quả kinh tế cao góp phần thúc đẩy cho sự phát triển kinh tế xã hội của đất nước.

Tuy nhiên, việc biểu diễn các trạng thái khuôn mặt 3D trên máy tính cũng gặp phải không ít những khó khăn trong việc mô phỏng do sự phức tạp của giải

phần trên khuôn mặt người và sự nhạy cảm tự nhiên tới sự xuất hiện trên khuôn mặt. Hơn nữa, việc mô hình hóa khuôn mặt 3D một cách tối ưu, làm sao cho mô hình vừa nhẹ lại vừa đảm bảo độ chi tiết phù hợp cho việc hoạt hóa đúng các trạng thái trên khuôn mặt đảm bảo cho hệ thống hoạt hóa khuôn mặt phải tạo ra các hoạt ảnh động thực tế của mặt, thao tác trong thời gian thực, .v.v. là một việc không hề dễ dàng. Có rất nhiều cách giải quyết được đưa ra, dưới đây luận văn xin trình bày một số cách tiếp cận trong việc mô hình hóa, hoạt hóa biểu diễn trạng thái biểu cảm khuôn mặt 3D.

1.3.2. Cách thức thể hiện trạng thái biểu cảm trên khuôn mặt người

Trạng thái khuôn mặt người là thể hiện của khuôn mặt trong một khoảng thời gian nhất định nào đó dùng để mô tả những tính chất hình dáng nhất định của khuôn mặt ở một thời điểm nhất định. Mỗi trạng thái khuôn mặt có thể xuất hiện dài hay ngắn tùy theo mức độ ảnh hưởng và làm thay đổi nó của những yếu tố tác động lên việc thể hiện cảm xúc của con người.

Theo Shaver và các đồng nghiệp [8] các biểu cảm trên khuôn mặt con người sẽ gồm có 6 trạng thái biểu cảm cơ bản là vui vẻ, buồn, ngạc nhiên, tức giận, sợ hãi, căm ghét.

- Trạng thái vui vẻ: vui vẻ là trạng thái tâm lý của con người khi gặp những chuyện vui, những việc làm mình cảm thấy hài lòng. Trạng thái vui vẻ được thể hiện với một số đặc điểm nổi bật trên khuôn mặt như được thể hiện qua nụ cười tươi làm cho hở hàm răng, hai khóe miệng căng sang hai bên và dịch chuyển lên phía trước, gò má nhô cao và đuôi mắt nhăn.



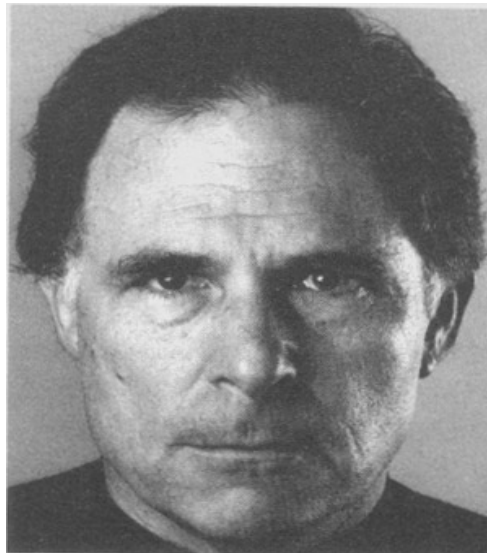
Hình 1.4. Trạng thái vui vẻ

- Buồn: là trạng thái khi con người gặp việc đau thương hoặc đang có điều gì đó không được như ý muốn. Trạng thái buồn được thể hiện trên khuôn mặt như vùng trán nhăn, góc của đôi môi kéo xuống, lông mày phía trong nhô lên và kéo gần vào nhau, mí mắt trùng xuống, và có thể là hàm giảm xuống đối với người đau đớn hoặc đang khóc.



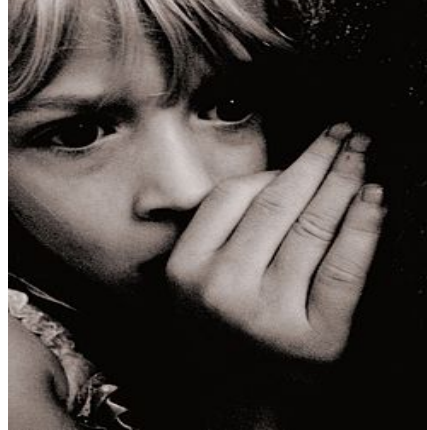
Hình 1.5. Trạng thái buồn

- Tức giận: là trạng thái của con người khi tức giận với một ai đó khi họ làm điều gì khiến bản thân mình không vui hoặc điều gì đó khiến bản thân cảm thấy bực bội trong người. Trạng thái tức giận thường biểu hiện qua điệu bộ đôi lông mày kéo lại gần nhau và hạ thấp, môi mím chặt hoặc hàm nghiến chặt, ánh mắt giận dữ, mí mắt căng, cũng có khi hở hai hàm răng tiếp xúc nhau thì miệng sẽ có dạng hình chữ nhật.



Hình 1.6. Trạng thái tức giận

- **Sợ hãi:** chính là trạng thái biểu đạt sự hoảng loạn về tinh thần trước một điều gì đó ảnh hưởng tới bản thân, nó xuất hiện từ việc nhận thức các mối đe dọa, nguy hiểm trước mắt. Trạng thái sợ hãi đi kèm với một số biểu cảm trên khuôn mặt như lông mày nhô lên, sát vào nhau, mắt mở to, mi mắt căng ra và môi kéo giật ra sau.



Hình 1.7. Khuôn mặt sợ hãi của một chú bé trong hoàn cảnh xa lạ

- **Ngạc nhiên** là trạng thái con người cảm thấy bất ngờ về một việc gì đó xảy ra một cách bất ngờ ngay trước mắt họ. Thông thường trạng thái này chỉ xảy ra trong một khoảng thời gian rất ngắn, nét mặt diễn ra nhanh, có đôi khi là chưa tới 1 giây. Trạng thái ngạc nhiên được biểu hiện trên khuôn mặt như lông mày nhô cao, mắt mở to, hàm dưới trề xuống và miệng há hốc.



Hình 1.8. Trạng thái ngạc nhiên

- **Căm ghét (hay khinh bỉ):** là trạng thái của con người trước một điều gì đó mà họ cảm thấy không thích, không hài lòng thậm chí là coi thường sự việc đó. Trạng thái căm ghét thường biểu hiện qua điệu bộ nhăn mũi, lông mày hạ thấp, mắt gằn như nhắm hấn và môi trên thì chun lên.



Hình 1.9. Trạng thái căm ghét

1.3.3. Mô hình hóa khuôn mặt 3D

Gần đây, mô hình hóa khuôn mặt người trở nên có nhiều ý nghĩa bằng cách đặt ra nhiều thách thức trong lĩnh vực y học, kỹ thuật, hoạt hình và đồ họa máy tính. Do đó, quá trình tạo ra những gương mặt tổng hợp một cách thực tế với các đặc trưng ba chiều và cố gắng để làm cho nó hấp dẫn đã thu hút nhiều sự chú ý quan tâm của các nhà nghiên cứu. Với sự ra đời của các thuật toán và kỹ thuật về mặt hình ảnh, nó có thể tạo ra các dạng hình học của khuôn mặt người một cách chi tiết sử dụng các phần mềm hình ảnh và máy quét 3D. Trong luận văn này, tôi xin trình bày một số mô hình khuôn mặt 3D và cách mô hình hóa các mô hình khuôn mặt 3D đã có trong một số nghiên cứu gần đây.

Mô hình hóa khuôn mặt được giới thiệu lần đầu tiên bởi Gouraud vào năm 1971. Sau đó Parke (1972) đã cải tiến mô hình này để tạo nên một mô hình khuôn mặt với 250 đa giác và 400 đỉnh. Sau này đã có một số nghiên cứu chi tiết hơn về việc mô hình hóa khuôn mặt như mô hình hóa khuôn mặt với lưới đa giác và mô hình hóa với bề mặt tham số được trình bày trong [10] của Bùi Thế Duy, Pighin và các đồng nghiệp [11] đã trình bày kỹ thuật để tạo ra các mô hình mặt 3D từ các hình ảnh của một đối tượng người. Zhang cùng các đồng nghiệp [12] giới thiệu một hệ thống biểu cảm khuôn mặt định hướng hình học bằng cách sử dụng cách tiếp cận dựa trên ví dụ. Phương pháp Bump mapping của Blinn [13] đã cố gắng thay đổi bề mặt bình thường trước khi tới các tính toán để đạt được một hình ảnh trực quan của nếp nhăn mà không sử dụng biến dạng hình học. Ý tưởng này đã được sử dụng một cách rộng rãi đặc biệt là cho các nếp nhăn trên khuôn mặt [14].

Conde [27] giải thích rằng ngày nay có thể có được hai loại dữ liệu khuôn mặt, dữ liệu hình ảnh (2D) và dữ liệu ba chiều (3D). Loại dữ liệu đầu tiên đại

diện cho các kết cấu của khuôn mặt trong khi loại thứ hai thu thập những cấu trúc hình học trên khuôn mặt. Cả hai loại dữ liệu bổ sung cho nhau bởi vì chúng cung cấp hai loại thông tin khác nhau. Tuy nhiên, hạn chế chính của hệ thống biểu diễn khuôn mặt dựa trên kết cấu là sự phụ thuộc vào các điều kiện ánh sáng và vị trí của khuôn mặt. Trong khi đó 3D rendering không phụ thuộc vào sự chiếu sáng và cho phép chuẩn hóa vị trí của khuôn mặt 3D.

Với đặc điểm và mục đích này, các mô hình khuôn mặt 3D lý tưởng theo Golovinskiy và các đồng nghiệp [29], phải được phổ cập, dễ dàng thu nhận và tính toán, mục tiêu của nó là để mô phỏng lại tất cả các khía cạnh khuôn mặt của một người thành một mô hình khuôn mặt 3D biểu diễn được trên máy tính. Cootes và Taylor [31] cũng chỉ ra rằng một mô hình là hữu ích nếu nó cụ thể, chi tiết, điều này có nghĩa là mô hình có khả năng biểu diễn cho các đối tượng mô hình hóa.

Bảng 1.1 trình bày tóm tắt về các nghiên cứu của các tác giả theo thời gian trong việc xây dựng các mô hình khuôn mặt 3D với các thông tin như tác giả, năm xuất bản, số đỉnh, số lượng đa giác và số lượng các đối tượng sử dụng để xây dựng mô hình khuôn mặt 3D.

Bảng 1.1. Các mô hình khuôn mặt 3D

Model/author/year	Purpose	Vertices	Polygons	Subjects
Park [25], 1974	Speech animation	356	334	----
Candide – 3, 2001	Facial expressions and animation	75	100	----
Morphable Model, Blaz y Vetter[24], 1999	Generation of human faces	70,000	----	200
Bronstein et al [28], 2003	Facial recognition	2,000 a 2,500	----	157
El-Husuna [32], 2003	Facial model	----	----	8
Golovinskiy et al [29], 2006	Facial details modeling	----	500,000	149
Basel Face Model, Paysan et al, 2009	Generation of human faces	53,490	160,470	200
Ramirez et al [30], 2009	Facial recognition	2,777	2,676	----

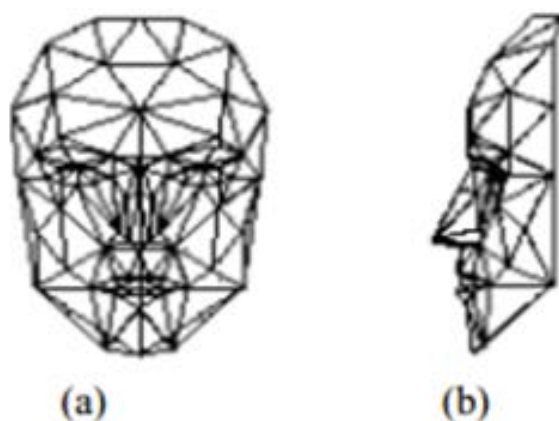
Phothisane et al [26], 2011	Head tracking	3,000	----	200
--------------------------------	---------------	-------	------	-----

Parke [25] trình bày một mô hình khuôn mặt ba chiều, đối xứng và tham số tạo hình ảnh khuôn mặt. Khuôn mặt được xây dựng bởi các bề mặt đa giác và thao tác thông qua các tham số để thực hiện nội suy điều khiển, dịch chuyển, xoay và co giãn các đặc trưng khuôn mặt khác nhau với mục đích cho phép biểu diễn biểu cảm khuôn mặt. Mô hình đề xuất bởi Parke[25] yêu cầu thông tin đầu vào nhỏ để xác định và tạo ra một khuôn mặt cụ thể trong một biểu cảm cụ thể. Các mô hình của Parke dựa trên thuật toán của Watkins để giải quyết vấn đề của bề mặt thấy được tới các đối tượng đa giác và kỹ thuật của Gouraud cho các đối tượng này một bề mặt mịn hơn. Hình 1.10 cho thấy mô hình mặt được phát triển bởi Parke. Parke thu thập dữ liệu từ 10 trạng thái khuôn mặt khác nhau và đã thực hiện một bộ phim hoạt hình để hiển thị các hiệu ứng chuyển tiếp từ trạng thái khuôn mặt này tới trạng thái khuôn mặt khác.



Hình 1.10. Mô hình khuôn mặt được đề xuất bởi Parke vào năm 1974

Mô hình mặt Candide là một mặt nạ tham số được thiết kế đặc biệt để mô hình hóa khuôn mặt của con người. Mô hình này xử lý khoảng 100 đa giác, cho phép xây dựng lại mô hình khuôn mặt một cách nhanh chóng với khả năng tính toán chấp nhận được và được điều khiển bởi các đơn vị hành động địa phương (AUs – Action Units). Mô hình CANDIDE ban đầu được xây dựng bởi Rydfalk tại trường đại học Linköping Thụy Điển vào năm 1987. Mô hình ban đầu có 75 đỉnh và 100 tam giác và nó đã được thay đổi 3 lần với các đỉnh được thêm vào để che đi toàn bộ đầu trục diện và vai để hoạt hóa một cách đơn giản bởi MPEG-4. Hình 1.11 biểu diễn mô hình mặt nạ CANDIDE.



Hình 1.11. Mô hình mặt CANDIDE

Blanz y Vetter [24] đề xuất một kỹ thuật tham số của mô hình khuôn mặt cho khuôn mặt con người tùy ý. Nghiên cứu thống kê dữ liệu khuôn mặt 3D (bao gồm cả hình dạng và kết cấu) của tác giả đã xây dựng một mô hình Morphable mặt mẫu và thu được tri thức về các biến thể trên khuôn mặt sử dụng mô hình phân lớp. Tính toán khuôn mặt trung bình và các biến thể chính trong nhóm dữ liệu, họ thu được một phân bố xác suất để tránh “các khuôn mặt không phải của con người”.

Phần nền tảng cơ bản của Blanz y Vetter là một mô hình khuôn mặt con người chung. Mô hình này giới hạn phạm vi của khuôn mặt được chấp nhận dựa trên các mẫu khuôn mặt người và mô hình này sử dụng hàng nghìn điểm trên mỗi khuôn mặt dẫn đến một tam giác trực tiếp của bề mặt mà không đòi hỏi bất kỳ kỹ thuật nội suy nào cho các biến thể của bề mặt và cũng thêm một mô hình biến thể kết cấu.

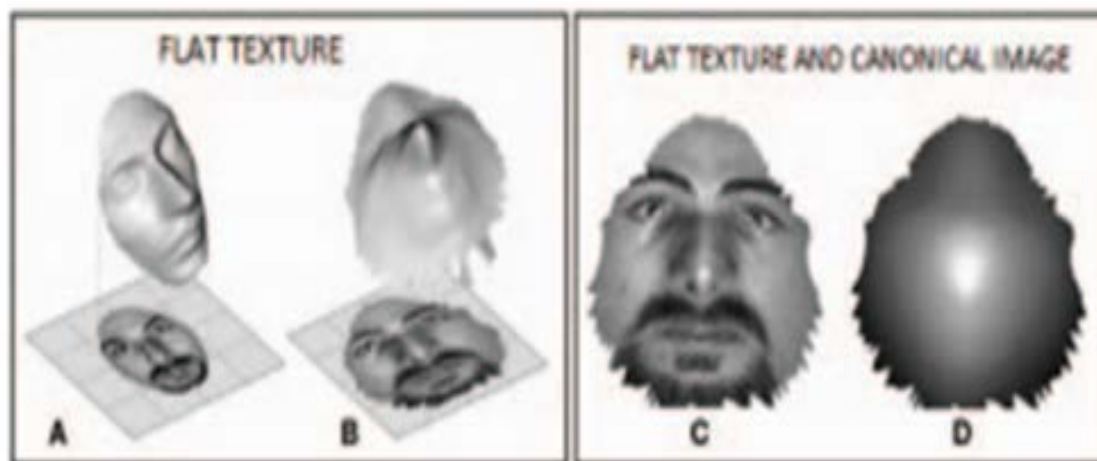
Mục tiêu của Blanz y Vetter là biểu diễn bất kỳ khuôn mặt nào như một sự kết hợp tuyến tính của một nhóm khuôn mặt mẫu. Các tác giả sử dụng các phép đo thống kê để duy trì một biến dạng đáng tin cậy. Các cơ sở dữ liệu được sử dụng bởi Blanz và Vetter bao gồm 200 đầu người được quét 3D của thanh niên (100 nam và 100 nữ) không có tóc, không trang điểm hoặc phụ kiện. Các mô hình thu được có 10,000 đỉnh cho vị trí và số lượng tương tự cho các giá trị màu. Hình 1.12 cho thấy khuôn mặt ý nghĩa là cơ sở để tạo ra các khuôn mặt mới.



Hình 1.12. Mô hình mặt Morphable

Bronstein và các đồng nghiệp [28] trình bày một phương pháp hình học để nhận dạng khuôn mặt sử dụng dữ liệu trong khuôn gian 3D. Phương pháp của họ là dựa vào các bất biến hình học của khuôn mặt người và tạo ra một so sánh của các bề mặt cho phép biến dạng địa phương của khuôn mặt cho các biểu cảm khuôn mặt. Nghiên cứu này giải thích rằng khuôn mặt con người không thể được coi là các đối tượng cứng nhắc và dựa trên những quan sát thực nghiệm các tác giả kết luận rằng các biểu cảm mặt có thể được mô hình hóa như biến đổi isometric. Các tác giả sử dụng các dạng chuẩn bất biến uốn cong như là trung tâm trong công trình nghiên cứu của họ cho quá trình nhận dạng khuôn mặt.

Về thu nhận hình ảnh, các tác giả[28] sử dụng kỹ thuật mã hóa ánh sáng để có được hình dạng của các khuôn mặt dựa vào các biến dạng mô hình. Những hình ảnh 3D được chuyển đổi trong một bề mặt tam giác bằng cách sử dụng bề mặt splines, bề mặt khuôn mặt giảm đến 2.500 đỉnh và sau đó dạng chuẩn được tính toán, các hình ảnh kết cấu có thể được ánh xạ trên bề mặt chuẩn. Phép nội suy kết cấu và hình thành một lưới Cartesian có đại diện phẳng cho kết cấu và hình dạng như trong hình 4.



Hình 1.13. Kết cấu phẳng trong lưới Cartesian

Bronstein và các đồng nghiệp sử dụng một cơ sở dữ liệu của khuôn mặt bao gồm 64 trẻ em và 93 người lớn, 115 nam và 42 nữ; cơ sở dữ liệu chứa các hình ảnh của cặp song sinh giống hệt nhau. Các tác giả đã phân tích bốn cách tiếp cận: phân rã eigen của các hình ảnh trong 3D, sự kết hợp của kết cấu và dạng hình ảnh trong eigenfaces, sự phân rã của các hình ảnh chuẩn và các thuật toán eigenforms. Bronstein và các đồng nghiệp kết luận rằng thuật toán của họ có thể trích xuất các đặc trưng hình học của bề mặt khuôn mặt, và việc sử dụng các biểu diễn chuẩn tạo nên sức mạnh trong cách tiếp cận của họ cho nhận dạng các biểu cảm khuôn mặt và biến dạng khuôn mặt.

Golovinskiy và các đồng nghiệp [29] trình bày một cách tiếp cận để mô hình hóa hình học của các đặc trưng nhỏ như nếp nhăn và 28 lỗ chân lông xuất hiện trên bề mặt da. Golovinskiy và các đồng nghiệp đề xuất một mô hình khuôn mặt thống kê có thể trích xuất, chuyên giao và tổng hợp các đặc trưng nhỏ trên khuôn mặt; cách tiếp cận của họ là dựa và phân tích scans ở độ phân giải cao, cung cấp hình học khuôn mặt ở dạng 3D, họ cũng sử dụng phân chia bề mặt để tách các chi tiết trên khuôn mặt. Bước tiếp theo là phân tích tham số để xử lý các biến thể của mô hình thống kê.

Golovinskiy và các đồng nghiệp chỉ ra rằng để khai thác các thông tin thống kê có thể sử dụng các kỹ thuật khác ví dụ như PCA. Về cơ bản hệ thống của họ bao gồm một bước phân tích và một bước tổng hợp, đầu tiên bắt đầu với khuôn mặt quét ở độ phân giải cao 500.000 đa giác, trong bước này các thông tin thống kê thu được; bước tổng hợp bao gồm điều chỉnh hình ảnh để tạo thành một khuôn mặt mới. Các tác giả đã đề cập tới một số ứng dụng trong phương pháp

của họ là các phân tích chi tiết khuôn mặt để nghiên cứu các ảnh hưởng của tuổi tác và giới tính lên các khuôn mặt. Thực nghiệm của họ trên 149 đối tượng tuổi khác nhau, giới tính, chủng tộc khác nhau; độ tuổi là từ 15 – 83 tuổi, 114 nam, 35 nữ tham gia và liên quan đến chủng tộc đã có 81 người da trắng, 63 châu á và 5 người Mỹ gốc Phi. Các gương mặt trung bình của Golovinskiy và các đồng nghiệp được biểu diễn trong hình 1.14 và từ mô hình này có thể tạo ra các khuôn mặt mới với các chi tiết về tuổi tác.



Hình 1.14. Mô hình được phát triển bởi Golovinskiy

Năm 2009, Ramirez cùng các đồng nghiệp [30] đã trình bày một mô hình để tạo ra nhiều biểu cảm từ một khuôn mặt tự nhiên được mô hình hóa với ba lớp lưới để biểu diễn cho da, mô mỡ và hộp sọ. Để tạo các biểu cảm trong 3D, mô hình được bổ sung thêm các cơ quan trọng nhất mà tham gia vào các biểu cảm khuôn mặt. Mô hình đề xuất bởi Ramirez và các đồng nghiệp được chia làm ba bước chính: lọc nhiễu, phát hiện các điểm điều khiển và trình diễn. Các mô hình khuôn mặt được tạo ra gồm 2,676 đa giác và 2,777 đỉnh. Hình 1.15 biểu diễn mô hình khuôn mặt được tạo ra của Ramirez.

Ramirez và các đồng nghiệp đã cho rằng các cơ mặt được xác định như một nhóm các đỉnh, có 18 cơ mặt và họ đã sử dụng hệ mã hóa hành động mặt (FACS) để mô tả và đo các hành vi trên khuôn mặt và thiết lập một cơ chế để phân loại các biểu cảm khuôn mặt. Sau khi xác định các điểm đánh dấu tại khuôn mặt được tạo ra và trong khuôn mặt đầu vào, các tác giả đã tiến hành chuyển đổi mô hình khuôn mặt chung trong mô hình khuôn mặt đầu vào sử dụng các điểm tương ứng. Để theo dõi những thay đổi trên mô hình khuôn mặt chung trong suốt quá trình điều chỉnh giai đoạn này được chia thành 3 bước: điều chỉnh toàn cục, biến dạng trong mặt phẳng xy và biến dạng theo trục z. Sau

đó quá trình điều chỉnh mô hình khuôn mặt chung có thể được sử dụng để tạo ra các biểu cảm khuôn mặt. Chương trình tổng hợp này sẽ chụp lại các biểu cảm mặt, sau đó so sánh với các dữ liệu đào tạo để xác định chúng có giống với thực tế hay không.



Hình 1.15. Mô hình khuôn mặt của Ramirez

Việc mô hình hóa khuôn mặt 3D trong các nghiên cứu trên được ứng dụng để tạo ra các khuôn mặt mới, phân tích biến dạng khuôn mặt, phát hiện khuôn mặt, nhận dạng khuôn mặt, theo dõi tư thế đầu, hoạt hóa khuôn mặt, tạo ra và phân tích các biểu cảm, các khuôn mặt lão hóa và chưa lão hóa cho phép nghiên cứu làm thế nào khuôn mặt thay đổi theo thời gian. Mỗi ứng dụng đưa ra ở trên được sử dụng trong các lĩnh vực như sinh trắc học, pháp y, tạo hình, phẫu thuật, v.v.

1.3.4. Một số phương pháp biểu diễn trạng thái khuôn mặt 3D

Hoạt hóa khuôn mặt trên máy tính có thể được mô tả một cách tương đối như là một tập các kỹ thuật để biểu diễn khuôn mặt đồ họa trên một hệ thống máy tính và sau đó hoạt hóa khuôn mặt này theo một cách nhất quán với các dịch chuyển trên khuôn mặt con người thực. Dưới đây luận văn trình bày bốn phương pháp cơ bản chính để biểu diễn trạng thái khuôn mặt, đó là key-frame animation, hoạt hóa theo tham số, hoạt hóa dựa trên cơ và hoạt hóa dựa trên cơ giả.

a. Hoạt hóa dựa vào Key-frame

Phương pháp này chủ yếu sử dụng các kỹ thuật nội suy để cung cấp một cách tiếp cận trực quan trong quá trình hoạt hóa khuôn mặt. Trong phương pháp này các hàm nội suy được sử dụng để tạo ra các frame giữa các khung hình chính hỗ trợ cho quá trình chuyển đổi giữa các khung hình chính một cách tron

tru. Các khung hình chính ở đây là những gương mặt với các biểu cảm khác nhau.

Kỹ thuật nội suy có thể dựa vào mô hình hình học hoặc mô hình tham số hóa của khuôn mặt 3D. Kỹ thuật nội suy dựa trên hình học trực tiếp điều chỉnh các đỉnh lưới của mô hình mặt. Vị trí của các vertex trong mô hình mặt trung gian được xác định bằng cách nội suy giữa các vị trí của các đỉnh tương ứng trong hai mô hình mặt chính. Các kỹ thuật nội suy trên mô hình mặt 3D tham số hóa xác định từng tham số được sử dụng để tạo ra khuôn mặt trung gian bằng cách nội suy các tham số tương ứng được sử dụng để tạo ra hai khuôn mặt biểu cảm chính. Toàn bộ quá trình nội suy bị chi phối bởi một hệ số nội suy.

Kỹ thuật nội suy có thể sử dụng các hàm nội suy tuyến tính hoặc phi tuyến. Để đơn giản thì có thể sử dụng nội suy tuyến tính. Tuy nhiên, các phép nội suy phi tuyến được sử dụng để bắt trước các chuyển động trên khuôn mặt một cách chính xác hơn. Để tính xấp xỉ các quá trình tăng tốc và giảm tốc giữa các chuyển động của khuôn mặt Parke (1972) đã sử dụng một phương trình nội suy cosin. Nội suy song tuyến cũng được sử dụng để tạo ra các biểu hiện khuôn mặt đa dạng khác nhau hơn so với nội suy tuyến tính.

Phương pháp hoạt hóa dựa vào key-frame có ưu điểm là nhanh chóng và dễ dàng để tạo ra các hoạt ảnh động trên khuôn mặt. Tuy nhiên, phương pháp này không tạo ra được các biểu cảm nằm ngoài giới hạn của tập khung hình chính. Sự kết hợp động của chuyển động mặt độc lập được tạo ra rất khó khăn. Do đó, nội suy là một phương pháp tốt để tạo ra một tập các hoạt ảnh động nhỏ từ một vài khung hình chính.

b. Hoạt hóa theo tham số

Phương pháp hoạt hóa theo tham số được phát triển lần đầu tiên bởi Parke vào năm 1974. Phương pháp này hoạt hóa khuôn mặt với một tập các tham số điều khiển. Các tham số này được nối cố định với các vùng cụ thể trên khuôn mặt. Một chuyển động mặt cụ thể được tạo ra bởi biến dạng của một tham số nối với vùng một vùng cụ thể trên khuôn mặt.

Một mô hình tham số hóa khuôn mặt được thiết kế đặc biệt để thực thi thiết lập các tham số điều khiển khi sử dụng các kỹ thuật hoạt hóa theo tham số. Mô hình khuôn mặt này được phân chia thành nhiều vùng. Các đa giác trong các

vùng này phải được thiết kế để các biến dạng địa phương có thể áp dụng để tạo ra các chuyển động mong muốn trên khuôn mặt.

Các chuyển động từng vùng của khuôn mặt (hay các biến dạng địa phương) có thể được tạo ra từ các phép nội suy địa phương, các biến đổi hình học và các kỹ thuật ánh xạ kết cấu. Các phép nội suy địa phương tạo ra hình dạng trung gian của hai hình dạng được xác định trước của từng vùng cụ thể dựa vào các giá trị tham số. Các biến đổi hình học, chẳng hạn như xoay, tỉ lệ, dịch chuyển vị trí, sử dụng các hàm toán học của tham số để biến đổi hình dạng của một vùng cụ thể trên khuôn mặt thành hình dạng khác. Các kỹ thuật ánh xạ kết cấu bề mặt thao tác trên kết cấu của vùng mặt để tổng hợp các chuyển động trên khuôn mặt. Từ việc biến đổi từng vùng cụ thể trên khuôn mặt chuyển động của khuôn mặt được tạo ra.

Không giống như phương pháp hoạt hóa dựa vào key-frame, hoạt hóa theo tham số cho phép kiểm soát một cách rõ ràng các khu vực cụ thể trên khuôn mặt. Phương pháp này cung cấp một chuỗi các biểu cảm khuôn mặt với chi phí tính toán tương đối thấp. Tuy nhiên, phương pháp hoạt hóa theo tham số ít khi tạo ra các biểu cảm con người một cách tự nhiên khi một xung đột giữa các tham số xảy ra. Do đó, phương pháp này chỉ phù hợp cho biến dạng các vùng cụ thể trên khuôn mặt.

c. Hoạt hóa dựa trên cơ giả

Sự hình thành của một biểu cảm khuôn mặt là một quá trình phức tạp, liên quan đến sự tham gia đồng bộ của một số các cơ. Ngoài ra, nó gây ra các biến dạng da nhìn thấy được, phù hợp với các chuyển động phức tạp của da. Mô hình hóa quá trình này về mặt giải phẫu đòi hỏi phải có khả năng tính toán tốt. Vì lý do này, các nhà khoa học quan sát các kết quả của các biểu cảm khuôn mặt để tìm ra một cách tái tạo lại những kết quả này mà không cần đến một mô hình giải phẫu phức tạp. Một kết quả cho nỗ lực này được gọi là cách tiếp cận dựa trên cơ giả. Hoạt hóa dựa trên cơ giả cố gắng mô phỏng lại các tập con các cơ trong thực tế bằng cách trực tiếp làm biến dạng lưới trên khuôn mặt.

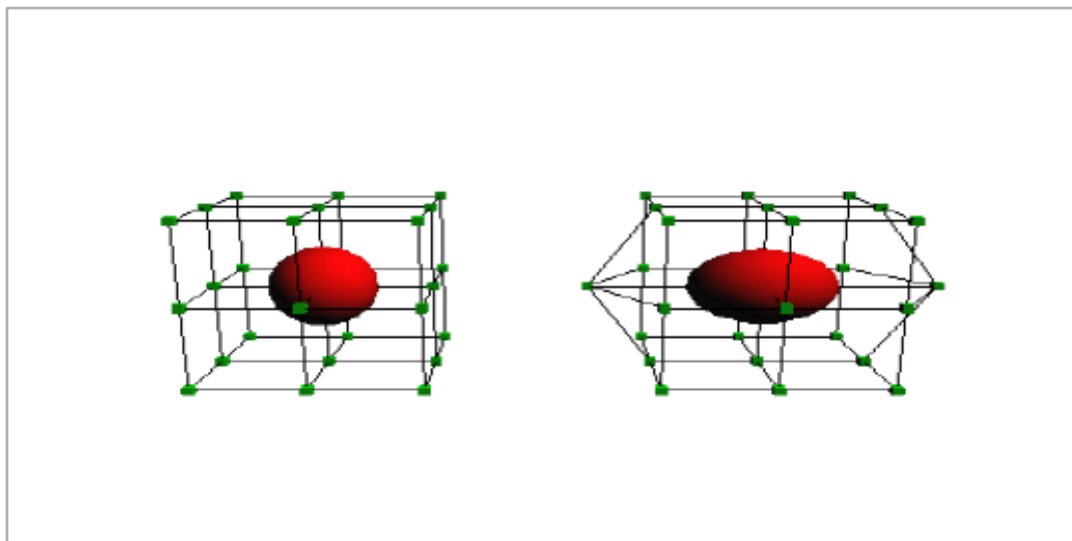
Nó cố gắng để cạnh tranh với tập con các cơ mặt cần thiết bằng cách trực tiếp làm biến dạng lưới trên khuôn mặt. Những kỹ thuật được sử dụng cho phương pháp này bao gồm hành động cơ trừu tượng (AMA – Abstract Muscle Action) và biến dạng tự do.

Năm 1988, Mangnemat – Thalmann cùng các đồng nghiệp đã sử dụng các đơn vị hành động để thiết kế một bộ các thủ tục được sử dụng để tạo ra các biểu cảm trên khuôn mặt. Họ gọi tập các thủ tục này là các hành động cơ trừu tượng. Mỗi thủ tục AMA hoạt động trên một khu vực cụ thể của khuôn mặt, và chúng được xác định khi khuôn mặt được xây dựng, để xấp xỉ gần đúng hành động của một cơ hay một nhóm cơ có liên quan một cách chặt chẽ với nhau. Đối với các cơ phức tạp (như môi), một vài thủ tục, trong đó mỗi thủ tục sẽ chịu trách nhiệm cho một chuyển động đơn giản đã được phân chia. Hệ thống của Magnemat - Thalmann cùng các đồng nghiệp bao gồm khoảng 30 thủ tục AMA. Bằng cách kết hợp các thủ tục này, một số lượng lớn các biểu cảm có thể được tạo ra. Ví dụ các thủ tục AMA là RIGHT_EYELID, MOVE_LEFT_EYE_VERTICAL và RIGHT_LIP_RAISER.

Các tác giả chia các biểu cảm khuôn mặt thành hai nhóm, cụ thể là các âm vị và cảm xúc. Âm vị là một biểu hiện trên khuôn mặt chỉ sử dụng các chuyển động miệng và trực tiếp góp phần vào tiếng nói. Cảm xúc là một biểu hiện trên khuôn mặt có tác dụng trên nhiều bộ phận của khuôn mặt. ví dụ như khóc, mỉm cười, và cười to. Trong khi các khuôn mặt ban đầu có thể được tạo ra sử dụng bất kỳ kỹ thuật mô hình hóa nào, các thủ tục AMA cũng đặt ra các yêu cầu nhất định, chẳng hạn như:

- Khuôn mặt được giả định là đối xứng
- Các thủ tục có thể cần phải được thu nhỏ lại bởi một nhân tố, cùng với mô hình khuôn mặt.
- Khuôn mặt cần được chia thành các khu vực cụ thể (chẳng hạn như da, răng và mí mắt)

Biến dạng tự do (FFD – Freeform Deformations) được giới thiệu lần đầu tiên bởi Sederberg và Parry vào năm 1986. FFD có thể biến dạng một đối tượng bằng cách bao quanh đối tượng đó bằng một cấu trúc mạng tinh thể hình hộp và di chuyển các đỉnh lưới hoặc các điểm điều khiển để làm biến dạng đối tượng. Cấu trúc mạng tinh thể hình hộp đó được gọi là một hộp điều khiển. Một đối tượng được nhúng vào một hộp điều khiển tương tự, rõ ràng và linh hoạt có chứa một lưới 3D các điểm điều khiển. Khi một hộp điều khiển bị phân rã, uốn cong, hoặc xoắn thành các hình dạng tùy ý, thì các đối tượng nhúng trong nó cũng bị biến dạng theo một cách phù hợp.



Hình 1.16. Biến dạng tự do

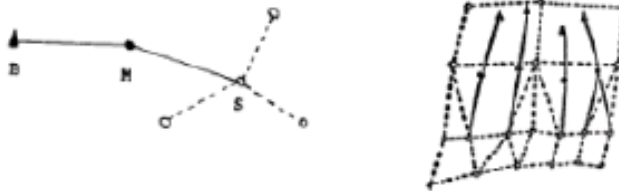
Năm 1989, Chadwick, Haumann và Parent đã sử dụng biến dạng tự do để thiết kế một kỹ thuật dùng để điều khiển các hành động cơ dựa trên việc biến dạng của bề mặt của da. Sau đó vào năm 1991, Kalra cùng các đồng nghiệp đã đề xuất một kỹ thuật là mở rộng của biến dạng tự do, đó là biến dạng xoay tự do. Biến dạng xoay tự do kết hợp các yếu tố trọng số cho mỗi điểm điều khiển và bổ sung thêm mức tự do khác trong các biến dạng cụ thể. Kalra đã kết hợp biến dạng xoay tự do với cách tiếp cận dựa trên vùng để mô phỏng các hiệu ứng hình ảnh của các cơ. Đầu tiên, khuôn mặt được chia thành các vùng tương ứng với mô tả giải phẫu của các hành động cơ. Sau đó, các vùng quan tâm được xác định như là các khối hình hộp dựa trên các vùng bề mặt. Một hành động cơ đạt được bằng cách tương tác di chuyển các điểm điều khiển và thay đổi các trọng số liên kết với mỗi điểm điều khiển. Biến dạng của các điểm biên nằm trong các vùng liền kề được xác định bằng cách nội suy tuyến tính.

Phương pháp hoạt hóa dựa trên cơ giả có ưu điểm là nó đơn giản hơn các phương pháp dựa trên cơ vì nó bỏ qua giải phẫu cơ bản của khuôn mặt. Phương pháp này cũng ít xử lý chuyên sâu hơn các phương pháp dựa trên cơ, do đó máy tính không phải tính toán các hành động cơ và động lực học của cơ mà nó biến dạng lưới một cách trực tiếp. Tuy nhiên, vì là phương pháp này bỏ qua giải phẫu cơ bản của khuôn mặt nên nó không thể mô hình hóa chính xác các nếp nhăn cũng như những chỗ lõm lõm trên khuôn mặt, do đó giảm tính hiện thực của khuôn mặt. Đồng thời phương pháp này cũng không chú ý lắm tới việc giải quyết các tương tác giữa các cơ trên khuôn mặt.

d. Hoạt hóa dựa trên cơ.

Phương pháp này chủ yếu dựa vào giải phẫu hoặc các mô hình nhiều lớp của da và các cơ để làm biến dạng lưới khuôn mặt tạo ra các chuyển động trên khuôn mặt người. Phương pháp tiếp cận dựa trên cơ được đặc trưng bởi việc mô phỏng các cơ và các hành động cơ phù hợp với các cơ trong thực tế của cơ thể con người.

Nghiên cứu đầu tiên trong việc hoạt hóa dựa trên cơ là của Platt và Badler vào năm 1981. Platt và Badler đã xây dựng một mô hình Mass-spring bao gồm ba thành phần là da, các xương và cơ. Da được biểu diễn bởi một tập hợp các điểm 3D di chuyển trong khi xương được biểu diễn bởi một tập hợp các điểm cố định, các cơ là một nhóm các điểm với các cung đàn hồi, cơ được xác định nằm giữa da và xương. Bằng việc áp dụng lực lên các lưới đàn hồi thông qua các cung cơ sẽ tạo ra các biểu cảm trên khuôn mặt người. Ở một cấp độ thấp nhất, phương pháp này bắt đầu với một điểm trong không gian 3D. Cấu trúc đơn giản nhất cho các ứng dụng là một sợi cơ (hình 1.14) nó bao gồm một điểm cơ M, một điểm xương B và một hoặc nhiều điểm da S.



Hình 1.17. Sợi cơ và cơ trong mô hình của Platt và Badler

Khi lực F được áp dụng, đỉnh S được dịch chuyển về phía đỉnh M một khoảng S'

$$S' = \frac{F}{k} \quad (1.1)$$

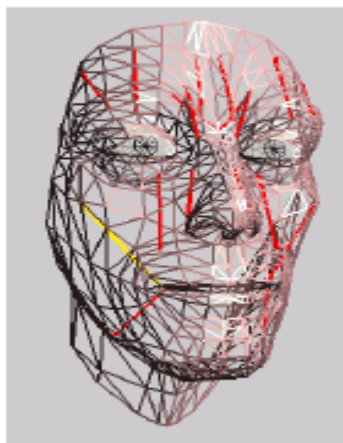
Trong đó, k là tổng các hằng số spring tại điểm đó. Khoảng S' này được tính cho tất cả các đỉnh mà bị tác động bởi lực F .

Một mức trừu tượng cao hơn từ các sợi cơ là một bộ các sợi cơ – các cơ thực tế, như được mô tả trong hình thứ hai bên phải của hình 1.14. Khi một hoạt động cơ được thực hiện, tất cả các sợi cơ trong một cơ đặc biệt được thực hiện song

song. Hạn chế trong phương pháp của Platt và Badler là với các cơ theo sau của xương (ví dụ như khu vực ở ngã ba của hốc mắt và lông mày) sẽ không được mô hình hóa đầy đủ khi có sự thay đổi hướng đột ngột, đồng thời các hành động như hành động của gò má hóp vào hoặc phồng ra, các hành động của xương hàm cũng không được thể hiện trong mô hình.

Sau đó vào năm 1987, Waters đã phát triển một mô hình khuôn mặt bao gồm ba loại cơ khác nhau dựa vào bản chất hành động của chúng, đó là cơ tuyến tính, cơ tấm và cơ vòng. Cũng giống như Platt và Badler, ông sử dụng hệ mã hóa hành động mặt cho giao diện và mô hình mass-spring để mô hình hóa cho da và các cơ. Tuy nhiên các cơ của Water là có hướng và độc lập với cấu trúc xương cho phép chúng có thể được chuyển đến cấu trúc lưới mặt đa dạng.

Cơ tuyến tính và cơ tấm được định nghĩa như là các cơ vector tuyến tính, trong khi cơ vòng được định nghĩa như một tương tác co kéo toàn bộ mặt phẳng xung quanh vào một tâm tưởng tượng. Sự khác nhau giữa cơ tấm và cơ tuyến tính là cơ tuyến tính hướng tới một điểm trung tâm, trong khi cơ tấm hướng tới toàn bộ vùng.



Hình 1.18. Mô hình cơ của Parke và Waters, 1996

Mô hình cơ của Waters là một mô hình rất thành công và đã trở thành một mốc quan trọng và khá tiên tiến trong một thời gian dài, mô hình này tương đối đơn giản theo tiêu chí hiện đại. Nó biểu diễn da như là một bề mặt hình học không có cấu trúc. Các biến dạng được thực hiện bằng cách biến dạng hình học đơn giản của bề mặt mà không biến dạng các mô một cách tinh tế. Terzopoulos và Waters (1990) đã giảm bớt một số các vấn đề được đề cập bằng cách tạo ra các mô hình cơ dựa trên giải phẫu, cùng với mô hình mô dựa trên vật lý. Mô

hình mô này cho phép biến dạng bề mặt một cách thực tế hơn so với các báo cáo trước đó. Zhang, Prakash và Sung (2001) mô hình hóa da bằng các frames lò xo phi tuyến tính mà có thể mô phỏng các động thái của da thật. Ưu điểm của phương pháp này là mô hình không cần phải được coi như là một bề mặt liên tục vì mỗi điểm mass và mỗi cơ spring có thể được truy cập riêng.

Pasquariello và Pelachaud (2001) và Bùi Thế Duy (2004) đã cải tiến điều khiển các hành động cơ bằng cách chia mô hình tương ứng thành một số vùng. Đối với mục đích mô phỏng da, cả Pasquariello và Pelechaud và Bui (2004) đều chuyển từ cách tiếp cận dựa trên vật lý. Mặc dù Pasquariello và Pelechaud hoạt hóa da một cách thực tế, họ không sử dụng mô phỏng vật lý của các cơ và hành vi đàn hồi của da. Hai kỹ thuật thay thế được họ sử dụng để mô phỏng chỗ lồi, lõm và nếp nhăn là Bump mapping và sự dịch chuyển vật lý của các đỉnh.

Ngoài ra, Bui đã tạo ra nếp nhăn bằng cách thay các đỉnh bị ảnh hưởng thành hướng của các hành động cơ. Ông cũng đề cập tới những gì xảy ra trên bề mặt da dưới sự ảnh hưởng của hai hoặc nhiều hơn các cơ vector của Waters. Cách xử lý này của Waters là để thêm các dịch chuyển tuần tự. Bui đề xuất mô phỏng song song bằng cách tính toán kết quả dịch chuyển bên trong sau đó áp dụng nó lên các đỉnh.

Tang, Liew và Yan (2004) đã giới thiệu một hệ thống dựa trên cơ NURBS, được xác định bởi 3-5 điểm điều khiển. Sử dụng hệ thống này, biến dạng cơ có thể đạt được bằng cách điều chỉnh trọng số của các điểm điều khiển này. Ngoài ra, việc thay đổi trọng số sẽ làm cho các nút dịch chuyển dẫn đến việc dịch chuyển lần lượt các đỉnh của mô hình. Để nâng cao tính hiện thực, các tác giả đã cố gắng để mô phỏng các phản ứng biến dạng của mô mỡ bằng cách thêm các điểm điều khiển giữa hai điểm điều khiển kết thúc. Một mô hình giải phẫu đầy hứa hẹn, cung cấp tính linh hoạt độc đáo, cũng đã được mô tả bởi Kahler, Haber và Seidel (2001) và Kahler cùng các đồng nghiệp (2002). Mô hình này được trang bị các cơ và tính toán lưới sọ dựa trên hình học của khuôn mặt, do đó làm giảm đáng kể sự can thiệp thủ công. Một kết quả của kỹ thuật này có thể được nhìn thấy trong hình 1.16 nơi mà mô hình của một cậu bé đã được tự động thích nghi với các lứa tuổi khác nhau của mình. Hoạt hóa này là dựa trên hệ thống Mass-spring.



Hình 1.19. Mô hình cậu bé của Kahler và các đồng nghiệp

Sifakis và các đồng nghiệp (2005) đã xây dựng một mô hình cơ mặt giải phẫu một cách chính xác sử dụng các nguyên tắc bắt nguồn từ các nguyên tắc xây dựng cơ tổng quát hơn của Teran và các đồng nghiệp (2005). Kỹ thuật này dựa trên các thuật toán phần tử hữu hạn. Một đặc trưng quan trọng của mô hình này là hành động cơ của nó có thể tương tác với môi trường, đó là, các lực cơ có thể được kết hợp với các lực bên ngoài như va chạm, tạo ra ảnh hưởng một cách hiệu quả.

Các thuật toán Mass-spring và phần tử hữu hạn dường như là hai công nghệ thống trị cho hoạt hóa dựa trên cơ ngày nay. Hai kỹ thuật này cùng tồn tại và vẫn chưa phân rõ được kỹ thuật nào chiếm được ưu thế nhiều hơn.

CHƯƠNG 2 - MỘT SỐ KỸ THUẬT NỘI SUY TRONG BIỂU DIỄN TRẠNG THÁI KHUÔN MẶT 3D

2.1. Kỹ thuật nội suy tuyến tính.

2.1.1. Giới thiệu

Trong toán học, nội suy tuyến tính là phương pháp đường cong phù hợp sử dụng đa thức tuyến tính để xây dựng các điểm dữ liệu mới trong phạm vi của một tập rời rạc các điểm dữ liệu đã biết.

Trong khoa học kỹ thuật, người ta thường có một số điểm dữ liệu đã biết giá trị bằng cách lấy mẫu thực nghiệm. Những điểm này là giá trị đại diện của một hàm số của một biến số độc lập có một lượng giới hạn các giá trị. Thường khi chúng ta phải nội suy (hoặc ước lượng) giá trị của hàm số này cho một giá trị trung gian của một biến độc lập.

Giả sử chúng ta có hai điểm đã biết tọa độ (x_0, y_0) và (x_1, y_1) thì nội suy tuyến tính chính là đường thẳng giữa hai điểm này. Cho giá trị x trong khoảng (x_0, x_1) thì giá trị y dọc theo đường thẳng được đưa ra từ phương trình

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0} \quad (2.1)$$

$$y = y_0 + (y_1 - y_0) \frac{x - x_0}{x_1 - x_0} \quad (2.2)$$

Hay

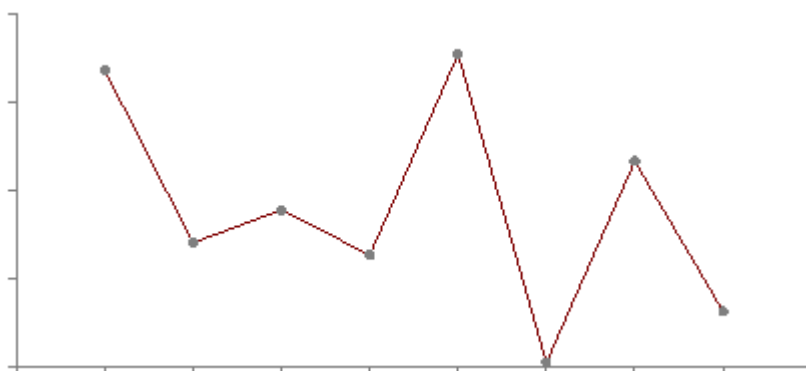
$$y = y_0 \left(1 - \frac{x - x_0}{x_1 - x_0} \right) + y_1 \left(\frac{x - x_0}{x_1 - x_0} \right) \quad (2.3)$$

Đây là công thức nội suy tuyến tính trong khoảng (x_0, x_1) .

Nội suy tuyến tính là phương pháp đơn giản nhất để nhận được các giá trị tại các vị trí giữa các điểm dữ liệu. Các điểm dữ liệu này được kết nối một cách đơn giản bởi một đoạn thẳng. Mỗi đoạn thẳng (giới hạn bởi hai điểm dữ liệu) có thể được nội suy một cách độc lập. Khi đó hàm nội suy giữa hai điểm có thể được viết như sau:

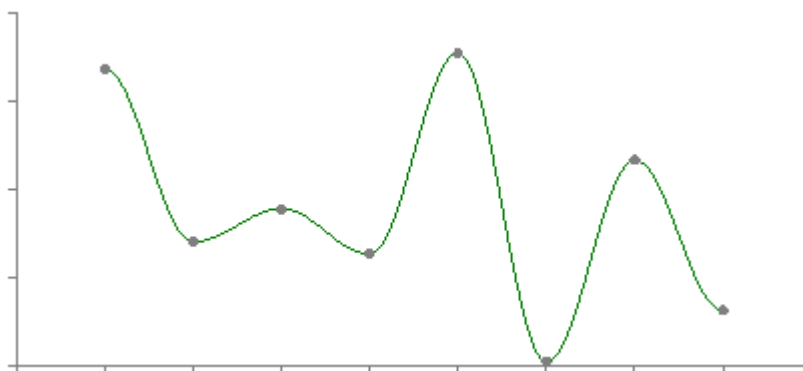
$$y = y_0(1 - t) + y_1 t \quad (2.4)$$

Tham số t xác định nơi để ước tính giá trị trên đường nội suy, nó là 0 tại điểm đầu tiên và là 1 tại điểm thứ hai. Các giá trị nội suy giữa hai điểm t nằm trong khoảng giữa 0 và 1.



Hình 2.1. Nội suy tuyến tính trên một tập dữ liệu

Kết quả của nội suy tuyến tính là không liên tục tại mỗi điểm. Do đó hàm nội suy cosin được đề xuất để cung cấp một chuyển đổi mịn giữa các đoạn trong một tập điểm dữ liệu của hàm nội suy tuyến tính.



Hình 2.2. Nội suy cosin trên một tập dữ liệu

Trong biểu diễn trạng thái khuôn mặt 3D, Kỹ thuật nội suy tuyến tính là phương pháp đơn giản nhất được sử dụng để hoạt hóa trạng thái khuôn mặt 3D. Kỹ thuật này đã được Parke nghiên cứu và sử dụng lần đầu tiên vào năm 1972 [6] với mô hình khuôn mặt được biểu diễn bởi 250 đa giác xác định bởi 400 đỉnh, sau đó kỹ thuật này được sử dụng nhiều trong các nghiên cứu về hoạt hóa khuôn mặt và hoạt hóa khuôn mặt với dữ liệu thực [7], [8]. Ngày nay kỹ thuật nội suy tuyến tính vẫn được sử dụng rộng rãi trong các gói phần mềm hoạt hóa thương mại như MAYA và 3D Studio max trong việc tạo ra các trạng thái biểu cảm khuôn mặt thực tế.

2.1.2. Sử dụng trong biểu diễn trạng thái khuôn mặt 3D

Giả sử rằng khuôn mặt 3D được biểu diễn bởi một đa giác thích hợp. Tư tưởng của kỹ thuật này tương tự như phương pháp tạo ra chuyển động của các phim hoạt hình thông thường. Các phim hoạt hình xác định chuyển động mong muốn bất kỳ bằng cách tạo ra một loạt các bản vẽ các chuyển động chính, sau đó người tạo chuyển động sẽ tạo ra một loạt các khung hình (frame) trung gian giữa các chuyển động của các bản vẽ chính. Đối với việc hoạt hóa trạng thái khuôn mặt trên máy tính thì các bản vẽ chính ở đây được thay thế bởi các tập tin dữ liệu mô tả các trạng thái khuôn mặt khác nhau. Các trạng thái khuôn mặt khác nhau này bao gồm các trạng thái như trạng thái khuôn mặt tự nhiên, trạng thái cười, trạng thái tức giận, ngạc nhiên, .v.v.

Các dữ liệu trong các tập tin mô tả trạng thái khuôn mặt là các vị trí ba chiều của các điểm xác định đa giác sử dụng để đại diện cho khuôn mặt hoặc vị trí ba chiều của các điểm tại các vùng đặt trung cơ bản nhất của khuôn mặt mà ở đó trạng thái khuôn mặt được thể hiện.

Các khung hình trung gian được xác định bằng cách sử dụng phép nội suy tuyến tính.

$$P_{interpolated}(t) = (1-t) * P_{prev} + t * P_{next} \quad (2.5)$$

$$0 \leq t \leq 1$$

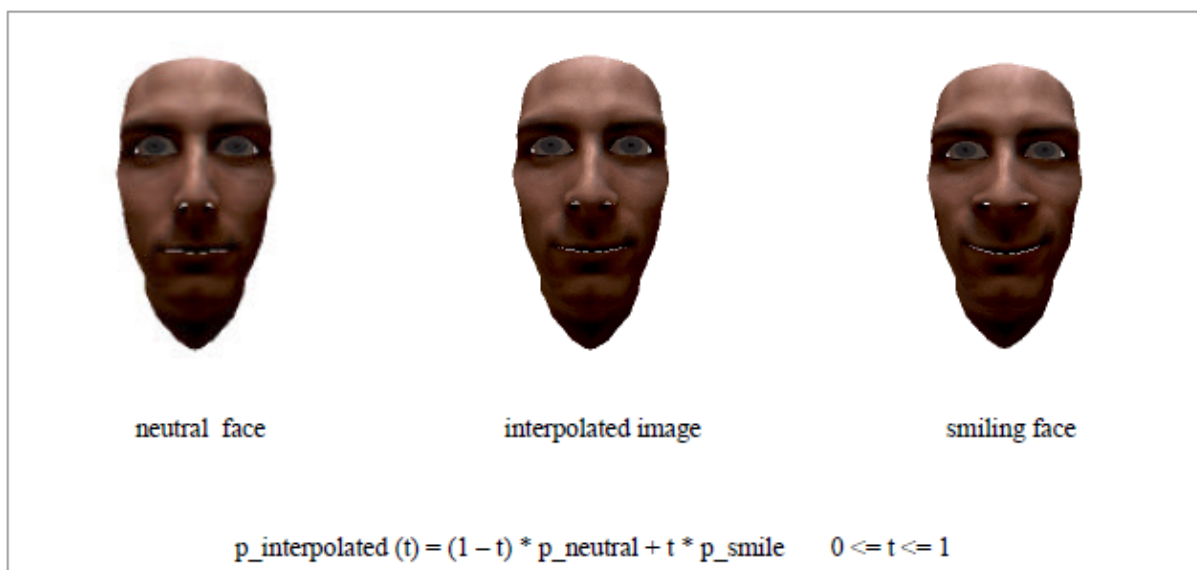
Trong đó:

$P_{interpolated}$: vị trí các điểm trong frame trung gian

P_{prev} : vị trí các điểm trong frame trạng thái khuôn mặt nguồn.

P_{next} : vị trí các điểm trong frame trạng thái khuôn mặt đích

t : Tham số nội suy ($0 \leq t \leq 1$)



Hình 2.3. Nội suy tuyến tính khuôn mặt từ trạng thái tự nhiên sang trạng thái cười

Tuy nhiên, khuôn mặt luôn được chi phối bởi các quy luật vật lý, chuyển động của nó không phải là tuyến tính mà có xu hướng tăng tốc và giảm tốc. Do đó, một hàm nội suy Cosin hoặc các biến thể khác có thể được sử dụng để cung cấp khả năng tăng tốc và giảm tốc độ thể hiện trạng thái khi bắt đầu và kết thúc một hoạt cảnh.

Thực hiện:

Đầu vào: Mô hình khuôn mặt 3D ở trạng thái nguồn và trạng thái đích.

Đầu ra: Các trạng thái khuôn mặt 3D thể hiện sự biến đổi từ trạng thái ban đầu sang trạng thái khuôn mặt đích.

Thuật toán:

Bước 1: Khởi tạo mô hình 3D khuôn mặt ở trạng thái nguồn và trạng thái đích.

Bước 2: Thực hiện nội suy ra trạng thái khuôn mặt trung gian bằng cách sử dụng phép nội suy tuyến tính.

Bước 3: Kết thúc

Kỹ thuật nội suy tuyến tính có thể được sử dụng để biểu diễn trạng thái khuôn mặt 3D một cách độc lập như được trình bày ở trên hoặc cũng có thể được kết hợp với các kỹ thuật khác như kết hợp với kỹ thuật Morphing, hệ mã hóa hành động mặt,..v.v. để tạo ra các chuyển động mặt tự nhiên và thực tế hơn.

❖ **Kết hợp giữa kỹ thuật nội suy tuyến tính và kỹ thuật Morphing:**

Morphing trạng thái khuôn mặt 3D nghĩa là tạo ra các chuyển tiếp liên tục và thực tế giữa các biểu cảm khác nhau của khuôn mặt. Phương pháp này đạt được các hiệu ứng bằng cách morphing giữa các mô hình mặt tương ứng. Trình tự morphing 3D có thể đạt được bằng cách sử dụng kỹ thuật nội suy tuyến tính đơn giản giữa các tọa độ hình học của các đỉnh tương ứng trong mỗi lưới của hai lưới khuôn mặt. Cùng với quá trình nội suy hình học thì các texture của mô hình khuôn mặt cũng cần được blend từ các textures có liên quan. Khi morphing hai biểu cảm khác nhau của cùng một mô hình mặt thì mô hình khuôn mặt trung gian được tạo ra bởi phép nội suy hình học. Sau đó, texture của mô hình khuôn mặt trung gian này trực tiếp được tạo ra từ khuôn mặt tự nhiên bằng cách thiết lập tương ứng giữa mô đa giác của mô hình khuôn mặt tự nhiên và mô hình biểu cảm trung gian cho mỗi dòng quét với mỗi pixel. Với morphing giữa hai trạng thái biểu cảm của bất kỳ hai mô hình khuôn mặt nào thì texture trung gian được tạo ra bằng cách sử dụng nội suy tuyến tính của nguồn tương ứng và tam giác đích cho từng dòng quét. Thuật toán morphing được đề xuất như sau:

Bước 1: Với mỗi giá trị của hệ số nội suy t lặp lại

Từ bước 2 tới bước 5 ($0 \leq t \leq 1$)

Bước 2: Với mỗi tam giác của mô hình lặp lại từ bước 3 tới bước 5 để tạo ra khung hình trung gian.

Bước 3: Sắp xếp tam giác dựa trên Y_{min} của chúng.

Bước 4: Tạo tam giác trung gian của mô hình nguồn và mô hình đích.

$$Imtriangle = (1-t)*source\ triangle + t*target\ Triangle$$

Bước 5: Với mỗi dòng quét từ Y_{min} tới Y_{max} lặp lại bước 4

$$Image(x, y) = (1 - t)*Image1(x1, y1) + t*Image2(x2, y2)$$

Bước 5 tạo ra kết cấu cho mô hình trung gian từ hình ảnh nguồn và hình ảnh đích. Đối với mỗi dòng quét của tam giác trung gian, các hệ số y tăng như sau:

$$Yinc(Imtriangle) = 1$$

$$Yinc(source\ triangle) = (ysmax - ysmmin) / (yimax - yimin)$$

$$Yinc(target\ triangle) = (ytmax - ytmin) / (yimax - yimin)$$

$ysmax$ và $ysmin$ là giá trị y lớn nhất và nhỏ nhất của tam giác nguồn, $yimax$ và $yimin$ là giá trị y lớn nhất và nhỏ nhất của tam giác trung gian và $ytmax$, $ytmin$ là giá trị y lớn nhất và nhỏ nhất của tam giác đích.

Với mỗi dòng quét, các hệ số x tăng như sau:

$$Xinc (ImtriangleC) = 1$$

$$Xinc (source triangle) = (xstart2 - xstart1) / swidth$$

$swidth = xend - xbegin$ của tam giác trung gian và $xstart1$, $xstart2$ là giá trị x bắt đầu của tam giác nguồn tương ứng với dòng quét hiện tại.

$$Xinc (target triangle) = (xstart4 - xstart3) / swidth.$$

$xstart3$, $xstart4$ là giá trị x bắt đầu của tam giác đích tương ứng với dòng quét hiện tại.

*Giá trị x bắt đầu tương ứng với dòng quét tiếp theo = giá trị x bắt đầu trước đó + $(1/slope) * yinc$.*

❖ **Kết hợp giữa kỹ thuật nội suy tuyến tính và hệ mã hóa hành động mặt:**

Hệ mã hóa hành động mặt (FACS – Facial Action Coding System) được giới thiệu bởi Ekman và Friesen năm 1976 để đánh giá và mô tả các hành động mặt bằng cách kiểm tra tất cả các cơ mặt. Hiện nay, nó được coi là một chuẩn để hiển thị sự xuất hiện khuôn mặt kích thích bởi những thay đổi trên từng cơ mặt. Nó xuất phát từ việc phân tích giải phẫu khuôn mặt bằng cách mô tả các hoạt động cơ mặt của con người. Nguyên tắc làm việc của FACS trong việc tìm hiểu hành vi của khuôn mặt được dựa trên các hành động mặt. FACS chia khuôn mặt con người thành 46 đơn vị hành động (AUS – Action Units). Mỗi đơn vị hiện thân của một hành động cơ riêng biệt hoặc một nhóm các cơ đặc trưng cho một vị trí duy nhất trên khuôn mặt. Các đơn vị hành động mặt được xây dựng sao cho phù hợp với các hành động nơi mà mỗi đơn vị hành động mặt có thể liên quan đến nhiều cơ.

Nguyên tắc là mỗi AU là một đơn vị nhỏ nhất mà không thể được giảm thành các đơn vị hành động nhỏ hơn. Bằng cách sắp xếp (phân loại) các cơ khác nhau trên khuôn mặt FACS có thể bắt trước tất cả các di chuyển cơ mặt. Biểu diễn các AUS khác nhau và tên FACS tương ứng được liệt kê trong bảng 2.1

Bảng 2.1: Biểu diễn hành động và các đơn vị của chúng.

AU	FACS name	AU	FACS name	AU	FACS name
6	Check raiser	23	Lip tightener	15	Lip corner depressor
1	Inner brow raiser	26	Jaw drop	5	Uppwe lid raiser
17	Raiser chin	20	Lip stretcher	9	Nose wrinkle
4	Brow lower	14	Dimpler	16	Lower lip depressor
10	Upper lip raiser	2	Outer brow raiser	12	Lid corner puller

Các đơn vị hành động khác nhau có thể được kết hợp để tạo ra một biểu cảm mặt được biến đổi. Ví dụ, sự kết hợp của AU4 (Brow Lower – lông mày hạ thấp xuống), AU15 (Lip Corner Depressor – góc môi kéo xuống), AU1 (Inner Brow Raiser – lông mày bên trong cao lên) và AU23 (Lip Tightener – căng môi) tạo ra biểu cảm buồn. FACS hỗ trợ quá trình hoạt hóa để minh họa và xây dựng các biểu cảm mặt thực tế qua tất cả các chi tiết có thể có của các đơn vị hoạt hóa mặt đồng thời nó là thước đo về chất lượng trong việc tạo ra các nét mặt thực tế, phản ánh chính xác các cử động cơ mặt một cách đáng tin cậy. Hơn nữa kỹ thuật nội suy tuyến tính là một phương pháp hiệu quả để tạo ra các biểu cảm tự nhiên của con người. Do đó, Mohamad và các đồng nghiệp [7] đã đề xuất một kỹ thuật mới bằng cách kết hợp giữa FACS và nội suy tuyến tính để đạt được nhiều biểu cảm mặt của con người một cách chính xác và hiệu quả.

Thuật toán tạo ra 4 trạng thái biểu cảm giận giữ, hạnh phúc, buồn và sợ hãi:

Bước 1. Thêm bốn các giá trị trạng thái biểu cảm, $E = \{\text{Happy, Anger, Sad, Fear}\}$ dựa vào FACS.

Bước 2. Tạo một bản đồ kiểm soát miền với bốn cực và điểm ở giữa (0,0)

Bước 3. Vẽ mô hình khuôn mặt tự nhiên N và đặt các giá trị của nó ở giữa bản đồ kiểm soát miền dựa vào nội suy hình dạng.

Bước 4. Thiết lập các giá trị biểu cảm tuyệt đối E_i ở mỗi cực tới bản đồ kiểm soát miền.

Bước 5. Người sử dụng điều khiển vị trí con trỏ (C_x, C_y) đến trạng thái biểu cảm mong muốn.

Bước 6. Đặt trọng số w của mỗi trạng thái biểu cảm theo vị trí của con trỏ chuột:

if $C_x \geq 0$ then

$wE_3 = x$

$wE_4 = 0$

end if

if $C_x \leq 0$ then

$wE_3 = 0$

$wE_4 = |x|$

end if

if $C_y \geq 0$ then

$wE_1 = y$

$wE_2 = 0$

end if

if $C_y \leq 0$ then

$wE_1 = 0$

$wE_2 = |y|$

end if

Bước 7. Tính toán giá trị nội suy tuyến tính I bằng cách dịch chuyển các giá trị tự nhiên tới vị trí trọng số biểu cảm α

$$I = \alpha (\text{value1}) + (1.0 - \alpha) (\text{value2})$$

$$0.0 < \alpha < 1.0$$

với *value1* và *value2* là các biểu cảm có các trọng số tích cực theo vị trí của con trỏ

Bước 8. Vẽ mô hình biểu cảm mới tương ứng với giá trị của I

2.1.3. Nhận xét

Kỹ thuật nội suy tuyến tính có một số ưu điểm và nhược điểm trong việc hoạt hóa trạng thái khuôn mặt như sau:

Ưu điểm:

- Đơn giản để hiểu biết và thực thi.
- Tốc độ hoạt hóa nhanh
- Không yêu cầu tính toán chuyên sâu.
- Dễ dàng tạo ra các hoạt ảnh trên khuôn mặt với kết quả được dự đoán là cao

Nhược điểm:

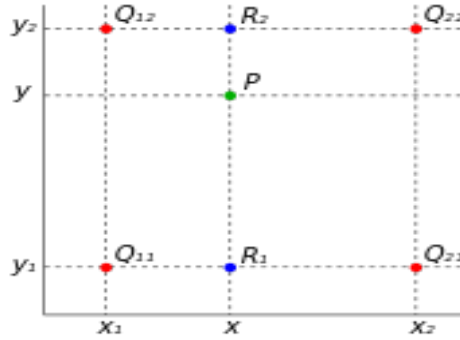
- Bị giới hạn bởi một số lượng các biểu cảm chính nhất định.
- Nếu sự chuyển đổi giữa hai key-frame trạng thái khuôn mặt là không tuyến tính cũng không phù hợp với bất kỳ hàm nào đang tồn tại thì sẽ khó khăn trong việc tính toán vị trí của các trạng thái khuôn mặt trung gian.
- Khó khăn trong việc tạo ra sự kết hợp động của các chuyển động mặt độc lập
- Không hiệu quả đối với các nét mặt có sự biến đổi phức tạp

2.2. Kỹ thuật nội suy song tuyến.

2.2.1. Giới thiệu

Trong toán học, nội suy song tuyến (Bilinear interpolation) là mở rộng của phép nội suy tuyến tính cho các hàm nội suy hai biến (ví dụ: x và y) trong một lưới phẳng 2D. Ý tưởng của phương pháp này là nội suy tuyến tính theo hướng đầu tiên, và sau đó nội suy tuyến tính một lần nữa theo hướng khác. Hay nói cách khác, nội suy song tuyến tính là phương pháp được sử dụng để tính trọng số khoảng cách trung bình của bốn điểm gần nhất để ước lượng giá trị điểm mới được tạo ra.

Giả sử rằng chúng ta muốn tìm giá trị của hàm f chưa biết tại điểm (x, y) . Giả sử chúng ta đã biết giá trị của f tại bốn điểm $Q_{11} = (x_1, y_1)$, $Q_{12} = (x_1, y_2)$, $Q_{21} = (x_2, y_1)$ và $Q_{22} = (x_2, y_2)$.



Hình 2.4. Minh họa nội suy song tuyến tính

Đầu tiên chúng ta thực hiện nội suy tuyến tính theo trục x.

$$f(x, y_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}) \quad (2.6)$$

$$f(x, y_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22})$$

Chúng ta tiếp tục nội suy theo trục y để đạt được kết quả như sau:

$$f(x, y) \approx \frac{y_2 - y}{y_2 - y_1} f(x, y_1) + \frac{y - y_1}{y_2 - y_1} f(x, y_2) \quad (2.7)$$

Hay

$$f(x, y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} [x_2 - x \quad x - x_1] \begin{bmatrix} f(Q_{11}) & f(Q_{12}) \\ f(Q_{21}) & f(Q_{22}) \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \quad (2.8)$$

Nếu chúng ta chọn một hệ tọa độ mà trong đó bốn điểm có giá trị f là $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ thì công thức nội suy được đơn giản hóa thành:

$$f(x, y) \approx [1 - x \quad x] \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1 - y \\ y \end{bmatrix} \quad (2.9)$$

2.2.2. Sử dụng trong biểu diễn trạng thái khuôn mặt 3D

Kỹ thuật nội suy song tuyến tính được áp dụng trong việc hoạt hóa khuôn mặt và biểu diễn trạng thái biểu cảm khuôn mặt trong thực tại ảo. Quá trình biểu diễn sự thay đổi giữa hai trạng thái biểu cảm bất kỳ trên khuôn mặt 3D được thực hiện bằng cách nội suy các vị trí của các điểm điều khiển trên khuôn mặt tại các khung hình trung gian giữa hai khung hình biểu diễn hai trạng thái biểu cảm khuôn mặt ban đầu và trạng thái biểu cảm khuôn mặt đích.

Tư tưởng của phương pháp nội suy song tuyến là sử dụng bốn khung hình nội suy thay vì sử dụng hai khung hình trong phép nội suy tuyến tính. Các khung hình (hay keyframe) này đạt được bằng cách kết hợp hai dữ liệu hình dạng của khuôn mặt và hai dữ liệu biểu cảm trên khuôn mặt. Mỗi dữ liệu cho mỗi keyframe được cập nhật theo các bước sau:

Bước 1. Chuẩn bị dữ liệu về hình dáng khuôn mặt 3D của trạng thái biểu cảm khuôn mặt nguồn và trạng thái biểu cảm khuôn mặt đích. Mỗi mốc biểu cảm đạt được bằng cách tương tác xác định các vector cho các điểm điều khiển.

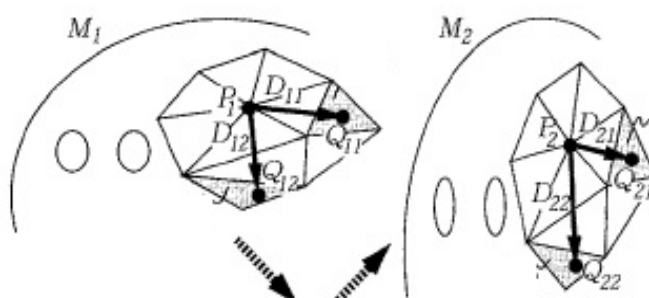
Bước 2. Hai dữ liệu biểu cảm được lựa chọn từ các dữ liệu đã chuẩn bao gồm hai mô hình chứa hai trạng thái biểu cảm khác nhau. Và bốn keyframes được khởi tạo bằng cách kết hợp dữ liệu của hai mô hình hai biểu cảm.

Bước 3. Bốn keyframes được nội suy với các tham số hiện tại cho việc điều khiển những thay đổi về biểu cảm khuôn mặt.

Bước 4. Mô hình biểu cảm mới được lựa chọn với keyframe mới được cập nhật.

Bước 5. Lặp lại bước 2

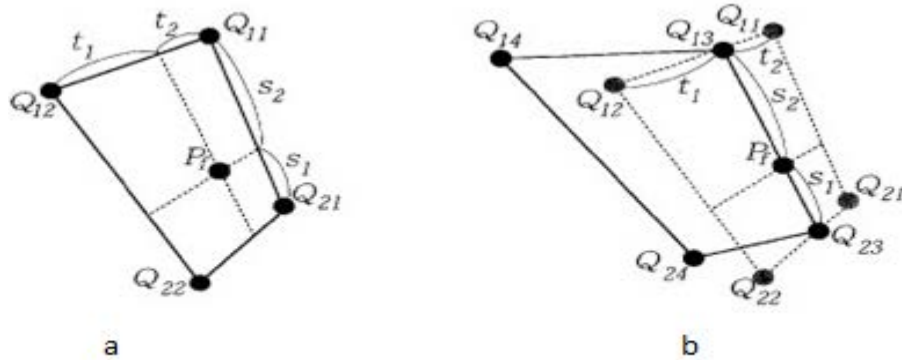
Giả sử, chúng ta biểu diễn sự thay đổi trạng thái biểu cảm từ trạng thái M_1 tới trạng thái trên mô hình M_2 thì chúng ta có thông tin về hai dữ liệu hình dạng M_1, M_2 và hai dữ liệu biểu cảm D_1, D_2 (Hình 2.5).



Hình 2.5. Dữ liệu hình dạng hai trạng thái khuôn mặt

Bằng cách kết hợp các dữ liệu này, bốn dữ liệu được tạo ra như các keyframes. $Q_{i,j}$ ($i = 1,2; j = 1,2$) là vị trí thay thế của một điểm điều khiển P tại các keyframes. Mỗi Q_{ij} là thành phần của keyframe đạt được bằng cách kết hợp i

dữ liệu hình dạng và j dữ liệu biểu cảm. Cho P_f là vị trí thay thế của điểm điểm tương ứng với điểm điều khiển P tại một frame cụ thể của quá trình hoạt hóa.



Hình 2.6. Nội suy song tuyến tính

Như được biểu diễn trong hình 2.6a, P_f đạt được bằng cách nội suy tuyến tính dựa vào Q_{11} , Q_{12} , Q_{21} và Q_{22} như sau:

$$P_f = s_1(t_1Q_{11} + t_2Q_{12}) + s_2(t_1Q_{21} + t_2Q_{22}) \quad (2.10)$$

$$(s_1 + s_2 = t_1 + t_2 = 1)$$

Với s_i ($i = 1, 2$) và t_j ($j = 1, 2$) là các tham số nội suy.

❖ Cập nhật các keyframe cho dữ liệu biểu cảm mới.

Hình 2.6b biểu diễn mô tả các thành phần của các keyframe khi dữ liệu biểu cảm mới được lựa chọn. Khi đó, vị trí của P_f được xác định bởi công thức 2.10. các thành phần Q_{13} , Q_{14} , Q_{23} , và Q_{24} đạt được theo các bước sau:

Bước 1. Đạt được Q_{13} và Q_{23} bằng cách gán các giá trị hiện tại của t_1 và t_2 vào công thức sau:

$$Q_{13} = t_1 Q_{11} + t_2 Q_{12}, \quad (2.11)$$

$$Q_{23} = t_1 Q_{21} + t_2 Q_{22}$$

Q_{13} và Q_{23} là hai trong bốn thành phần của keyframes mới cho biểu cảm mới.

Bước 2. Cho Q_{14} và Q_{24} là vị trí thay thế đạt được bằng cách kết hợp dữ liệu biểu cảm mới với mô hình biểu cảm ban đầu và mô hình biểu cảm thứ hai. Q_{14} và Q_{24} là hai thành phần của các keyframe mới.

Bước 3. Với việc gán giá trị 0 và 1 tới t_1 và t_2 khi đó chúng ta lại tiếp tục tính được giá trị của P_f theo công thức sau :

$$P_f = s_1(t_1Q_{13} + t_2Q_{14}) + s_2(t_1Q_{23} + t_2Q_{24}) \quad (2.12)$$

$$(s_1 + s_2 = t_1 + t_2 = 1)$$

2.2.3. Nhận xét

Ưu điểm :

- Nhanh, dễ dàng tạo ra các hoạt ảnh trên khuôn mặt
- Do sử dụng bốn khung hình chính có liên quan chứ không phải hai, nên kỹ thuật nội suy song tuyến tạo ra các biểu hiện khuôn mặt đa dạng khác nhau hơn so với nội suy tuyến tính. Nội suy song tuyến, khi kết hợp với mô phỏng morphing hình ảnh sẽ tạo ra một loạt các thay đổi biểu cảm khuôn mặt thực tế [9].

Nhược điểm :

Cũng giống như các kỹ thuật nội suy khác, kỹ thuật nội suy song tuyến cũng có các nhược điểm như :

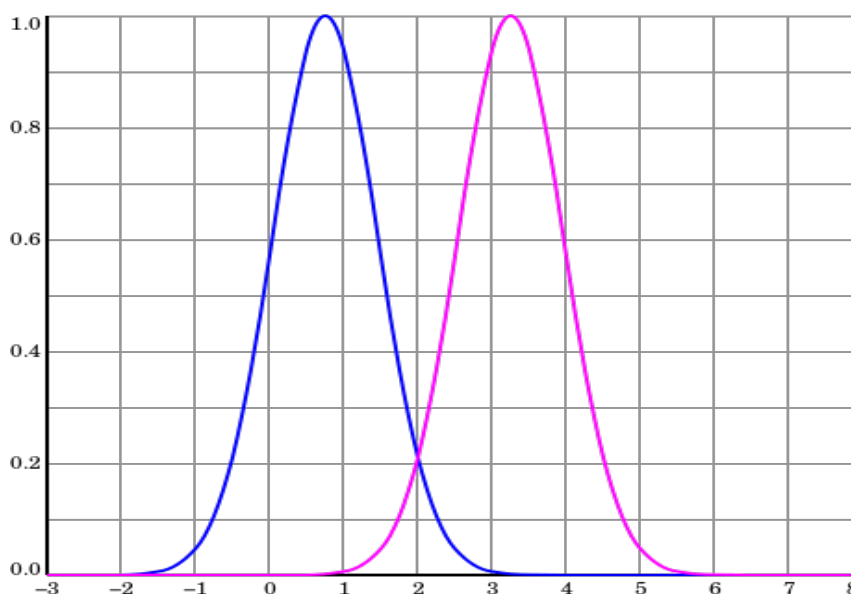
- Khó khăn trong việc tạo ra sự kết hợp động của các chuyển động mặt độc lập
- Không hiệu quả đối với các nét mặt có sự biến đổi phức tạp
- Chỉ phù hợp trong các bài toán nhỏ, sử dụng một vài khung hình cơ bản để sinh ra một tập ảnh nhỏ.

2.3. Kỹ thuật nội suy dựa trên hàm cơ sở bán kính

2.3.1. Giới thiệu

Hàm cơ sở bán kính (RBF) được xây dựng lần đầu tiên bởi R.L. Hardy dựa trên việc đặt câu hỏi, " Cho một tập các dữ liệu nằm rải rác thưa thớt $\{f_j\}_{j=1}^N$ tại các vị trí nút $\{x_j\}_{j=1}^N \subset R^d$, có thể sử dụng một phép nội suy để tính toán lại đầy đủ các điểm chưa biết trên bề mặt?" [14]. Nó lần đầu tiên được đề cập bởi Mairhuber [15], trong trường hợp nhiều hơn một chiều, nội suy không được đề

cập đến khi sử dụng mở rộng của hàm cơ sở $\{\phi_j(x)\}_{j=1}^N \in R^{d \geq 2}$ độc lập tại các vị trí nút. Do đó, có tồn tại một số lượng vô hạn các nút điều đó sẽ tạo nên một vấn đề về nội suy một biến. Hardy bỏ qua vấn đề này bằng cách xây dựng hàm nội suy một biến từ một hàm cơ sở bao gồm sự dịch chuyển của một hàm đối xứng xuyên tâm đơn với một tâm dữ liệu tại mỗi vị trí. Hàm nội suy một biến và các dẫn xuất của nó cho bất kỳ tập hợp các nút phân tán riêng biệt nào trong bất kỳ chiều nào sẽ được đảm bảo nếu loại bỏ phép chiếu trực giao. Mặc dù không có điều kiện về vấn đề của điểm kỳ dị trong phép nội suy, nhưng nó có thể được dự báo sớm trong một số trường hợp đặc biệt [16,17], chẳng hạn như năm 1986 thì multiquadric (MQ) RBFs được đảm bảo là không có điểm kỳ dị [18] đã làm tăng sự phát triển và sự chấp nhận của RBFs. Công trình tiên phong của M.J.D. Powell và các cộng sự tại Đại học Cambridge cũng đóng một vai trò quan trọng trong lịch sử ban đầu của RBFs [19]. Một minh họa cho RBF (xem hình 2.7).



Hình 2.7. Hàm cơ sở bán kính (RBF)

Từng phần RBFs mịn đặc trưng cho một bước nhảy trong một đạo hàm nào đó và do đó có thể dẫn đến hội tụ. Ví dụ, hình cầu với bán kính $|r|^3$, trong đó $r = \|x - x_j\|$ là chuẩn Öclit, có một bước nhảy trong đạo hàm bậc ba tại $x = x_j$, dẫn đến hội tụ, với thứ tự hội tụ ngày càng tăng như là tăng số chiều. Mặt khác, việc nội suy với nhiều RBFs mịn, như $\sqrt{1 + (\epsilon r)^2}$, $\exp(-(\epsilon r)^2)$, và $1/(1+(\epsilon r)^2)$ sẽ dẫn đến quang phổ hội tụ [20, 21]. Trong trường hợp này cần chú ý rằng độ mịn của RBFs phụ thuộc và tham số ϵ . Nó được đưa ra lần đầu tiên bởi Driscoll

và Fornberg [22], trong giới hạn của $\varepsilon \rightarrow 0$ (tức là RBFs phẳng) kỹ thuật RBF tái tạo kỹ thuật pseudo-spectral nếu các nút được đặt phù hợp (ví dụ như các nút cách đều cho phương pháp Fourier, đối với phương pháp Chebyshev thì các nút là Gauss- Chebyshev). Tương tự, Fornberg và Piret [23] cho thấy, trên bề mặt của một quả cầu, trong giới hạn của $\varepsilon \rightarrow 0$, RBFs tạo ra sóng hình cầu theo nghĩa là chúng kéo dài một không gian tương đương với bất kỳ tập nút rời rạc nào.

❖ Hàm cơ sở bán kính (Radial Basic Function - RBF):

Giả sử rằng giá trị của một hàm vô hướng $F: R^3 \rightarrow R$ là đại diện cho sự biến đổi của n điểm rời rạc khác biệt x_i trong không gian R^3 . Khi đó hàm cơ sở bán kính (RBF) cung cấp một phương thức cho việc tạo ra một phép nội suy tron cho hàm F trong không gian R^3 . Hàm này được viết dưới dạng tổng của n xấp xỉ bởi một hàm cơ sở bán kính $g(r_i): [0, rMax] \rightarrow [0,1]$ với r_i là khoảng cách giữa các điểm $p = (x, y, z)$ được xấp xỉ và các điểm $p_i = (x_i, y_i, z_i)$

$$F(p) = \sum_{i=1}^n a_i g(\|\mathbf{x} - \mathbf{x}_i\|) + c_0 + c_1 x + c_2 y + c_3 z \quad (2.13)$$

$$\text{với } \mathbf{x} = (x, y, z)$$

Trong đó:

a_i : là các hệ số vô hướng.

c_0 đến c_3 : là hệ số của đa thức bậc một, các hệ số này mô tả một phép biến đổi affine mà không được thực hiện bởi hàm cơ sở bán kính.

Công thức (2.13) cho thấy rằng: Với n điểm, để biết được các giá trị $F(x_i, y_i, z_i) = F_i$ chúng ta cần phải mở rộng thêm bốn giá trị $n = n + 4$ do có thêm bốn tham số c_0, c_1, c_2, c_3 là các hệ số của đa thức, khi đó phương trình (2.13) được viết dưới dạng ma trận:

$$\mathbf{F} = \mathbf{GA} \quad (2.14)$$

Trong đó:

$$\mathbf{F} = (F_1, F_2, \dots, F_n, 0, 0, 0, 0)^T$$

$$\mathbf{A} = (a_1, a_2, \dots, a_n, c_0, c_1, c_2, c_3)^T$$

\mathbf{G} là một ma trận cấp $(n + 4) \times (n + 4)$:

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdot & \cdot & \cdot & g_{1n} & 1 & x_1 & y_1 & z_1 \\ g_{21} & g_{22} & \cdot & \cdot & \cdot & g_{2n} & 1 & x_2 & y_2 & z_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ g_{n1} & g_{n2} & \cdot & \cdot & \cdot & g_{nn} & 1 & x_n & y_n & z_n \\ 1 & 1 & \cdot & \cdot & \cdot & 1 & 0 & 0 & 0 & 0 \\ x_1 & x_2 & \cdot & \cdot & \cdot & x_n & 0 & 0 & 0 & 0 \\ y_1 & y_2 & \cdot & \cdot & \cdot & y_n & 0 & 0 & 0 & 0 \\ z_1 & z_2 & \cdot & \cdot & \cdot & z_n & 0 & 0 & 0 & 0 \end{bmatrix}$$

Với: $g_{ij} = g(\|x_i - x_j\|)$

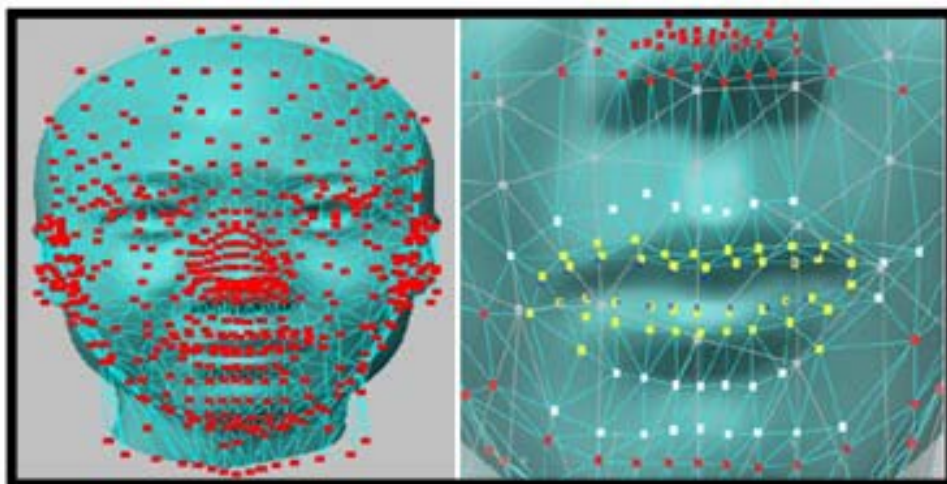
Chúng ta có thể lựa chọn các hàm cơ sở bán kính g_{ij} là một số hàm phổ biến sau: x – biharmonic (\mathbb{R}^3), x^{2n+1} – polyharmonic (\mathbb{R}^3) và $\sqrt{x^2 + c^2}$ (với c là một hằng số) - multiquadric (\mathbb{R}^3):

$$g(x) = \begin{bmatrix} x \\ x^{2n+1} \\ \sqrt{x^2 + c^2} \end{bmatrix}$$

Giải hệ phương trình (2.14) chúng ta có hệ số của \mathbf{A} chính là các hệ số a_i của phương trình (2.13).

2.3.2. Sử dụng trong biểu diễn trạng thái khuôn mặt 3D

Biểu diễn trạng thái khuôn mặt 3D dựa vào hàm cơ sở bán kính là kỹ thuật được nghiên cứu và sử dụng phổ biến hiện nay. Tư tưởng của kỹ thuật này là khuôn mặt 3D sẽ được đặc trưng bởi một tập điểm gọi là điểm điều khiển, việc biểu diễn trạng thái khuôn mặt 3D sẽ dựa trên việc điều chỉnh các điểm điều khiển của mô hình khuôn mặt ban đầu thành mô hình khuôn mặt đích (mô hình trạng thái khuôn mặt 3D mà chúng ta muốn biểu diễn. Ví dụ trạng thái vui, buồn, ngạc nhiên, .v.v.). Bắt đầu chúng ta sẽ tính các hệ số biến đổi dựa vào sự biến đổi của tập điểm điều khiển từ mô hình khuôn mặt ban đầu tới mô hình khuôn mặt đích, sau đó việc tính toán lại dữ liệu của khuôn mặt sẽ được thực hiện bởi nội suy dựa trên hàm cơ sở bán kính với các giá trị vừa tính được.



Hình 2.8. Các điểm điều khiển của mô hình khuôn mặt chung

Giả sử rằng sự biến đổi tập điểm điều khiển được biết đến với n vị trí x_i trong không gian R^3 và rằng thông tin này được đại diện bởi một vector mô tả 3D u_i rời rạc của hình học của khuôn mặt đã được đặt ở vị trí x_i trong bề mặt khuôn mặt ban đầu. Chúng ta có thể xem các vị trí x_i như các điểm điều khiển đã được chuyển đến vị trí $x_i + u_i$. Phương pháp nội suy RBF lúc này có thể được sử dụng cho các nội suy chuyển vị cho các vị trí khác.

Khi đó để giải quyết bài toán này RBF đưa ra hai pha, pha thứ nhất là xác định sự biến đổi của tập điểm điều khiển và tính ra các tham số \mathbf{A} , pha thứ hai là tính toán sự biến đổi của các điểm dữ liệu dựa vào vector \mathbf{A} và ma trận khoảng cách \mathbf{G}' .

Sử dụng các ký hiệu: $x_i = (x_i, y_i, z_i)$ và $u_i = (u_i^x, u_i^y, u_i^z)$ ba hệ thống tuyến tính được thiết lập như trên cho phép chúng ta tính các dịch chuyển u dựa vào các tham số \mathbf{A} tính được ở phần trước:

$$\begin{aligned} \mathbf{G}'\mathbf{A}_x &= (u_1^x, u_2^x, \dots, u_m^x)^T \\ \mathbf{G}'\mathbf{A}_y &= (u_1^y, u_2^y, \dots, u_m^y)^T \\ \mathbf{G}'\mathbf{A}_z &= (u_1^z, u_2^z, \dots, u_m^z)^T \end{aligned} \quad (2.15)$$

Với m là số điểm trên bề mặt, và \mathbf{G}' là ma trận được xây dựng là $(m*n+4)$ tương tự như ma trận \mathbf{G} trong công thức (2.14).

Thực hiện:

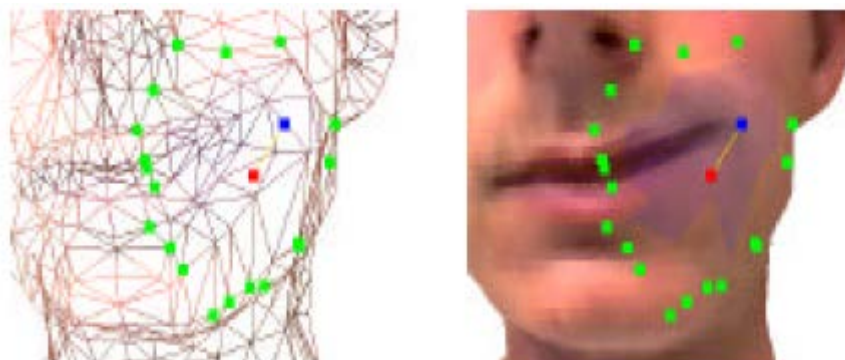
Đầu vào: Mô hình khuôn mặt 3D ban đầu, thông tin điểm điều khiển

Đầu ra: Mô hình trạng thái biểu cảm đích của khuôn mặt 3D

Thuật toán:

Bước 0: Khởi tạo mô hình 3D khuôn mặt và thông tin về tập điểm điều khiển.

Bước 1: Xác định thông tin về sự biến đổi của tập điểm điều khiển.



Hình 2.9. Minh họa sự biến đổi của điểm điều khiển.

Bước 2: Tính tham số $A = (a_1, a_2, \dots, a_M, c_0, c_1, c_2, c_3)$ theo (2.14).

Bước 3: Thực hiện nội suy dựa trên hàm RBF của trạng thái biểu cảm ban đầu để thu được trạng thái mới của khuôn mặt 3D

Bước 4: Kết thúc.

2.3.3. Nhận xét**Ưu điểm:**

- Kỹ thuật nội suy này sử dụng các điểm điều khiển để nội suy ra các trạng thái biểu cảm khuôn mặt, do đó chúng ta có thể dễ dàng thiết kế các điểm điều khiển này tại các vùng đặc trưng có sự biến đổi rõ rệt giữa các trạng thái biểu cảm sao cho phù hợp với khuôn mặt và phù hợp với các vị trí phức tạp có sự biến đổi trên khuôn mặt.
- Trong kỹ thuật nội suy này việc thêm các điểm điều khiển mới không làm ảnh hưởng nhiều tới thuật toán.
- Kỹ thuật này không làm thay đổi được cấu trúc lưới bề mặt của khuôn mặt.

Nhược điểm:

- Khả năng tính toán chậm, khó cài đặt.

- Nếu số điểm điều khiển tăng lên thì độ phức tạp thời gian tính toán của thuật toán cũng tăng lên khá nhanh.

CHƯƠNG 3 – THỰC NGHIỆM

3.1. Giới thiệu bài toán

Khuôn mặt con người là một đối tượng nghệ thuật và khía cạnh trung tâm, đóng một vai trò quan trọng trong quá trình tương tác. Các tỷ lệ và những biểu cảm khuôn mặt là rất quan trọng để xác định nguồn gốc, khuynh hướng tình cảm, chất lượng sức khỏe và thường là nền tảng tương tác xã hội của con người. Rất nhiều loại thông tin quan trọng khác nhau có thể nhìn thấy trên khuôn mặt. Vì vậy, hoạt hóa mặt người là một sự thay thế quan trọng cho phép tương tác người máy một cách tự nhiên. Hoạt hóa khuôn mặt có sự hỗ trợ của máy tính làm cho các ứng dụng đa dạng bao gồm cả con người ảo thực trong ngành công nghiệp giải trí, y tế, pháp y và chuẩn đoán.

Các khuôn mặt nói chuyện tương tác trong các ứng dụng truyền thông làm cho người sử dụng tương tác máy tốt hơn bằng cách cung cấp một giao diện thân thiện để thu hút người dùng. Về tính hiện thực của tổng hợp hoạt ảnh động, việc tạo ra các biểu cảm như con người là một trong những vấn đề quan trọng. Sự phức tạp liên quan đến biến dạng khuôn mặt di chuyển và sự nhạy cảm vốn có của con người đến sự tinh tế của chuyển động mặt làm cho hoạt hóa mặt là một chủ đề thách thức với cộng đồng đồ họa máy tính. Hơn nữa, việc giải thích các cảm xúc của con người vẫn còn là một đề tài nghiên cứu đa ngành cực kỳ khó khăn trong đồ họa máy tính, trí tuệ nhân tạo, giao tiếp, tâm lý học, v.v.

Bài toán biểu diễn trạng thái biểu cảm khuôn mặt 3D được mô tả như sau:

Đầu vào: Mô hình khuôn mặt 3D ở trạng thái ban đầu và khuôn mặt 3D ở trạng thái đích.

Đầu ra: Các khuôn mặt trung gian thể hiện cho quá trình hoạt hóa khuôn mặt 3D từ trạng thái ban đầu sang trạng thái đích. Các khuôn mặt trung gian này được tính toán dựa trên hai trạng thái khuôn mặt ở đầu vào.

Để giải quyết bài toán trên luận văn tiến hành cài đặt chương trình thực nghiệm quá trình hoạt hóa khuôn mặt 3D từ trạng thái ban đầu là trạng thái tự nhiên tới các trạng thái đích là các trạng thái biểu cảm cơ bản của khuôn mặt như vui vẻ, buồn, ngạc nhiên, giận giữ, căm ghét, sợ hãi. Quá trình xây dựng chương trình thực nghiệm sẽ bao gồm các bước chính như sau:

Bước 1: Xây dựng mô hình khuôn mặt 3D. Quá trình xây dựng mô hình khuôn mặt có thể dựa vào các phần mềm thiết kế mô hình có sẵn hiện nay. Chẳng hạn như Maya, 3Dmax, .v.v.

Bước 2: Hoạt hóa khuôn mặt 3D thể hiện các trạng thái biểu cảm bằng cách sử dụng kỹ thuật nội suy trong thuật toán morphing. Thực chất của quá trình hoạt hóa khuôn mặt 3D là quá trình hoạt hóa các thành phần mô phỏng khuôn mặt. Thuật toán morphing sử dụng kỹ thuật nội suy tuyến tính hoạt hóa khuôn mặt 3D ở những vùng đặc trưng cơ bản nhất của khuôn mặt 3D. Quá trình hoạt hóa khuôn mặt 3D chính là quá trình biến đổi các khung hình thể hiện khuôn mặt 3D phù hợp theo thời gian.

3.2. Phân tích chương trình thực nghiệm

3.2.1. Các công cụ kỹ thuật

Như đã trình bày ở trên, việc đầu tiên trong quá trình xây dựng thực nghiệm là đi xây dựng mô hình 3D khuôn mặt. Chúng ta có thể xây dựng mô hình 3D khuôn mặt bằng cách sử dụng các phần mềm thiết kế như 3Dmax, Maya hoặc một số phần mềm thiết kế khác. Các mô hình 3D được thiết kế sẽ làm cơ sở dữ liệu trong quá trình hoạt ảnh khuôn mặt.

Sau khi xây dựng xong các mô hình 3D khuôn mặt, chúng ta có thể sử dụng các ngôn ngữ lập trình như VC++, C++ .v.v. và OpenGL để lập trình điều khiển, hoạt hóa các trạng thái của khuôn mặt 3D.

OpenGL (Open Graphics Library) là một thư viện đồ họa có khoảng 150 hàm nhằm giúp xây dựng các đối tượng và các giao tác cần thiết trong các ứng dụng tương tác 3D. OpenGL cũng là một tiêu chuẩn kỹ thuật đồ họa có chức năng tiêu chuẩn hóa giao diện lập trình ứng dụng đồ họa 3 chiều [33].

Thư viện OpenGL chứa các hàm đồ họa dùng để xây dựng các đối tượng phức tạp từ các thành phần hình học cơ bản như điểm, đoạn thẳng, đa giác, ảnh,..., các hàm sắp xếp đối tượng trong 3D và chọn điểm thuận lợi để quan sát, các hàm đồ họa tính toán màu sắc của các đối tượng, các hàm biến đổi những mô hình toán học của đối tượng và thông tin màu sắc thành các pixel trên màn hình [34]. Nhưng thư viện này lại không hỗ trợ một số hàm nhập xuất hay thao tác trên window, cũng không có sẵn các hàm để xây dựng các mô hình đối tượng mà người lập trình phải tự xây dựng từ các thành phần hình học cơ bản.

Do đó, luận văn có sử dụng thêm một số thư viện đồ họa hỗ trợ khác như GLUT, GLEW.

C++ là một ngôn ngữ lập trình hướng đối tượng mạnh và linh hoạt. C++ có thể được sử dụng để phát triển những ứng dụng bậc cao và cả những chương trình bậc thấp hoạt động tốt trên phần cứng. Ngôn ngữ C++ cũng hay được sử dụng trong các dự án về lập trình đồ họa, xử lý văn bản, bảng tính, .v.v. C++ cũng có sẵn rất nhiều các trình biên dịch, các thư viện được viết sẵn, đồng thời người lập trình cũng có thể tự tạo thêm các hàm thông dụng vào thư viện để có thể tái sử dụng sau này.

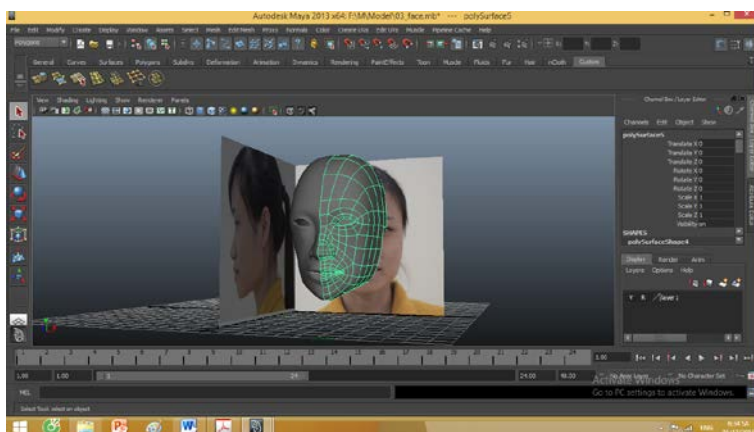
Môi trường sử dụng cho ngôn ngữ C++ là Microsoft Visual C++. Đây là một môi trường phát triển tích hợp có các công cụ cho phát triển và gỡ lỗi mã nguồn C++ trợ giúp cho các lập trình viên trong việc xây dựng và phát triển phần mềm. Ngoài ra, Microsoft Visual C++ còn có thể bao gồm hệ thống quản lý các phiên bản, các công cụ nhằm giúp cho việc đơn giản hóa công việc xây dựng giao diện người dùng đồ họa. Microsoft Visual C++ là một môi trường phát triển đồ họa mạnh.

Với những lợi thế và ưu điểm có được của C++ và thư viện OpenGL nên luận văn đã lựa chọn ngôn ngữ lập trình C++ kết hợp với thư viện OpenGL để lập trình điều khiển quá trình hoạt hóa khuôn mặt 3D.

3.2.2. Các bước triển khai cụ thể

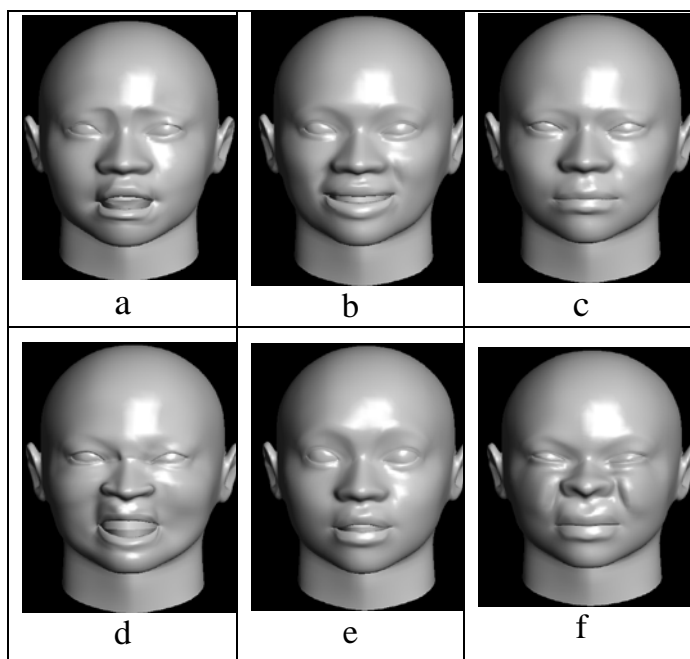
Bước 1: Xây dựng mô hình khuôn mặt 3D.

- Luận văn sử dụng phần mềm Autodesk Maya 2013 để xây dựng mô hình khuôn mặt đầu vào cho chương trình thực nghiệm ở trạng thái tự nhiên.



Hình 3.1. Mô hình khuôn mặt 3D được xây dựng trong Maya

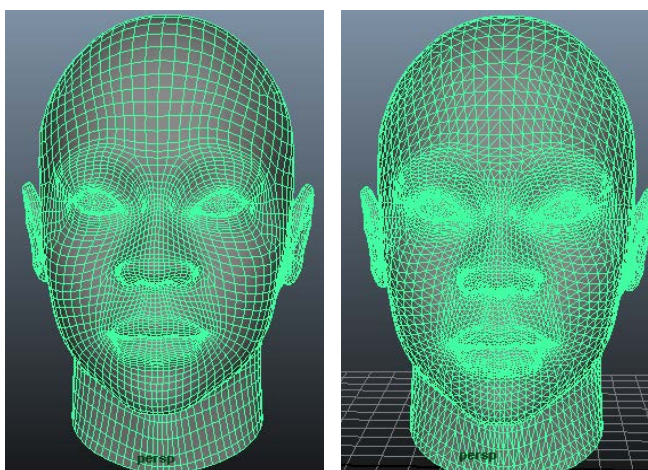
- Sau đó dựa vào mô hình ban đầu này, luận văn tiến hành xây dựng 6 mô hình trạng thái khuôn mặt 3D chính dựa vào một số đặc trưng cho từng trạng thái biểu cảm như đã được mô tả trong chương 1.



Hình 3.2. Các trạng thái biểu cảm cơ bản của khuôn mặt

a. sợ hãi, b. trạng thái vui vẻ, c. buồn rầu, d. giận giữ, e. Ngạc nhiên, f. căm ghét.

- Để thuận lợi cho quá trình xây dựng chương trình thực nghiệm và quá trình hoạt hóa dựa vào kỹ thuật nội suy, tôi chuyển mô hình khuôn mặt 3D từ lưới tứ giác đã xây dựng thành lưới tam giác.



Hình 3.3. Mô hình khuôn mặt 3D dưới dạng lưới

- Lưu lại các mô hình trạng thái khuôn mặt 3D dưới dạng file .Obj

Bước 2: Load dữ liệu vào chương trình thực nghiệm

Dữ liệu được đưa vào chương trình thực nghiệm dưới dạng file mô hình đuôi .obj (dạng lưu trữ mô hình 3D). Sau đó được lưu trữ dưới dạng mảng các đỉnh với các tọa độ x, y, z của mỗi đỉnh và thông tin về texture.

Hàm LoadOBJ() trong chương trình nguồn được sử dụng để đưa dữ liệu vào chương trình thực nghiệm.

Bước 3: Tính toán các mô hình khuôn mặt trung gian dựa vào kỹ thuật nội suy tuyến tính.

- Các mô hình khuôn mặt trung gian được tạo ra bằng cách tính toán các giá trị vị trí các đỉnh trên mô hình này. Vị trí các đỉnh trên mô hình khuôn mặt trung gian được tính toán dựa vào công thức (2.5) như sau:

```
GLvoid COGLView::morphModel(t_Bone *curBone)
{
    int loop,pointloop;
    float *dest,*src1,*src2,ratio;

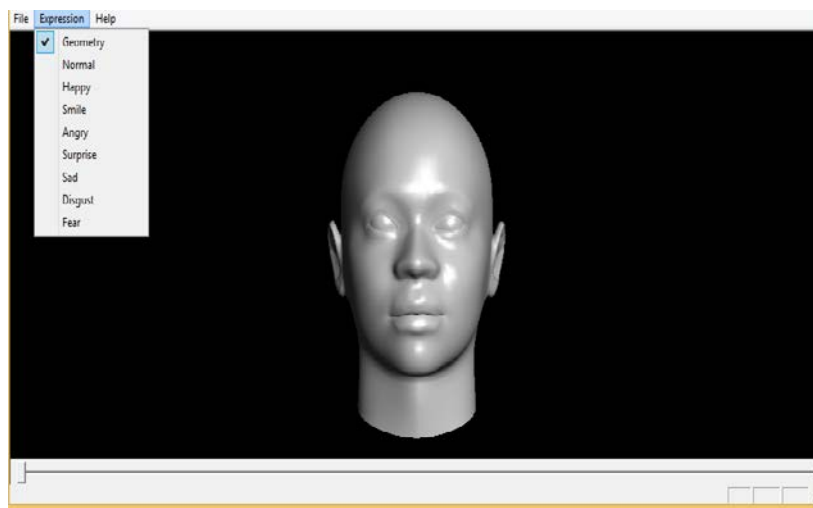
    if (curBone->visualCnt > m_curVisual && curBone->visuals[0].vertexData != NULL)
    {
        src1 = curBone->visuals[0].vertexData;    // lưu trữ mô hình ban
        dau
        src2 = curBone->visuals[1].vertexData;    // mô hình đích
        dest = curBone->visuals[2].vertexData; // mô hình trung gian
        ratio = m_Slider->GetSetting();    // tham số nội suy

        for (loop = 0; loop < curBone->visuals[0].triCnt * 3; loop++)
        {
            for (pointloop = 0; pointloop < curBone->visuals[0].vSize;
            pointloop++)
            {
                dest[(loop * curBone->visuals[0].vSize) + pointloop] =
                LERP(src1[(loop * curBone->visuals[0].vSize) +
            pointloop],
                src2[(loop * curBone->visuals[0].vSize) +
            pointloop],ratio);
            }
        }
    }
}
```

Bước 4: Hiện thị mô hình khuôn mặt trung gian.

3.3. Kết quả thực nghiệm

Khi chạy chương trình thực nghiệm sẽ xuất hiện với giao diện như sau:



Hình 3.4. Mô hình khuôn mặt 3D ở trạng thái ban đầu

Khi chạy chương trình, mô hình khuôn mặt 3D được nạp vào chương trình ở trạng thái bình thường của khuôn mặt và sẵn sàng cho quá trình hoạt hóa. Chúng ta có thể sử dụng chuột điều chỉnh để quan sát mô hình.

Chúng ta có thể bắt đầu quá trình hoạt hóa bằng cách lựa chọn các trạng thái biểu cảm cơ bản cần hoạt hóa ở menu Expression và kéo thanh trượt ở phía cuối màn hình để bắt đầu quá trình hoạt hóa.

Một số trạng thái biểu cảm cơ bản trong quá trình hoạt hóa:



a



b



c



d



f



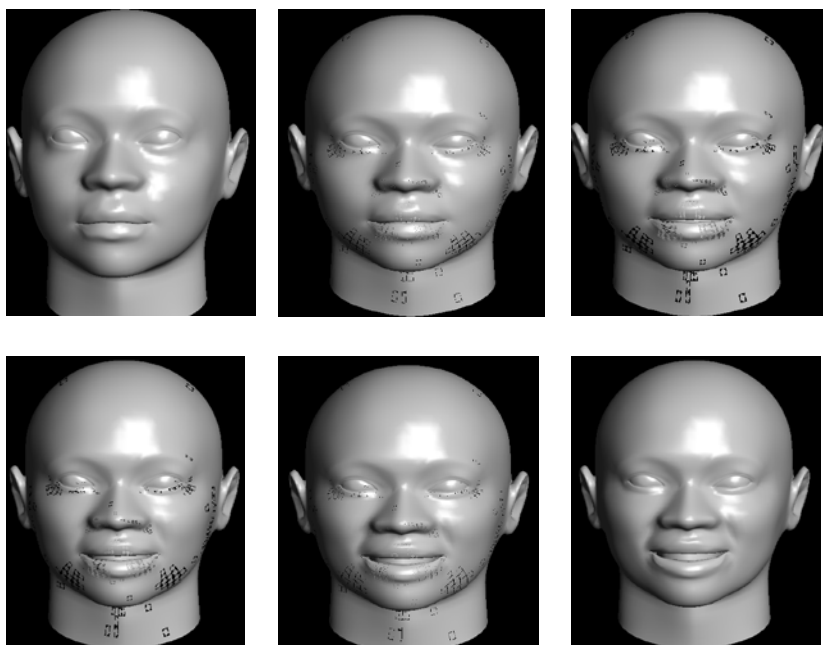
e



g

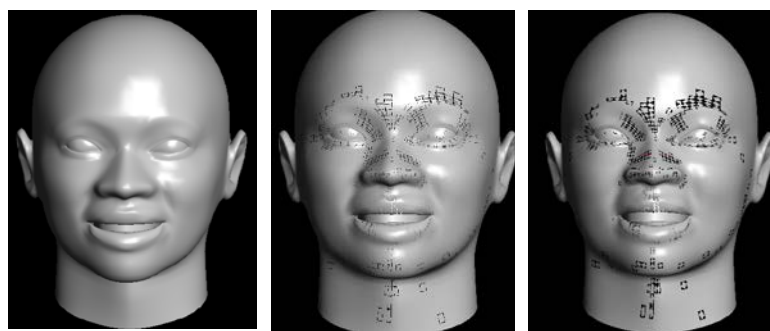
Hình 3.5. Các trạng thái biểu cảm cơ bản của khuôn mặt

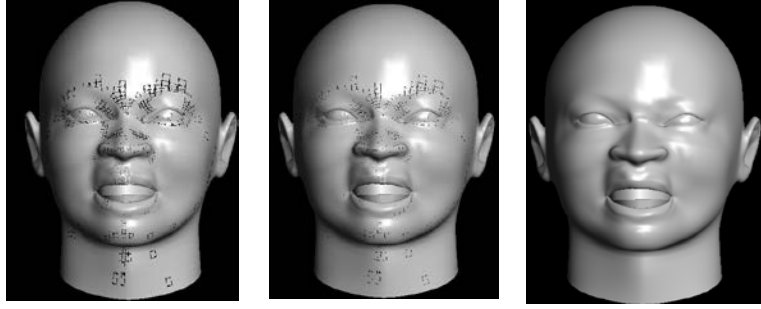
a. Trạng thái tự nhiên, b. trạng thái vui vẻ, c. buồn rầu, d. giận giữ, e. Ngạc nhiên, f. căm ghét, g. sợ hãi



Hình 3.6. Quá trình hoạt hóa từ trạng thái khuôn mặt tự nhiên sang trạng thái vui

vẻ





Hình 3.7. Hoạt hóa từ trạng thái vui về sang trạng thái tức giận

Để thoát khỏi ứng dụng chúng ta chọn chức năng Exit từ menu File trong cửa sổ chương trình.

KẾT LUẬN

Với những nội dung đã tìm hiểu, nghiên cứu và được trình bày trong luận văn cả về lý thuyết và thực nghiệm, luận văn đã đạt được những kết quả như sau:

- Tìm hiểu khái quát về đối tượng 3D, mô hình 3D và cách xây dựng mô hình 3D.

- Tìm hiểu khái quát về cách thức thể hiện một số trạng thái biểu cảm khuôn mặt cơ bản trong thực tế, cách thức mô hình hóa khuôn mặt 3D và một số kỹ thuật sử dụng để hoạt hóa trạng thái khuôn mặt 3D.

- Hệ thống được 3 kỹ thuật nội suy sử dụng trong biểu diễn trạng thái khuôn mặt 3D như nội suy tuyến tính, kỹ thuật nội suy song tuyến tính, kỹ thuật nội suy dựa vào hàm cơ sở bán kính, nhận xét phù hợp đối với từng kỹ thuật.

- Cài đặt thử nghiệm chương trình biểu diễn trạng thái biểu cảm khuôn mặt 3D trong thực tại ảo dựa vào kỹ thuật nội suy.

Hướng mở rộng nghiên cứu trong tương lai:

Khuôn mặt của con người ngày càng được nghiên cứu một cách chuyên sâu hơn để xác định những ảnh hưởng lên những nhận thức về sức khỏe sắc đẹp, tình cảm, tuổi và hình dạng của con người. Biểu diễn trạng thái khuôn mặt 3D có nhiều ứng dụng trong thực tế. Luận văn hoàn thành cài đặt bài toán biểu diễn trạng thái khuôn mặt 3D dựa vào kỹ thuật nội suy. Tuy nhiên để có thể biểu diễn, hoạt hóa khuôn mặt một cách chính xác, thực tế còn nhiều vấn đề cần giải quyết như các vấn đề về màu da, kết cấu da, các ảnh hưởng của các vùng khuôn mặt khác nhau trong quá trình hoạt hóa, v.v. Đây là các hướng nghiên cứu mở rộng của biểu diễn trạng thái khuôn mặt 3D trong tương lai.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1] Nguyễn Ái Việt, Đỗ Năng Toàn, Trịnh Xuân Đàn, Phạm Bá Máy, Hồ Xuân Nhàn (2013), “*Mô hình ba chiều và xây dựng mô hình bộ phận cơ thể ảo*”, Tạp chí Y học Việt Nam - Tập 411, năm 2013.
- [2] Phạm Việt Bình, Đỗ Năng Toàn (2008), *Xử lý ảnh*, NXB Khoa học kỹ thuật, Hà Nội.
- [3] Đặng Quang Á, “*Tiếp cận dựa trên hàm RBF tới bài toán khôi phục mặt người từ hộp sọ*”, Hội thảo quốc gia “Ứng dụng CNTT và truyền thông trong hoạt động đào tạo, nghiên cứu khoa học và quản lý điều hành tại trường đại học”, Thái nguyên, 7-8/12/2007
- [4] Nguyễn Thị Nguyệt, *Nội suy ảnh và một số ứng dụng*, Luận văn thạc sĩ, Khoa công nghệ thông tin, Đại học Thái Nguyên, Tr(36-42).

Tiếng Anh

- [5] Mattos, A.B., Mena-Chalco, J.P., Cesar, R., & Velho, L. (2010). *3D linear facial animation based on real data*. In Proceedings of the 23rd SIBGRAPI conference on graphics, patterns and images (pp. 271-278).
- [6] M. H. Alkawaz, Dzulkifli Mohamad, A. H. Basori, T. Saba (2015), *Blend shape interpolation and FACS for realistic avanta*. In Springer
- [7] Shaver, Phillip, et al (1987), “Emotion knowledge: further exploration of a prototype approach”, *Journal of personality and social psychology*, Vol 52, No 6, pp. 1061-1086.
- [8] 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.
- [9] The Duy Bui (2004), “*Creating emotions and facial expressions for Agents*”, the Thesis.
- [10] Pighin, F., & Lewis, J. (2006). *Performance – driven facial animation*. In ACM SIGGRAPH.

- [11] Zhang, Q., Liu, Z., Quo, G., Terzopoulos, D., et al (2006). *Geometry – driven photorealistic facial expression synthesis*. IEEE Transactions on Visualization and Computer Graphic,. 12(1), 48-60
- [12] Blinn, J.F.(1978). *Simulation of wrinkled surfaces*, In ACM SIGGRAPH computer graphics (pp. 286-292). New York: ACM Press.
- [13] Boissieux, L., Kiss, G., Thalmann, N. M., & Kalra, P.(2000). *Simulation of skin aging and wrinkles with cosmetics insight*. In Proceedings of the eurographic workshop on computer animation and simulation (pp. 15-27) Berlin: Springer.
- [14] Hardy RL, “*Multiquadric equations of topography and other irregular surfaces*”, J Geophy Res 76:1905–1915, 1971
- [15] Mairhuber JC, “*On Haar’s theorem concerning Chebyshev approximation problems having unique solutions*”, Proc Amer Math Soc 7(4):609–615, 1956
- [16] Bochner S, “*Monotone Functionen Stieltjes Integrale und Harmonische Analyse*”, Ann Math 108:378–410, 1933
- [17] Schoenberg IJ (1938), “*Metric spaces and completely monotone functions*”, Ann Math 39:811–841
- [18] Micchelli CA, “*Interpolation of scattered data: Distance matrices and conditionally positive definite functions*”, Constr Approx 2:11–22, 1986
- [19] Powell MJD, “*The theory of radial basis function approximation*”, In: Light W(ed) *Advances in Numerical Analysis, Vol. II: Wavelets, Subdivision Algorithms and Radial Functions*, Oxford University Press, Oxford, UK, pp 105–210, 1992
- [20] Madych WR, Nelson SA, “*Bounds on multivariate polynomials and exponential error estimates for multiquadric interpolation*”, J Approx Theory 70:94–114, 1992
- [21] Yoon J, “*Spectral approximation orders of radial basis function interpolation on the Sobolev space*”, SIAM J Math Anal 33(4):946–958, 2001

- [22] Driscoll TA, Fornberg B, “*Interpolation in the limit of increasingly flat radial basis functions*”, *Comput Math Appl* 43:413–422, 2002
- [23] Fornberg B, Piret C, “*A stable algorithm for flat radial basis functions on a sphere*”, *SIAM J Sci Comput* 30:60–80, 2007
- [24] V. Blanz, T. Vetter, “*A morphable model for the synthesis of 3d faces*”, SIGGRAPH '99 Proceedings of the 26th annual conference on Computer graphics and interactive techniques Pages 187-194.
- [25] Park, F. I. “*A parametric model for human faces*”, 1974, Universidad de Utah, USA.
- [26] Phothisane, P., Bigorgne, E., Collot, L., Prévost, L, “*a robust composite metric for head pose tracking using an accurate face model*”, 2011, Automatic Face & Gesture Recognition and Workshops (FG 2011), IEEE International Conference.
- [27] Conde, C. “*verificación facial multimodal 2D y 3D*”, 2006. Tesis Doctoral. Departamento de Arquitectura y tecnología de computadores, ciencia de la computacion e inteligencia artificial.
- [28] Bronstein, A., Bronstein, M., Kimmel, R., “*Expression invariant 3d face recognition*”, Springer-Verlag Berlin Heidelberg 2003.
- [29] Golovinskiy, A., Wojciech Matusik, Hanspeter Pfister, Szymon Rusinkiewicz, Thomas Funkhouser. “*A statistical model for synthesis of detailed facial geometry*”, *ACM Transactions on Graphics (Proc. SIGGRAPH)*, July 2006.
- [30] Ramírez Valdez Leonel, Hasimoto Beltran Rogelio, “*3D-facial expression synthesis and its application to face recognition systems*”, *Journal of Applied Research and Technology* 2009, Redalyc.
- [31] Cootes, T.F., Taylor, C.J. “*Statistical models of appearance for computer vision*”. *Imaging Science and Biomedical Engineering*, University of Manchester. 2004.
- [32] El-Hussuna, A. “*Statistical variation of three dimensional face models*”. Master Thesis. ITUniversity of Copenhagen, Multimedia Technologies. March 2003.
- [33] <http://vi.wikipedia.org/wiki/OpenGL>
- [34] <http://glprogramming.com/red>