

00327

3



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

FACULTAD DE CIENCIAS

Técnica No Supervisada de Generación de Modelos Faciales  
3D Texturizados usando Análisis de Componentes Principales

T E S I S

QUE PARA OBTENER EL TITULO DE:  
LICENCIADO EN CIENCIAS DE LA COMPUTACION

PRESENTA:

HECTOR CUEVAS VAZQUEZ DEL MERCADO

DIRECTOR DE TESIS

DR. ISAAC JUAN RUDOMIN GOLDBERG

TESIS CON  
FALLA DE ORIGEN

2003



FACULTAD DE CIENCIAS  
UNAM

DIVISION DE ESTUDIOS PROFESIONALES



FACULTAD DE CIENCIAS  
SECCION ESCOLAR



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**DRA. MARÍA DE LOURDES ESTEVA PERALTA**  
Jefa de la División de Estudios Profesionales de la  
Facultad de Ciencias  
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

Técnica No Supervisada de Generación de Modelos Faciales 3D Texturizados  
usando Análisis de Componentes Principales

realizado por Héctor Cuevas Vázquez del Mercado

con número de cuenta 9650529-0 , quién cubrió los créditos de la carrera de Ciencias de la  
Computación.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario Dr. Isaac Juan Rudomín Goldberg

Propietario Mat. Ana Luisa Solís González-Cosío

Propietario M. en C. Miguel Miranda Miranda

Suplente M. en I. Ma. de Luz Gasca Soto

Suplente M. en C. María Guadalupe Ibargüengoitia González

Consejo Departamental de Matemáticas



FACULTAD DE CIENCIAS  
DRA. AMPARO LÓPEZ ZAMORA  
DE  
MATEMÁTICAS

TESIS CON  
FALLA DE ORIGEN

# **Agradecimientos**

**En primer lugar, a mi director de tesis, Isaac Rudomín, por su ayuda y paciencia durante la realización de este trabajo.**

**Al jurado que revisó esta tesis, y a todas las personas que en diversos momentos y formas han contribuido a la realización de este trabajo. En este último grupo quisiera mencionar a José Luis Villarreal, del Departamento de Visualización en DGSCA.**

**Finalmente, a todas las personas que me han enseñado algo.**

## **Resumen**

La aportación principal de este trabajo es una extensión de la técnica de rostros característicos para lograr la generación de un modelo facial 3D texturizado a partir de una fotografía frontal de un rostro bajo iluminación simple.

Para lograr esto se analiza previamente un conjunto facial de inicialización compuesto por rostros 2D y 3D, utilizando Análisis de Componentes Principales, una técnica de ordenamiento y reescritura de datos que obtiene una base generadora ortonormal para los mismos. La técnica original de rostros característicos utiliza esta base generadora para generar imágenes a partir de imágenes; en este trabajo se analiza el problema de usarla para generar modelos faciales 3D a partir de imágenes faciales.

Como resultado de este trabajo se aporta la técnica de *rostros característicos con datos mixtos*, y se realizan diferentes pruebas para determinar empíricamente su validez y robustez frente a condiciones de ruido y oclusión.

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Aplicaciones de la generación de modelos faciales 3D	1
1.2. Rostros característicos con datos mixtos	2
1.3. Técnicas de generación de modelos faciales	3
1.3.1. Ajuste de puntos característicos	5
1.3.2. Reconstrucción estéreo	7
1.3.3. Síntesis por estimación	8
1.3.4. Modelado basado en imágenes	10
1.4. Análisis de Componentes Principales aplicado a rostros	10
1.4.1. Rostros Característicos	11
1.4.2. Análisis de Componentes Principales	12
1.4.3. Ejemplo de aplicación a rostros	14
<b>2. Generación de modelos 3D usando rostros característicos</b>	<b>17</b>
2.1. Planteamiento del problema	17
2.1.1. Limitaciones de los rostros característicos	18
2.2. Bosquejo de la solución	19
2.2.1. Requerimiento de codificaciones intercambiables	20
2.2.2. Obtención de modelos 3D generadores	22
2.3. Solución: Rostros Característicos con Datos Mixtos	24
<b>3. Implementación</b>	<b>27</b>
3.1. Características del conjunto facial de inicialización	27
3.1.1. Imágenes Faciales	27
3.1.2. Modelos Faciales	28

3.2.	Obtención del conjunto facial de inicialización . . . . .	29
3.2.1.	Normalización geométrica . . . . .	29
3.3.	Inicialización . . . . .	33
3.3.1.	Obtención de imágenes características de rostros . . . . .	33
3.3.2.	Obtención de modelos 3D generadores de rostros . . . . .	33
3.4.	Generación de modelos faciales 3D . . . . .	34
<b>4.</b>	<b>Experimentos y Resultados</b>	<b>35</b>
4.1.	Conjunto Facial de Inicialización . . . . .	35
4.2.	Perturbaciones a analizar . . . . .	37
4.3.	Medidas . . . . .	40
4.4.	Conjuntos de prueba . . . . .	47
4.5.	Resultados . . . . .	49
4.5.1.	Imágenes espurias . . . . .	49
4.5.2.	Imágenes perturbadas y ocluidas . . . . .	50
4.5.3.	Rostros fuera del conjunto de inicialización . . . . .	50
<b>5.</b>	<b>Conclusiones y Trabajo Futuro</b>	<b>53</b>
5.1.	Conclusiones . . . . .	53
5.2.	Trabajo Futuro . . . . .	54
<b>A.</b>	<b>Formato OBJ</b>	<b>55</b>
<b>B.</b>	<b>Uso, diseño y programación de la implementación</b>	<b>57</b>
B.1.	Uso . . . . .	57
B.2.	Programas implementados . . . . .	57
B.2.1.	Procesamiento de modelos faciales 3D . . . . .	57
B.2.2.	Obtención de rostros característicos y generadores . . . . .	59
B.2.3.	Generación de un modelo . . . . .	60
B.2.4.	Realización de experimentos . . . . .	62
B.3.	Tipos de Datos Abstractos . . . . .	63
B.4.	Plataforma de Programación . . . . .	64
B.4.1.	Rendimiento y uso de disco . . . . .	65
<b>C.</b>	<b>Gráficas de resultados</b>	<b>67</b>
C.1.	Descripción de las gráficas . . . . .	67
C.1.1.	Oclusiones . . . . .	68
C.1.2.	Perturbaciones . . . . .	77

<b>D. Modelos Generados</b>	<b>87</b>
D.1. 120 × 120 . . . . .	<b>88</b>
D.1.1. Original . . . . .	<b>88</b>
D.1.2. Barra horizontal de 18 pixels en gris . . . . .	<b>89</b>
D.1.3. Ruido Aleatorio 10 por ciento . . . . .	<b>90</b>
D.2. 60 × 60 . . . . .	<b>91</b>
D.2.1. Original . . . . .	<b>91</b>
D.2.2. Barra horizontal de 9 pixels en gris . . . . .	<b>92</b>
D.2.3. Ruido Aleatorio 10 por ciento . . . . .	<b>93</b>
 <b>Literatura Citada</b>	 <b>95</b>



# Índice de figuras

1.1. Deformación de un modelo genérico en base a puntos característicos	6
1.2. Obtención de la textura a partir de dos vistas . . . . .	6
1.3. Mapa de profundidad obtenido a partir de las disparidades de un par estéreo . . . . .	7
1.4. Rostros obtenidos a partir de diferentes parametrizaciones de un modelo genérico deformable . . . . .	9
1.5. El análisis de un conjunto facial mixto permite generar rostros nuevos . . . . .	9
1.6. El rostro promedio es sobrepuesto en la imagen original para comenzar la aproximación . . . . .	10
1.7. De imágenes a rostros característicos . . . . .	14
3.1. Imagen y modelo faciales antes del procesamiento . . . . .	29
3.2. Selección de la máscara facial . . . . .	30
3.3. Normalización geométrica por mitades y aplicación de la máscara facial . . . . .	32
4.1. Algunas de las personas en el conjunto facial de inicialización . . . . .	37
4.2. Imagen de prueba original . . . . .	38
4.3. Filtro de medianas con aperturas de 3, 5 y 7 pixels . . . . .	38
4.4. Ruido Aleatorio del 5 y 10 por ciento . . . . .	39
4.5. Ruido Uniforme . . . . .	39
4.6. Filtro promedio con apertura de 3 pixels . . . . .	41
4.7. Mosaico de 2 y 3 pixels de apertura . . . . .	41
4.8. Oclusiones. Barras horizontales en tres anchos (6,12 y 18 pixels) en tres colores: blanco, gris y negro centradas en la imagen. . . . .	42
4.9. Oclusiones: Barras verticales en tres anchos (6,12 y 18 pixels) en tres colores: blanco, gris y negro centradas en la imagen. . . . .	43

4.10. Máscaras en tres colores: blanco, gris y negro. De arriba a abajo, solamente son visibles la boca, la nariz y finalmente nariz y boca. . . . .	44
4.11. Máscaras en tres colores: blanco, gris y negro. De arriba a abajo, solamente son visibles ojos, ojos y boca, ojos y nariz . . . . .	45
4.12. Máscaras en tres colores: blanco, gris y negro, mostrando solamente los rasgos del rostro: ojos, nariz y boca. . . . .	46
4.13. Resultados de usar una imagen espuria con el conjunto generador de hombres . . . . .	49
4.14. Resultados de usar una oclusión horizontal negra de 18 pixeles con el conjunto generador de hombres . . . . .	50
4.15. Las dos imágenes de entrada y las generadas . . . . .	50
4.16. Generación de un rostro femenino fuera del conjunto de inicialización . . . . .	51
4.17. Generación de un rostro masculino fuera del conjunto de inicialización . . . . .	52
<b>B.1. La textura original es recortada y posteriormente escalada . . . . .</b>	<b>61</b>

# Índice de cuadros

1.1. Comparación de técnicas de generación de modelos faciales . . .	5
4.1. Nombrado de las imágenes de prueba (de 120 × 120 pixels) . . .	46

## CONTENTS

1. Introduction	1
2. Theoretical background	4

**PAGINACION**

**DISCONTINUA**

# Capítulo 1

## Introducción

En este capítulo se menciona la motivación para resolver el problema de la generación de modelos faciales tridimensionales y se describen las soluciones propuestas comparándolas entre sí y con la técnica propuesta en este trabajo: los *rostros característicos con datos mixtos*, llamados así por ser una extensión de la técnica para manipulación de rostros 2D, los *rostros característicos*.

### 1.1. Aplicaciones de la generación de modelos faciales 3D

Sin lugar a dudas el rostro es la principal característica fisonómica de una persona: a través de él se observan las emociones, se reconoce su identidad y es el gran portal del contacto con el mundo exterior.

Los avances tecnológicos han favorecido la aparición de aplicaciones que utilizan modelos faciales 3D para facilitar la interacción entre los usuarios y de la aplicación con los usuarios, pero la obtención de un modelo facial sin equipo especializado continúa siendo un proceso costoso.

Este trabajo presenta una solución al problema de generación rápida de modelos faciales 3D texturizados: tomando como entrada una fotografía frontal del rostro del individuo se obtiene un modelo texturizado de la parte visible del rostro. Los requerimientos para que esto sea posible se describen más adelante.

La aplicación principal de una técnica como la presentada en este trabajo es la generación rápida de rostros tridimensionales de avatares para mundos virtuales, bajo el esquema de un servidor centralizado que recibe fotografías de los usuarios. Así cada usuario tendrá un avatar con su rostro rápidamente solamente con enviar

una fotografía digital de su rostro. Aunque no se llevó a cabo en este trabajo, la técnica también permite la generación de modelos faciales animables basados en vértices asociados a movimientos faciales.

Sin embargo, teniendo el modelo facial es posible usarlo para cualquier aplicación de realidad virtual o mundos 3D: guías virtuales, salones virtuales donde los participantes son representados mediante un modelo facial por razones de privacidad o ancho de banda, una interfaz basada en expresiones faciales para motivar al usuario a usar el sistema, etcétera.

Otras aplicaciones para las que esta técnica es apropiada son la codificación de rostros tridimensionales para su transmisión o almacenamiento usando la menor cantidad de ancho de banda o espacio en disco, respectivamente. Como estas aplicaciones están fuera del alcance del presente trabajo solamente se mencionan brevemente.

## **1.2. Rostros característicos con datos mixtos**

La principal contribución de este trabajo es la técnica de rostros característicos con datos mixtos. Se muestra la forma en que esta técnica resuelve el problema de generación de modelos faciales 3D texturizados a partir de una fotografía de un rostro y se describe un grupo de pruebas y sus resultados para validar empíricamente la técnica.

Los rostros característicos con datos mixtos son una técnica holística (es decir, que trabaja sobre la totalidad de la entrada y no por partes) derivada de los rostros característicos de Turk y Pentland (1991).

Los rostros característicos originales fueron planteados para el reconocimiento de rostros representados como imágenes. Este reconocimiento se hacía transformando las imágenes de rostros a una representación de menor dimensión, llamada *codificación*, un pequeño vector de escalares. Al ser de menor dimensión, era más fácil comparar dos codificaciones que los rostros de los cuales se habían obtenido. Ya que es posible reconstruir un rostro a partir de su codificación, los rostros característicos también sirven para la generación de imágenes de rostros, ya sea utilizando codificaciones obtenidas de rostros o propuestas por el usuario.

Mediante la extensión propuesta en este trabajo es posible usar la codificación obtenida por los rostros característicos para generar ya no una imagen, sino el modelo facial correspondiente.

### **1.3. Técnicas de generación de modelos faciales**

Para fijar el contexto de este trabajo se describen las diferentes soluciones al problema, comparándolas entre sí y con la técnica propuesta en este trabajo.

Se discutirán los siguientes factores caracterizadores de las técnicas de generación de modelos faciales tridimensionales:

#### **Requerimientos de Inicialización**

- El requerimiento más frecuente es el de un conjunto facial compuesto por modelos 3D y generalmente complementado con texturas.

#### **Forma en la que se llega al modelo resultado**

- Deformación controlada de un modelo genérico. Esta deformación puede ser guiada visualmente por el usuario en los sistemas más sencillos o estar sujeta a una función de evaluación, basándose en las características topológicas del rostro (curvatura, puntos de inflexión) o en su apariencia (una medida matemática de su similitud visual, sin necesidad de despliegue).
- Construcción directa. En este caso el modelo resultado se construye directamente a partir de la información disponible sobre la superficie. El mejor ejemplo de construcción directa es cuando se utiliza la información de un mapa de profundidades (generalmente obtenido por métodos estereoscópicos) para obtener la malla que representa la superficie de un objeto.

#### **Intervención del usuario**

- El usuario puede supervisar el proceso de la reconstrucción modificando parámetros conforme avanza.
- En otros basta con que indique las condiciones iniciales y espere el resultado final.

#### **Criterio de terminación**

- La técnica puede terminar después de un número fijo de pasos.
- Al cumplirse una condición de convergencia, dada por la cercanía al resultado deseado o la falta de avance.

- El usuario decide cuándo detener el proceso.

### **Características y requerimientos de la entrada**

- Las técnicas basadas en la apariencia utilizan directamente la imagen o imágenes del objeto a modelar. Estas imágenes suelen estar normalizadas geoméricamente (posición, orientación, tamaño) aunque en algunos casos la técnica puede trabajar con rostros en diferentes poses. En otros el usuario debe indicar manualmente puntos característicos en el rostro como nariz, ojos y boca. Por lo general, se requiere que las condiciones de iluminación sean simples: una sola fuente de luz blanca de frente al objeto.
- La reconstrucción estéreo no solamente necesita varias imágenes (por definición requiere de por lo menos dos imágenes) sino de información de las condiciones en que fueron tomadas. Por lo mismo las imágenes no son normalizadas, ya que hacerlo haría inútil la información sobre las condiciones que fue tomada. Las imágenes en sí no se usan, solamente la posición que ocupan en ellas puntos correspondientes del objeto. La posición de los puntos correspondientes y no su apariencia determina el modelo.

### **Características de la salida**

- Los modelos faciales generados difieren no sólo en su calidad, sino en su grado de estructuración. Si hubo un conjunto de inicialización, los modelos resultado suelen tener la misma estructura, si ésta existía. Este suele ser el caso de técnicas basadas en el ajuste de puntos característicos o modelos genéricos deformables.
- Otras técnicas generan modelos sin ninguna estructura predecible, siendo éste el caso de la reconstrucción estéreo.
- Sobre la textura, ésta puede generarse independientemente o simplemente reusar la entrada dada al sistema.

En el Cuadro 1.1 se resumen estos aspectos para varias técnicas en su planteamiento tradicional. A continuación se discute cada una de ellas tomando como ejemplo práctico a sistemas que las implementan.

Técnica	Modelo Genérico	Inicialización/ Parametrización	Interactividad	Criterio de Terminación	Características de la Entrada	Características de la Salida
Ajuste de Puntos Característicos	Si	Si/No	Semimanual	Decisión del usuario	Una o más imágenes	Modelo Estructurado
Reconstrucción estéreo	No	No/Si	No Supervisado	Determinista	Das o más imágenes	Modelo No Estructurado
Síntesis por estimación	Si	Si/Si	NoSupervisado	Convergencia	Una imagen	Modelo Estructurado
Modelado Basado en Imágenes	No	Si/No	NoSupervisado	Determinista	Una imagen	Modelo Estructurado

Cuadro 1.1: Comparación de técnicas de generación de modelos faciales

### 1.3.1. Ajuste de puntos característicos

Lee, Gu y Magnenat-Thalmann (2000) describen un sistema para generar cuerpos humanos animables descritos en VRML. Dicho sistema tiene dos componentes: replicación de cuerpo y replicación de rostro, que es el que se discutirá a continuación. Para replicar un rostro se marcan puntos característicos en imágenes frontal y lateral del rostro; después se hace un mapeo de textura continuo de 360 grados. Este sistema se enfoca en lograr una apariencia convincente de la forma obtenida, aunque la superficie no sea muy exacta. Como entrada se utilizan dos fotografías del rostro, lateral y frontal, tomadas con una cámara común en condiciones simples de iluminación, como se muestra en la Figura 1.1.

Para obtener el rostro se deforma un rostro genérico con partes animables y triangulación fina en las partes curvas o móviles y burda en el resto guiándose con puntos característicos del rostro como la posición de ojos, nariz y boca, ayudándose con restricciones de deformación para evitar rostros irreales. La localización de los puntos es semiautomática: el usuario localiza manualmente algunos puntos clave y el resto se ajusta mediante transformaciones afines y técnicas iterativas. Entonces se combinan las posiciones de ambas fotografías, que representan vistas de los planos  $XY$  y  $XZ$ , para obtener las coordenadas 3D correspondientes. Estas coordenadas se mapean al rostro genérico, que será deformado semiautomáticamente para ajustarse a esos puntos, cuidando de conservar la curvatura natural en un rostro humano.

Para crear la textura se utilizan tres fotografías: la frontal y dos laterales. Como ya se tiene la posición de los puntos característicos se definen regiones basados en éstos mediante conectividades predefinidas. Estas regiones son deformadas

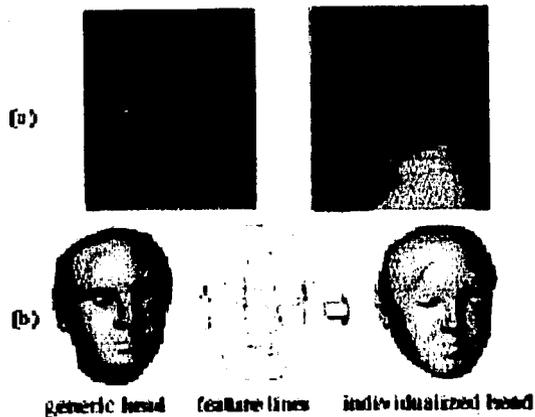


Figura 1.1: Deformación de un modelo genérico en base a puntos característicos

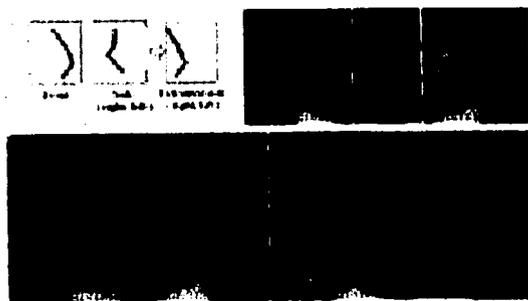


Figura 1.2: Obtención de la textura a partir de dos vistas

geoméricamente y unidas para generar una primera textura, de la que se obtiene la textura final suavizando los puntos de unión. Después se generan coordenadas de textura para cada vértice del modelo genérico deformado tomando en cuenta en qué región quedó cada uno de sus puntos característicos, según se observa en la Figura 1.2.

El modelo final es parecido al sujeto de la fotografía y la aplicación de la textura mejora la apariencia general del mismo.

### 1.3.2. Reconstrucción estéreo

Al tener dos o más imágenes del mismo objeto tomadas desde ángulos diferentes es posible utilizar el conocimiento de la posición de las cámaras para determinar las posiciones diferentes del mismo punto en todas las imágenes y obtener un mapa de profundidades. A esto se le llama reconstrucción estéreo. En Lengagne, Tarel y Monga (1996) se describe un sistema de modelado facial basado en la reconstrucción estéreo para la generación de modelos faciales.

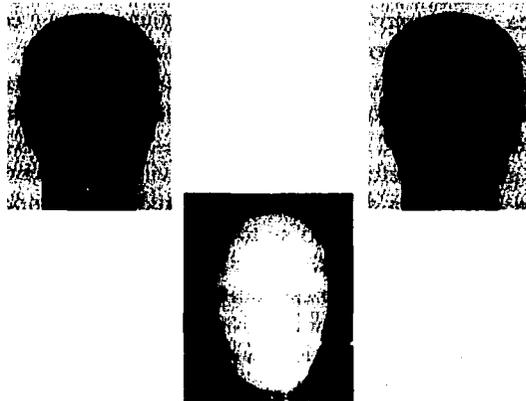


Figura 1.3: Mapa de profundidad obtenido a partir de las disparidades de un par estéreo

A partir del mapa de profundidades se obtiene un modelo facial deformable de la superficie fotografiada. Para lograr que los vértices se coloquen sobre los rasgos faciales en una forma que facilite la definición de parches faciales en vez

de la posición arbitraria dada por el mapa de profundidad, los vértices se deslizan sobre la superficie para colocarse sobre los puntos de inflexión que caracterizan a cada rasgo del rostro, cuidando además que las aristas del modelo coincidan con las líneas naturales del rostro. De este modo, es posible descomponer la superficie del rostro en parches cuadráticos que determinan un modelo analítico del rostro, a diferencia de los modelos poligonales obtenidos por la mayoría de las técnicas. Sin embargo, los autores no tratan de obtener un modelo parametrizable en general, sino que para cada rostro se genera uno según las características particulares de su superficie.

### **1.3.3. Síntesis por estimación**

En (Blanz y Vetter 1999) se describe un sintetizador/modelador facial casi automático basado en un conjunto facial mixto de 200 sujetos. Estos modelos y sus texturas fueron obtenidos con un escáner láser, por lo que tienen muy alta resolución. Este conjunto facial mixto es analizado para obtener el rostro promedio y aislar las variaciones correspondientes a cuatro segmentos del rostro: ojos, nariz, boca y el resto de la cabeza. Aplicando una combinación lineal de estas variaciones al rostro promedio es posible modelar rostros arbitrarios poniendo especial atención en suavizar los bordes donde las regiones se unen. En la Figura 1.4 se muestra al rostro promedio modificado según distintos parámetros de textura y forma.

Este modelo genérico deformable puede servir para obtener el modelo de un rostro al imitar la apariencia del rostro nuevo mediante la modificación iterativa controlada de sus parámetros, como se muestra en la Figura 1.5. Por ello a este tipo de técnicas se les llama de síntesis por estimación.

Dada una fotografía, el usuario especifica la orientación y posición del rostro, además de la principal fuente de iluminación. Después de recibir estos parámetros el sistema ya no requiere intervención del usuario. El rostro promedio se coloca en esa posición y orientación para ser dibujado, como se muestra en la Figura 1.6. Su apariencia es comparada contra la del rostro deseado tal como aparece en la fotografía y se estima cómo deberán modificarse los parámetros para aumentar el parecido. Este proceso es iterativo, deformando simultáneamente la superficie y la textura del rostro promedio para que su apariencia se parezca a la del rostro en la imagen de entrada. Cuando el parecido es suficientemente bueno (lo que se decide con una función de similitud) el proceso se detiene y se entrega como resultado el modelo y la textura obtenidos con la última parametrización.

Los resultados son de gran calidad, visualmente idénticos al rostro de entrada,

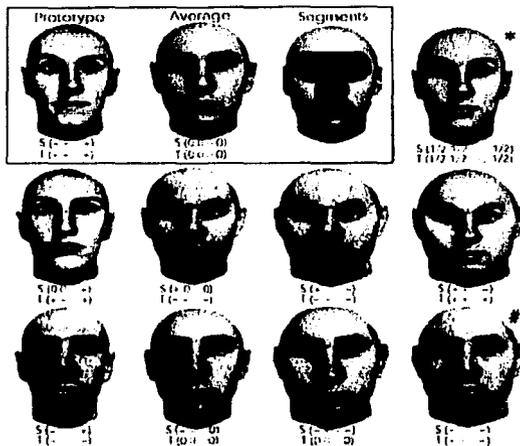


Figura 1.4: Rostros obtenidos a partir de diferentes parametrizaciones de un modelo genérico deformable

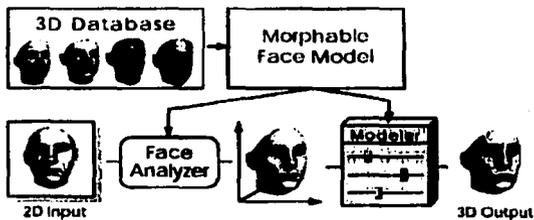


Figura 1.5: El análisis de un conjunto facial mixto permite generar rostros nuevos



Figura 1.6: El rostro promedio es sobrepuesto en la imagen original para comenzar la aproximación

y al tener el modelo es posible cambiar la expresión del individuo y las condiciones de iluminación.

### **1.3.4. Modelado basado en imágenes**

A las técnicas de este grupo también se les llama basadas en apariencia, porque determinan la superficie del objeto a partir de su aspecto. La técnica propuesta en este trabajo, los rostros característicos con datos mixtos, pertenece a ésta categoría, ya que asume que existe una correspondencia lineal entre la apariencia (imagen) de un rostro y su superficie (modelo 3D). Analizando un conjunto facial de inicialización esta correspondencia permite obtener un conjunto generador de rostros 3D. La técnica propuesta en este trabajo es una extensión de una técnica de generación de imágenes, los rostros característicos.

En la siguiente sección se describe la técnica de rostros característicos, basada en el Análisis de Componentes Principales.

En el siguiente capítulo se describirá a detalle la extensión propuesta en este trabajo, los rostros característicos 3D con datos mixtos.

## **1.4. Análisis de Componentes Principales aplicado a rostros**

La técnica de los rostros característicos consiste en aplicar el Análisis de Componentes Principales (ACP) a un conjunto de imágenes de rostros. Siendo la

base de la técnica propuesta en este trabajo es necesario detallar sus características.

### 1.4.1. Rostros Característicos

La técnica de rostros característicos (*eigenfaces*) fue planteada por Turk y Pentland (1991) originalmente para tareas de reconocimiento facial mediante la representación eficiente de imágenes faciales obtenidas por el ACP, pero su flexibilidad ha permitido su uso para otras tareas, entre ellas la generación. Su nombre se debe a que para realizar el ACP es necesario obtener los vectores característicos (*eigenvectors*) de una matriz.

Para comprender la utilidad de un cambio de representación hay que recordar que existen dos formas básicas de representar un vector de datos:

- Como una  $n$ -ada de valores (la representación usual).
- Como el conjunto ordenado de los coeficientes de la combinación lineal de elementos de una base generadora que lo determina. A este conjunto de coeficientes se le llama isomorfismo o mapeo de coordenadas. En la terminología de los rostros característicos se le conoce como *codificación*.

Tratándose de vectores en  $R^n$  en la base canónica y para fines prácticos, el vector y su mapeo de coordenadas son idénticos. Solamente cuando se trata con conjuntos generadores diferentes de la base canónica se observa la conveniencia de utilizar el mapeo de coordenadas para representar un vector.

Aunque es posible utilizar conjuntos generadores que no sean una base generadora propiamente dicha para expresar un vector, tener una base generadora ortonormal tiene grandes ventajas, una de ellas es que para obtener un mapeo de coordenadas basta con multiplicar el vector por la matriz cuyas columnas son la nueva base.

#### Características deseables de un cambio de representación

La nueva base generadora estará determinada por el conjunto de vectores que se desea representar con ella.

Si se trata de representar un subespacio vectorial, la nueva base generadora podrá tener menos elementos que la base canónica del espacio vectorial subyacente. En el caso de imágenes de rostros es evidente que se trata de un subconjunto relativamente pequeño del espacio de las imágenes en general —por ejemplo la forma oval de los rostros humanos hace innecesario generar píxeles individuales en las

esquinas de las imágenes. Para caracterizar al subespacio vectorial a representar se utilizará un conjunto finito de vectores proporcionado por el usuario. Esto permite garantizar que la nueva base podrá representar correctamente los datos conocidos y si éstos son lo suficiente representativos, al subespacio al que pertenecen.

Si todos los vectores en el conjunto son linealmente independientes, la nueva base generadora deberá tener tantos elementos como vectores existen en el conjunto. Este hecho se puede demostrar fácilmente notando que todo conjunto puede generarse como combinación lineal de sus propios elementos.

Otra característica importante del subespacio vectorial a representar es su linealidad, que permitirá representar a sus elementos como una combinación lineal de los elementos de la nueva base generadora. En el caso general de imágenes de rostros, esta suposición es falsa debido a diferencias en posición, orientación y proporción. Por ejemplo, si la posición de los ojos no coincide en todos los rostros a representar, forzosamente la posición de los píxeles de los ojos tampoco coincidirá en todos los elementos de la base generadora, por lo que algunos de los rostros generados con ella tendrán cuatro ojos de apariencia velada. La solución es sencilla para la mayoría de los casos: normalizar geoméricamente los rostros haciendo que ocupen la misma posición y orientación en la imagen, por ejemplo enderezando y recortándolos. La solución al problema de proporción no es tan sencilla -simplemente escalar los rostros para ser del mismo tamaño no basta. En este caso se deforman los rasgos del rostro para forzarlos a tener un tamaño y área semejantes. La necesidad de estas deformaciones es evidente si se considera a las personas de labios gruesos y labios delgados con diferente distancia entre los labios y la nariz. Como los rostros utilizados para este trabajo son relativamente parecidos simplemente se deformaron uniformemente para centrar los ojos en la imagen, escalando por separado la parte superior e inferior del rostro.

Dado un conjunto linealmente independiente de datos, el ACP encuentra una base ortonormal que lo expresa perfectamente y es adecuado para el subespacio al que pertenece el conjunto de entrada.

#### **1.4.2. Análisis de Componentes Principales**

La importancia del ACP es que para cualquier conjunto linealmente independiente de datos encuentra una base ortonormal que captura la variación entre los datos proporcionados sin intervención del usuario.

Al capturar la totalidad de la variación entre los datos es posible analizar un subespacio usando solamente a sus elementos más representativos; en el caso de imágenes de rostros es impráctico (por no decir imposible) procesar todos los

rostros existentes.

Por otra parte, el usuario no debe estimar o proponer la estructura del subespacio que debe analizar, con el consiguiente riesgo de error, sino que el ACP encuentra matemáticamente una base generadora ortonormal, cuyas propiedades generales son conocidas.

A partir de esta base generadora se obtienen codificaciones (mapeos de coordenadas) de los datos originales. El tamaño de las codificaciones es mucho menor que el de los datos originales, lo cual facilita su comparación, transmisión y almacenamiento. Por definición, se puede recuperar el dato original usando su codificación.

La aportación de este trabajo es mostrar cómo se puede recuperar un dato de dimensiones diferentes del dato original (un modelo 3D a partir de la codificación de una imagen 2D) pero con algo en común: ambos representan al rostro de la misma persona.

### Planteamiento del ACP

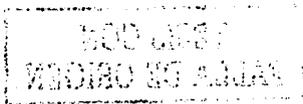
Sea  $X = (x_1|x_2|\dots|x_k)$  la matriz cuyas columnas son los vectores normalizados que representan el subespacio que se desea analizar. Cada  $x_i$  tiene  $p$  elementos. Por lo general  $p$  será mucho mayor que  $k$ .

Existen dos matrices que proveen información sobre  $X$ : el producto interno,  $X^T X$ , y el producto externo,  $XX^T$ . Los vectores característicos del producto externo son los llamados componentes principales. Calcularlos directamente es muy difícil computacionalmente, ya que el producto externo es una matriz de  $p \times p$ , por lo que se obtienen indirectamente usando el producto interno, solamente de  $k \times k$ , utilizando un resultado del álgebra lineal (ver por ejemplo Turk y Pentland (1991)) que dice:

Sea  $V = X^T X$  la matriz de vectores característicos del producto interno, denominados como  $\{v_i\}$ . Sea  $P = XX^T$  la matriz de vectores característicos del producto externo, denominados como  $\{p_i\}$ . Sea  $D$  la matriz cuya diagonal contiene los valores característicos de  $V$ , denominados por  $\{d_i\}$ .

Entonces los valores característicos de  $XX^T$  son la raíz cuadrada de los valores característicos de  $X^T X$ , y cada  $p_i = d_i^{-1/2} X v_i$ .

Los vectores característicos obtenidos, los  $\{p_i\}$  de dimensión  $p \times 1$ , forman la base generadora ortonormal para el subespacio determinado por las columnas de  $X$ .



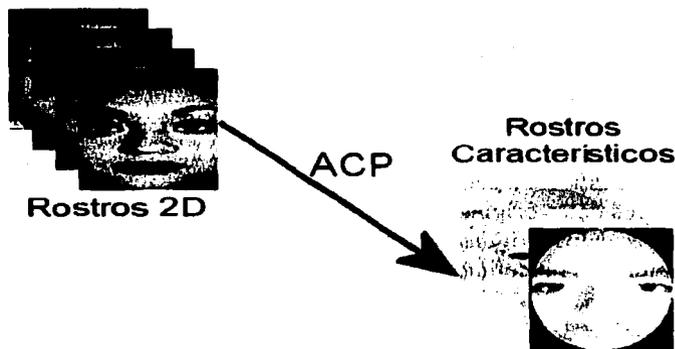


Figura 1.7: De imágenes a rostros característicos

### 1.4.3. Ejemplo de aplicación a rostros

La Figura 1.7 resume los resultados de aplicar el ACP a un conjunto de imágenes. Los rostros característicos mostrados fueron obtenidos de un conjunto de 54 imágenes pero por razones de espacio solamente se muestran los primeros 4 de ambos. Antes del procesamiento se aplicó una máscara circular a los rostros, y esta máscara puede verse en los rostros característicos obtenidos.

El procedimiento es el siguiente:

1. Cada imagen, previamente normalizada, es cubierta por una máscara circular para eliminar los píxeles en las esquinas.
2. Los valores de pixel se escriben como un vector columna, tomándolos en orden lexicográfico y agrupándolos en una matriz.
3. Cuando todas las imágenes han sido procesadas se normalizan las columnas, obteniendo con esto la matriz  $X$ . Esto es muy importante para lograr las simetrías que se requieren para obtener los vectores característicos.
4. Al tener los rostros característicos se agrupan en una matriz de columnas, la cual se utilizará tanto para codificar como decodificar rostros.

### **Codificación y decodificación de rostros**

Sea  $R^T$  un vector hilera obtenido de una imagen similar a las utilizadas en el conjunto de inicialización y procesada de la misma manera. Sea  $E$  la matriz con los rostros característicos en sus columnas. La codificación  $C = (c_1, c_2, \dots, c_k)$  de  $R$  está dada por  $R^T E$ . La decodificación está dada por  $EC^T$ .

Los rostros característicos no sólo generan, por combinación lineal, rostros en el conjunto original sino que permiten generar rostros fuera de éste dentro de ciertos límites. Como puede observarse en la Figura 1.7, el primer rostro característico, designado como tal por ser el más representativo de la totalidad del conjunto, tiene la apariencia de un rostro normal, lo cual indica que el conjunto facial utilizado es representativo del subespacio facial. Si hubiera habido algún rostro diferente a los demás, por ejemplo un rostro de perfil, el primer rostro característico tendría un "fantasma" del rostro de perfil sobre un rostro frontal.

Como el resto de los rostros característicos expresan una menor variación en el conjunto facial su apariencia es velada, aunque en la mayoría de los casos son reconocibles como rostros de algún tipo.

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that proper record-keeping is essential for the success of any business and for the protection of the interests of all parties involved. The document outlines the various methods and systems that can be used to ensure the accuracy and reliability of financial records.

The second part of the document provides a detailed overview of the accounting process, from the initial recording of transactions to the final preparation of financial statements. It covers the various steps involved in the accounting cycle, including the identification of transactions, the recording of transactions in the journal, the posting of transactions to the ledger, and the preparation of trial balances and financial statements.

The third part of the document discusses the importance of internal controls and the role of the auditor. It explains how internal controls can be used to prevent and detect errors and fraud, and how the auditor's role is to provide an independent opinion on the accuracy and reliability of the financial statements.

The fourth part of the document provides a summary of the key points discussed in the document and offers some final thoughts on the importance of maintaining accurate records and the role of the auditor. It concludes by stating that proper record-keeping and internal controls are essential for the success of any business and for the protection of the interests of all parties involved.

## **Capítulo 2**

# **Generación de modelos 3D usando rostros característicos**

En este capítulo se detalla la principal aportación de este trabajo, la técnica de rostros característicos con datos mixtos, llamada así por ser una extensión de los rostros característicos para permitir la generación de modelos faciales 3D a partir de imágenes.

### **2.1. Planteamiento del problema**

Los rostros característicos han mostrado su efectividad en la manipulación de rostros, pero están limitados a codificar y decodificar datos del mismo tipo. Se dice datos porque el Análisis de Componentes Principales ha sido utilizado con éxito en el procesamiento de datos de varios tipos: imágenes, modelos, grupos de parámetros, tensores, etcétera, pero en todos los casos en que se realiza generación el tipo de salida es el mismo que el de entrada.

En el caso de rostros humanos, si existe una relación entre apariencia y forma, esa relación debería poder plantearse de modo que sea posible expresar ambos usando los mismos parámetros. Pero este objetivo es demasiado amplio.

Por ello este trabajo se restringe a contestar la siguiente pregunta: ¿es posible utilizar la codificación de la imagen de un rostro, obtenida con los rostros característicos, para generar el modelo facial 3D correspondiente?

### **2.1.1. Limitaciones de los rostros característicos**

Hay dos grandes formas de proceder para resolver este problema: procesando rostros 2D y 3D por separado o como una unidad. A continuación se discute cada opción, asumiendo por supuesto que los conjuntos de rostros contienen a las mismas personas.

#### **Rostros 2D y 3D procesados por separado**

En este caso se obtienen dos conjuntos de rostros característicos, uno para los rostros de cada dimensión. Las codificaciones son del mismo tamaño, dado que el número de rostros es el mismo en ambos conjuntos, pero es evidente que no son idénticas.

Para obtener un rostro 3D a partir de una codificación 2D es necesario dar una función no trivial entre las codificaciones 2D y 3D que preserve la correspondencia de las codificaciones de los rostros analizados. Aún suponiendo que existiese una función así, queda la pregunta de si esa función será válida para todas las codificaciones posibles. Más aún: ¿es lineal, como lo son por separado ambos conjuntos de rostros característicos? La linealidad es una de las grandes ventajas del uso del ACP, y aunque ambos conjuntos tienen el mismo número de rostros característicos, éstos pertenecen a subespacios muy diferentes.

Es probable que exista una relación que tenga todas estas propiedades, pero la respuesta queda fuera del alcance de este trabajo. Por otra parte, uno de los factores de mayor influencia en la elección de los rostros característicos como técnica de generación fue su sencillez y esa sencillez debería conservarse.

#### **Rostros 2D y 3D procesados como una sola entidad**

En este caso todos los datos de imagen y de modelo 3D se agregan en un solo vector para cada rostro a procesar, y se encuentra un problema diferente: ¿cómo obtener la codificación de una imagen para la que no se tiene el modelo 3D? Una solución rápida es poner en cero (o en cualquier valor arbitrario) los valores correspondientes a la información del modelo y codificar este rostro con apariencia pero sin forma. Pero entonces ¿cómo obtener un modelo facial a partir de una codificación que no contiene ninguna información del modelo a construir? El modelo resultante tendrá los valores arbitrarios que se usaron para "completar" la imagen de entrada, lo cual es claramente incorrecto.

En este caso no es necesario mapear o intercambiar codificaciones: sólo hay una. Pero como esta codificación depende directamente de la apariencia y de

la forma del rostro no es posible obtener una a partir de la otra. Aunque existen técnicas para el ACP con datos incompletos, no sólo se perdería tiempo analizando modelos faciales que serán estimados en cada codificación, sino que se cae en un problema circular: si se tuviera la capacidad de estimar correctamente el modelo correspondiente a la imagen de un rostro, no habría necesidad de aplicar el ACP para obtener modelos 3D a partir de imágenes.

Como puede verse, la primera opción separa demasiado a los modelos 3D de las imágenes y la segunda los hace interdependientes. Lo que se requiere es un punto medio, una conexión lo suficientemente fuerte para que sea fácil usar una codificación de rostro 2D para generar uno 3D pero no tanto como requerir el conocimiento previo de ambos.

## **2.2. Bosquejo de la solución**

Lo que se necesita es la combinación de ambos enfoques: una codificación única para rostros de ambos tipos (imagen y modelo 3D) pero con un análisis por separado para evitar dependencias entre uno y otro.

A este requerimiento, que será planteado formalmente más adelante, se le llama *requerimiento de codificaciones intercambiables*. Su importancia radica en ser el punto de partida para solucionar el problema.

Pero antes de seguir adelante, es necesario plantear cuidadosamente lo que se quiere hacer. Repitiendo la pregunta hecha anteriormente: ¿es posible utilizar la codificación de la imagen de un rostro, obtenida con los rostros característicos, para generar el modelo facial 3D correspondiente?

Analizando esta pregunta se tiene que:

1. Debe obtenerse la codificación de una imagen con los rostros característicos.
2. Esta codificación se usará (tal vez con otra cosa) para generar un modelo 3D.
3. No se pide usar los mismos rostros característicos para obtener la codificación y decodificarla. De hecho, no se pide usar rostros característicos de ningún tipo para obtener un modelo facial 3D –solamente se pide usar la codificación.
4. No se pide la codificación de los modelos faciales 3D.

Entonces no se requiere obtener rostros característicos, independientes o comunes, para los modelos faciales 3D —solamente para las imágenes de rostros.

Pero si se debe usar la codificación de una imagen para obtener un modelo facial 3D —entonces debe existir una forma de generar modelos faciales 3D como combinación lineal de los coeficientes de la codificación. Esto último no se deduce directamente de la pregunta, pero sería la solución congruente con el uso de rostros característicos para las imágenes, además de que usar la codificación tal cual facilita el uso de la técnica.

Habiendo acotado la respuesta por su pregunta, queda la interrogante de si es posible cumplir con estas condiciones.

### 2.2.1. Requerimiento de codificaciones intercambiables

Para formalizar estas ideas, se escribe el problema a resolver en notación matemática:

Sea  $I$  la imagen de un rostro. Sea  $D$  el modelo facial 3D correspondiente. Sea  $R$  la matriz de rostros característicos asociada a  $I$ . Sea  $C$  la codificación de  $I$  en  $R$ . Sea  $G$  un conjunto generador de modelos faciales 3D a partir de una codificación  $C$ . Por razones prácticas  $G$  será escrito como una matriz.

Entonces la pregunta es: dados  $C$  y  $G$ , ¿es cierto que  $D = CG$ ?

Pero el que  $C$  sea la codificación de  $I$  significa algo importante:  $I = CR$ .

Por lo que la pregunta, en su versión final, queda planteada como:

$$¿I = CR \Rightarrow D = CG?$$

Esta implicación es el requerimiento de codificaciones intercambiables mencionado con anterioridad. Al asumir su cumplimiento se deduce la solución al problema.

Examinando la expresión matemática del requerimiento de codificaciones intercambiables se observa que la “variable dependiente”, por llamarla de algún modo, es  $G$ .

$I$  es la imagen dada por el usuario.  $C$  está determinada por  $I$ , y para cada  $I$  existe una  $D$  que es la salida “correcta”.  $G$  no tiene mayor restricción que hacer cumplir la igualdad por lo menos para los rostros del conjunto de inicialización, la misma garantía que se le pide a un conjunto de rostros característicos.

Pero, ¿cómo calcular  $G$ ?  $G$  debe tener el mismo número de elementos que la codificación, ya que de lo contrario no es posible hacer una combinación lineal. Pero la codificación tiene tantos coeficientes como rostros característicos, y a su

vez el número de rostros característicos coincide con el número de rostros en el conjunto facial de inicialización. Por lo tanto,  $G$  debe cumplir con el siguiente grupo de implicaciones:

$$\begin{aligned} I_1 = C1R &\Rightarrow D_1 = C1G \\ I_2 = C2R &\Rightarrow D_2 = C2G \\ &\dots \\ I_K = CKR &\Rightarrow D_K = CKG \end{aligned}$$

donde  $K$  es el número de rostros. El lado izquierdo de la implicación es independiente y por lo tanto se omite, dejando solamente las ecuaciones del lado derecho:

$$\begin{aligned} D_1 &= C1G \\ D_2 &= C2G \\ &\dots \\ D_K &= CKG \end{aligned}$$

Dado que las  $C_i$  y  $G$  son conjuntos de elementos, expandimos la expresión:

$$\begin{aligned} D_1 &= C1_1G_1 + C1_2G_2 + \dots + C1_KG_k \\ D_2 &= C2_1G_1 + C2_2G_2 + \dots + C2_KG_k \\ &\dots \\ D_K &= CK_1G_1 + CK_2G_2 + \dots + CK_KG_k \end{aligned}$$

Esto, escrito como un sistema linealmente independiente de  $K$  ecuaciones con  $K$  incógnitas parece bastante razonable, por lo que podría decirse que resolviéndolo se obtiene el conjunto generador que resuelve el problema.

Sin embargo, esto no es un sistema de ecuaciones. Cada uno de los  $G_i$  es un modelo 3D, con vértices de geometría, textura y normales. Debe serlo, porque de otra forma su combinación lineal con los escalares  $C$  no podría resultar en un modelo 3D,  $D_i$ .

Habiendo notado esto, a continuación se explica cómo obtener  $G$ .

## 2.2.2. Obtención de modelos 3D generadores

Para concretar, supóngase que ya se obtuvieron rostros característicos de un conjunto de tres imágenes de rostros ( $K = 3$ ), y para cada rostro del conjunto facial se obtuvo su correspondiente codificación.

Por simplicidad y sin pérdida de generalidad supóngase que los modelos faciales 3D texturizados correspondientes a las imágenes del conjunto facial tienen cuatro vértices de geometría, de textura y de normal. La siguiente exposición será para la obtención de los vértices de geometría de los modelos 3D generadores pero es igualmente válida para los demás tipos de vértice.

Sea  $D = \{D_1, D_2, D_3\}$  el conjunto de modelos faciales 3D, que son conocidos. Sean  $k = (k_1, k_2, k_3)$ ,  $h = (h_1, h_2, h_3)$  y  $q = (q_1, q_2, q_3)$  las codificaciones correspondientes de  $I_1, I_2$  e  $I_3$ . Se desea obtener  $G = \{G_1, G_2, G_3\}$ . Lo único que se sabe de los  $G_i$  es que son isomorfos topológicamente con los modelos  $D_i$  -tienen el mismo número de vértices, unidos por las mismas conectividades, y la posición relativa en el rostro es la misma para todos los vértices correspondientes. Esto se deriva del hecho de que los  $G_i$  generan por combinación lineal a los  $D_i$ .

Como ya se vio, el requerimiento de codificaciones intercambiables se expresa por  $I = CR \Rightarrow D = CG$ . Para la primera imagen facial el requerimiento es  $I_1 = kR \Rightarrow D_1 = kG$ . Cambiando el orden de los términos del lado derecho y sustituyendo, se tiene

$$k_1 \begin{pmatrix} G_1 \\ x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \\ x_3 \\ y_3 \\ z_3 \\ x_4 \\ y_4 \\ z_4 \end{pmatrix} + k_2 \begin{pmatrix} G_2 \\ x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \\ x_3 \\ y_3 \\ z_3 \\ x_4 \\ y_4 \\ z_4 \end{pmatrix} + k_3 \begin{pmatrix} G_3 \\ x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \\ x_3 \\ y_3 \\ z_3 \\ x_4 \\ y_4 \\ z_4 \end{pmatrix} = \begin{pmatrix} D_1 \\ x_1 \\ y_1 \\ z_1 \\ x_2 \\ y_2 \\ z_2 \\ x_3 \\ y_3 \\ z_3 \\ x_4 \\ y_4 \\ z_4 \end{pmatrix}$$

donde los componentes de los vértices fueron arreglados en una sola columna. Tomando la parte correspondiente a uno de los vértices, en este caso el segundo,

se tiene:

$$k_1 \begin{pmatrix} G_1 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix} + k_2 \begin{pmatrix} G_2 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix} + k_3 \begin{pmatrix} G_3 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} D_1 \\ x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

que puede escribirse como un verdadero sistema de 3 ecuaciones con 9 incógnitas:

$$k_1 G_1 [x_2] + k_2 G_2 [x_2] + k_3 G_3 [x_2] = D_1 [x_2]$$

$$k_1 G_1 [y_2] + k_2 G_2 [y_2] + k_3 G_3 [y_2] = D_1 [y_2]$$

$$k_1 G_1 [z_2] + k_2 G_2 [z_2] + k_3 G_3 [z_2] = D_1 [z_2]$$

Este sistema no puede resolverse por la disparidad entre el número de ecuaciones y de variables. Sin embargo, si escribimos todos los  $K$  sistemas generados por el mismo vértice bajo codificaciones diferentes, se tiene

$$k_1 G_1 [x_2] + k_2 G_2 [x_2] + k_3 G_3 [x_2] = D_1 [x_2]$$

$$\mathbf{k}_1 \mathbf{G}_1 [y_2] + \mathbf{k}_2 \mathbf{G}_2 [y_2] + \mathbf{k}_3 \mathbf{G}_3 [y_2] = \mathbf{D}_1 [y_2]$$

$$k_1 G_1 [z_2] + k_2 G_2 [z_2] + k_3 G_3 [z_2] = D_1 [z_2]$$

$$h_1 G_1 [x_2] + h_2 G_2 [x_2] + h_3 G_3 [x_2] = D_2 [x_2]$$

$$\mathbf{h}_1 \mathbf{G}_1 [y_2] + \mathbf{h}_2 \mathbf{G}_2 [y_2] + \mathbf{h}_3 \mathbf{G}_3 [y_2] = \mathbf{D}_2 [y_2]$$

$$h_1 G_1 [z_2] + h_2 G_2 [z_2] + h_3 G_3 [z_2] = D_2 [z_2]$$

$$q_1 G_1 [x_2] + q_2 G_2 [x_2] + q_3 G_3 [x_2] = D_3 [x_2]$$

$$\mathbf{q}_1 \mathbf{G}_1 [y_2] + \mathbf{q}_2 \mathbf{G}_2 [y_2] + \mathbf{q}_3 \mathbf{G}_3 [y_2] = \mathbf{D}_3 [y_2]$$

$$q_1 G_1 [z_2] + q_2 G_2 [z_2] + q_3 G_3 [z_2] = D_3 [z_2]$$

Seleccionando las ecuaciones correspondientes a la segunda ordenada del segundo vértice (en negrita) se obtiene un sistema determinado de 3 variables:

$$\mathbf{k}_1 \mathbf{G}_1 [y_2] + \mathbf{k}_2 \mathbf{G}_2 [y_2] + \mathbf{k}_3 \mathbf{G}_3 [y_2] = \mathbf{D}_1 [y_2]$$

$$\mathbf{h}_1 \mathbf{G}_1 [y_2] + \mathbf{h}_2 \mathbf{G}_2 [y_2] + \mathbf{h}_3 \mathbf{G}_3 [y_2] = \mathbf{D}_2 [y_2]$$

$$\mathbf{q}_1 \mathbf{G}_1 [y_2] + \mathbf{q}_2 \mathbf{G}_2 [y_2] + \mathbf{q}_3 \mathbf{G}_3 [y_2] = \mathbf{D}_3 [y_2]$$

Al resolver este sistema se obtienen los valores de la segunda ordenada del segundo vértice para todos los modelos generadores,  $G_1 [y_2]$ ,  $G_2 [y_2]$  y  $G_3 [y_2]$ . Agrupando las otras ecuaciones en forma similar se obtienen  $G_1 [x_2]$ ,  $G_2 [x_2]$ ,  $G_3 [x_2]$ ,  $G_1 [z_2]$ ,  $G_2 [z_2]$  y  $G_3 [z_2]$ , con lo que se completa el segundo vértice de geometría de todos los modelos generadores. Usando el mismo procedimiento se obtiene el resto de los vértices.

Es importante notar por qué estos sistemas siempre serán determinados. Si se tienen  $K$  rostros, habrá  $K$  vértices  $i$ -ésimos y codificaciones de  $K$  elementos porque habrá también  $K$  rostros característicos. Cada vértice genera  $d$  ecuaciones, donde  $d = 3$  en el caso de vértices de geometría y normales y  $d = 2$  en el caso de vértices de textura ( $d$  es el número de ordenadas). Para obtener el valor de una ordenada, se toma una ecuación de cada grupo —de  $K$  grupos se obtienen  $K$  ecuaciones. Las  $K$  ecuaciones seleccionadas tendrán  $K$  incógnitas —son  $K$  ordenadas correspondientes combinadas con  $K$  coeficientes conocidos. Por lo tanto, siempre será un sistema determinado, sin importar la dimensión de los vértices que componen el modelo 3D o el número de rostros en el conjunto facial.

Más aún, cada ordenada se calcula independientemente de las demás, por lo que la obtención de los modelos generadores es altamente paralelizable aunque en el presente trabajo se implementó en forma secuencial.

Las conectividades no afectan al método porque las mallas son isomorfas topológicamente, como se dijo antes.

### **2.3. Solución: Rostros Característicos con Datos Mixtos**

A la pregunta de si es posible utilizar la codificación de la imagen de un rostro, obtenida con los rostros característicos, para generar el modelo facial 3D correspondiente se responde con la técnica de *rostros característicos con datos mixtos*, la principal aportación de este trabajo.

Se llama así por ser una extensión de la técnica de rostros característicos y por utilizar un conjunto facial de inicialización mixto (2D y 3D).

A partir del requerimiento de codificaciones intercambiables se obtienen modelos generadores que por construcción proporcionan modelos correctos para los rostros en el conjunto facial de inicialización, pero al tratarse de un conjunto generador no trivial también permiten generar modelos para imágenes no analizadas previamente. Es importante notar que estos modelos generadores no son una base generadora ortonormal como la proporcionada por los rostros característicos; sin

embargo bastan para cumplir con las tareas de generación.

Los rostros característicos con datos mixtos heredan muchas de las ventajas y desventajas de la técnica original y tienen las propias:

### **Ventajas**

- **Sencillez.** La obtención de una codificación solamente requiere una multiplicación de matrices, y la generación de un modelo se hace con una combinación lineal.
- **Rapidez.** Una vez calculados los rostros característicos y los modelos generadores la obtención de un modelo a partir de una imagen en una PC de escritorio toma segundos.
- **Completez.** A partir de una sola codificación se obtienen vértices de geometría, vértices de textura y normales.
- **Paralelismo.** La obtención de los modelos generadores y su combinación lineal para generar un modelo facial 3D son altamente paralelizables.

### **Desventajas**

- **Sensibilidad a oclusiones.** Al trabajar sobre la totalidad de la entrada, los rostros característicos son muy sensibles a obstrucciones parciales.
- **Requiere normalizaciones geométricas.** Las imágenes de rostros a procesar deben estar normalizadas por su tamaño, posición y orientación como las del conjunto de inicialización. Aunque estas normalizaciones pueden realizarse semiautomáticamente, no es posible cambiar las características del conjunto facial sin recalcular rostros característicos.

El resto de este trabajo analizará los resultados prácticos de los rostros característicos con datos mixtos para determinar su viabilidad como una técnica de generación de modelos faciales 3D.



## **Capítulo 3**

# **Implementación**

Para mostrar la viabilidad práctica de la técnica desarrollada en este trabajo, los rostros característicos con datos mixtos, se realizó una implementación en lenguaje Java, descrita en este capítulo. También se describen los requerimientos prácticos de uso de la técnica, principalmente los concernientes a la obtención de un buen conjunto facial de inicialización. Los detalles de uso, diseño y programación se encuentran en el apéndice correspondiente.

### **3.1. Características del conjunto facial de inicialización**

Para la inicialización se requiere un conjunto facial mixto, llamado así porque para cada persona contiene tanto una imagen de su rostro como el modelo facial correspondiente. La obtención de este conjunto puede ser un problema en sí mismo, como se verá a continuación.

#### **3.1.1. Imágenes Faciales**

Las imágenes del conjunto facial de inicialización deberán tener las siguientes características:

1. El sujeto mirando al frente con expresión neutra
2. Una sola fuente de luz, de frente al rostro

3. Abarcar solamente el área del rostro que va desde la barbilla hasta el nacimiento del cabello verticalmente y desde el borde exterior de un ojo hasta el otro horizontalmente. Esta región será llamada *máscara facial*.
4. Haber sido escalada de modo que los ojos estén centrados verticalmente
5. Estar en tonos de gris
6. Todas las imágenes deberán tener el mismo tamaño, fijado previamente.
7. El cabello no debe ocultar el rostro. En el caso de los hombres, no tener barba o bigote.

Con esto se proveen las normalizaciones geométricas requeridas por los rostros característicos. La iluminación es frontal para minimizar las sombras. La máscara facial sirve para delimitar los rasgos faciales y evitar zonas en ángulos inconvenientes. Aunque no es estrictamente necesario que las imágenes estén en tonos de gris, es común que sea así, para analizar a la intensidad como resultado de la iluminación de la superficie del rostro. Tener la cara despejada es importante para no sesgar el análisis, lo mismo que reducir al mínimo el vello facial.

Para condicionar una buena relación entre la apariencia y la superficie de los rostros en el conjunto de inicialización, se utilizarán imágenes que también sirven como textura para los modelos faciales 3D.

### **3.1.2. Modelos Faciales**

Los modelos faciales del conjunto facial de inicialización deben tener características similares a las requeridas para las imágenes y además ser isomorfos topológicamente. Esto quiere decir que todos los modelos tienen el mismo número de vértices con las mismas conectividades, y la posición relativa en el rostro es la misma para todos los vértices correspondientes. Esto es importante porque estos modelos serán expresados como combinaciones lineales y sin estas condiciones ello no sería posible.

Obtener modelos de estas características no es fácil bajo la mayoría de métodos de modelado. En el caso de modelos faciales basados en puntos característicos los modelos más completos tienen menos de 100 vértices y por definición son isomorfos topológicamente, pero su obtención automática es un problema completo en sí mismo. Para este trabajo se utilizaron modelos densos de la superficie del rostro que fueron muestreados para obtener modelos isomorfos topológicamente



Figura 3.1: Imagen y modelo faciales antes del procesamiento

y recortados para coincidir con la máscara facial. Estos modelos tienen vértices de geometría, textura y normales.

## 3.2. Obtención del conjunto facial de inicialización

Debido a las características requeridas, es necesario tomar un conjunto facial existente y procesarlo para cumplir con los supuestos de los rostros característicos con datos mixtos.

Como es importante que las imágenes de rostros sean también las texturas de los modelos faciales, se buscó un conjunto facial con estas características. El conjunto usado en este trabajo consta de 54 personas, 31 hombres y 23 mujeres, y fue obtenido mediante una cámara 3D que implementa en hardware una variante de estéreo. De este modo se tiene una fotografía frontal del individuo y la misma fotografía sirve como textura del modelo correspondiente.

En la Figura 3.1 pueden verse la fotografía y un despliegue del modelo tal y como fueron obtenidos por la cámara 3D.

Para realizar el procesamiento de este conjunto se desarrollaron distintas herramientas que realizan las tareas descritas a continuación.

### 3.2.1. Normalización geométrica

Como la imagen y el modelo facial están delimitados por la misma máscara facial, la normalización geométrica de ambos es simultánea.

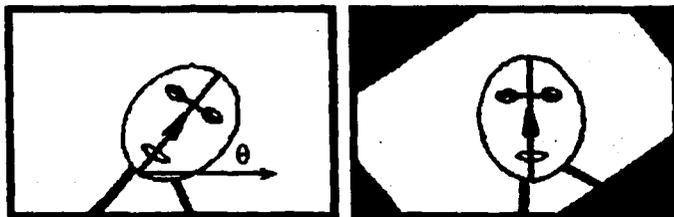


Figura 3.2: Selección de la máscara facial

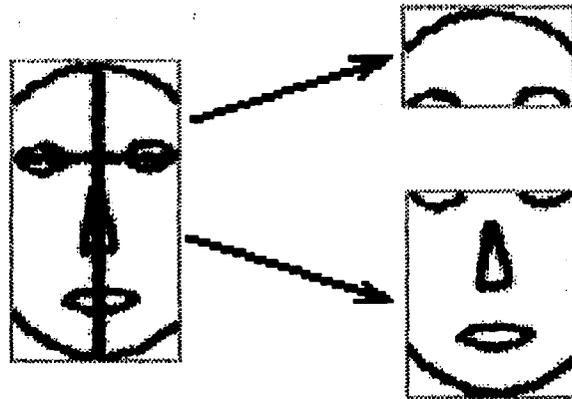
La aplicación lee el modelo facial obtenido por la cámara en formato OBJ y utiliza la textura para seleccionar la máscara facial, como se ve en la Figura 3.2.

El rombo en pantalla representa los ejes mayor y menor de la elipse que enmarca al rostro y que corresponde a la máscara facial. Los vértices del rombo son colocados en la posición correcta por el usuario usando el mouse. Al hacer click en el botón de *Crear Raster/Textura* la aplicación realiza los siguientes pasos en memoria:

1. Rota la imagen de modo que el eje que va desde la barbilla hasta el nacimiento del cabello quede vertical. El otro eje, que va de un extremo a otro de los ojos, es rotado alrededor de su centro hasta quedar horizontal.



2. Después de rotar la imagen se rotan las coordenadas de textura de la malla, para que caigan sobre las posiciones correctas en la imagen rotada.
3. Se recorta de la imagen el rectángulo que contiene a la máscara facial.
4. La imagen se divide en dos partes por el eje menor. Cada parte será procesada por separado.



5. Para cada mitad la aplicación hace lo siguiente:
  - a) Calcula a qué tamaño debe escalarse la textura original de modo que la mitad seleccionada sea de tamaño  $a \times h/2$ , donde  $a$  es el ancho deseado para la textura final, y  $h$  el alto correspondiente.
  - b) Realiza una rasterización sobre una cuadrícula que coincide con el tamaño de la textura original. Esta rasterización, realizada por interpolación lineal, tiene por objetivo llenar la cuadrícula con vértices, normales y coordenadas de textura, de modo que a cada pixel corresponda un punto muestreado de la superficie original.
  - c) Después de haber hecho esto, se recorta el pedazo correspondiente a la mitad seleccionada. El resultado es una nube de puntos donde cada vértice corresponde a cada pixel de la textura.

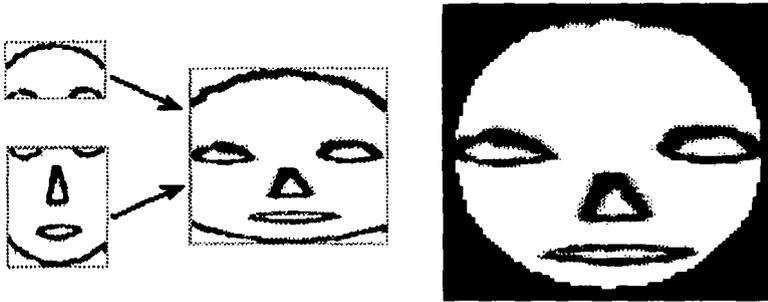


Figura 3.3: Normalización geométrica por mitades y aplicación de la máscara facial

6. Ahora se pegan las dos nubes de puntos para formar una sola nube que determina la superficie del rostro, y las dos mitades de textura para formar la imagen/textura normalizada, tal como se muestra en la Figura 3.3. Se aplica la misma máscara elíptica sobre ambos, que por tener forma cuadrada tendrá una apariencia circular. A partir de ahora no se tomarán en cuenta las regiones fuera de la máscara. Recorriendo la nube verticalmente, conectando a las hileras contiguas como bandas triangulares, se obtiene un modelo facial texturizado con una correspondencia directa con su textura, y por construcción isomorfo con el resto de los modelos obtenidos.

Después de escribir a disco la imagen y el modelo facial normalizados, la aplicación termina y se procesa el siguiente rostro.

Siguiendo este procedimiento se obtiene el conjunto facial de inicialización. Una ventaja del muestreo de modelos faciales es que permite añadir nuevos rostros al conjunto procesando rostros de bases existentes, sin importar sus convenciones de conectividad.

La textura se escribe a disco en formato PNG y por razones de eficiencia, cada modelo facial se escribe a disco en formato binario. Esto es importante porque por construcción el número de vértices en los modelos faciales será similar al número de pixels en la textura. En una PC de escritorio, la diferencia en tiempos de lectura es de 2 minutos en formato OBJ (ASCII) contra un par de segundos en formato binario.

Como los modelos son isomorfos, para ahorrar un poco de espacio las conectividades son escritas en binario por separado. Estas conectividades solamente se usarán cuando se genere un modelo facial 3D a partir de una imagen.

Teniendo un conjunto facial de inicialización adecuado, comienza el análisis.

### **3.3. Inicialización**

#### **3.3.1. Obtención de imágenes características de rostros**

Este paso se realizó en una aplicación independiente, ya que sólo requiere hacerse una vez.

Esta aplicación lee todas las imágenes de disco y genera una matriz de imágenes características de rostros que se escribe a disco en binario, usando el método descrito en la sección referente al Análisis de Componentes Principales.

Los cálculos se realizaron con las clases de álgebra lineal para matrices JAMA (JAMA: A Java Matrix Package), desarrollado en conjunto por NIST (*National Institute of Standards and Technology*) y The MathWorks, que puede obtenerse en <http://math.nist.gov/javanumerics/jama/>.

La matriz de rostros característicos es escrita a disco en formato binario para su uso posterior.

#### **3.3.2. Obtención de modelos 3D generadores de rostros**

El siguiente paso es la obtención de los modelos 3D generadores. La aplicación correspondiente toma como entrada las imágenes de rostros y utilizando los rostros característicos calculados en el paso anterior obtiene para cada imagen su codificación.

Después lee de disco los modelos faciales correspondientes a los rostros característicos y utilizando como se describió anteriormente el requerimiento de codificaciones intercambiables, obtiene la matriz de rostros generadores que se escribe a disco en un formato binario.

Para resolver los sistemas de ecuaciones lineales derivados del requerimiento de codificaciones intercambiables también se utilizó JAMA.

Con esto termina la inicialización de los rostros característicos con datos mixtos.

### **3.4. Generación de modelos faciales 3D**

Habiendo terminado la inicialización de los rostros característicos con datos mixtos ya se puede generar un modelo facial 3D a partir de una imagen.

El procedimiento es simple:

1. Se carga de disco una imagen normalizada geoméricamente y que puede ser a colores. La normalización se hace con una aplicación similar a la utilizada para obtener el conjunto facial de inicialización.  
Si es necesario, la imagen es convertida a tonos de gris.
2. La imagen es cubierta por una máscara que pone en cero los pixeles de las esquinas.
3. Los valores de pixel se toman en orden lexicográfico para formar un vector, que es normalizado para tener norma 1.
4. Los rostros característicos son leídos de disco y se multiplican por el vector normalizado para obtener una codificación.
5. La codificación obtenida sirve para combinar linealmente los rostros generadores, que también son leídos de disco, para obtener un modelo facial. Este modelo, con información de geometría, texturas y normales, es el resultado de la técnica y se escribe a disco ya sea en formato binario o VRML.

#### **Generación de texturas**

Las texturas para los modelos faciales generados pueden obtenerse de dos formas: generando una imagen usando la codificación de la imagen de entrada, lo que permite evaluar visualmente la representatividad del conjunto facial de inicialización, o simplemente copiar la imagen de entrada original, con la ventaja de que es la correcta y puede incluso estar a colores.

Para fines de estudio siempre se genera la textura a partir de la codificación de la imagen de entrada; un sistema de producción reusaría la imagen de entrada por razones de tiempo y exactitud.

## Capítulo 4

# Experimentos y Resultados

En este capítulo se describe la construcción de los distintos conjuntos de prueba, las pruebas realizadas y los resultados obtenidos.

### 4.1. Conjunto Facial de Inicialización

El conjunto facial de inicialización consta de 54 personas, 23 mujeres y 31 hombres, con edades entre los 17 y 23 años.

Se desea conocer los siguientes aspectos de la técnica:

**Sensibilidad al Ruido.** Para medir la sensibilidad de la técnica se aplicaron distintos tipos de ruido a una imagen dentro del conjunto facial, cuya generación es por definición perfecta. Se quiere demostrar que la técnica reacciona predeciblemente al ruido y es resistente a él, apoyando esta hipótesis en la resistencia conocida de los rostros característicos en su forma original.

**Reacción a Imágenes Espurias.** Esto es importante cuando no hay un supervisor humano que garantice que la entrada realmente es la imagen de un rostro. Para esto se revisa el primer coeficiente de la codificación de la imagen de entrada con los rostros característicos 2D. Este coeficiente indica el grado de pertenencia de la imagen al espacio facial representado por los rostros característicos, siendo mayor la pertenencia conforme se acerca a 1.

**Peso de Rasgos Específicos.** Para medir la importancia de los rasgos prominentes del rostro como son los ojos, la nariz y la boca se aplicaron máscaras cubriendo uno o más de ellos y se midió su influencia en el resultado final.

**Sensibilidad a las oclusiones.** Los rostros característicos son muy sensibles a las oclusiones, y por lo mismo se probaron varias oclusiones sobre las imágenes perfectas. En el caso de las oclusiones la sensibilidad de los rostros característicos es muy grande, por lo que es de esperar que esta sensibilidad se encuentre también en los rostros característicos con datos mixtos. Se generaron oclusiones en diferentes orientaciones y anchos.

**Influencia del color de las oclusiones.** El efecto de las oclusiones se ha discutido ampliamente para la técnica de rostros característicos, pero el influjo del color de la oclusión no se menciona, por lo que las pruebas realizadas servirán tanto para la técnica original como para la extensión propuesta en este trabajo. Para determinarlo cada oclusión se aplicó en tres colores: blanco, gris y negro.

**Efectos de la resolución del conjunto de inicialización.** Una resolución baja puede provocar errores por la pérdida de detalle, mientras que una resolución alta puede causar un tiempo demasiado alto y errores al procesar una gran cantidad de información innecesaria. Se probó con imágenes en dos resoluciones:  $60 \times 60$  y  $120 \times 120$  píxeles. Las comparaciones son principalmente visuales, aunque también se comparan las diferencias en las pruebas de sensibilidad al ruido y las oclusiones en términos relativos.

**Calidad Visual.** Los resultados de las distintas pruebas se despliegan en pantalla para ser inspeccionados visualmente y comprobar la fidelidad de las medidas de error además de detectar casos especiales.

**Efecto del género de los sujetos.** Para determinar la influencia del género de los sujetos en el conjunto facial de inicialización se crearon tres conjuntos de inicialización diferentes: uno sólo con mujeres, otro sólo con hombres y un tercero conteniendo a ambos.

En total se utilizaron los rostros de 23 mujeres y 31 hombres para formar los conjuntos de inicialización. Para formar los subconjuntos de prueba se tomaron 6 personas de cada conjunto de inicialización: los 3 más y menos representativos, según el valor del primer coeficiente de su codificación. Cuando el conjunto de inicialización estaba compuesto por personas del mismo género también se realizaron pruebas con las del género opuesto. En la Figura 4.1 se muestran las imágenes de algunos de los rostros en el conjunto de inicialización.

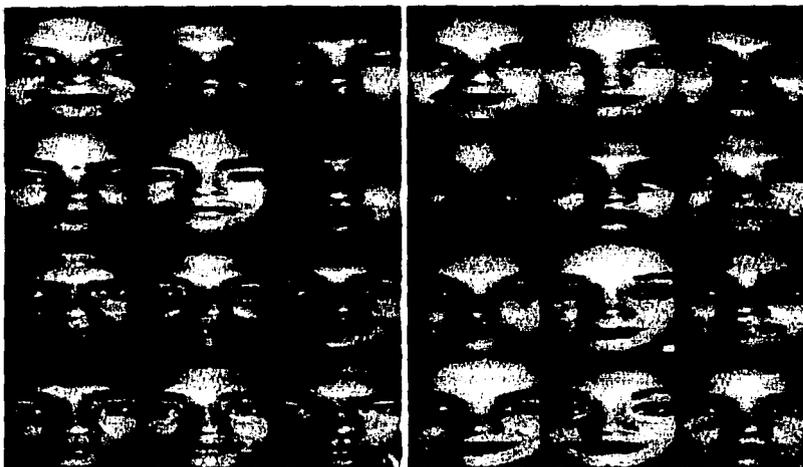


Figura 4.1: Algunas de las personas en el conjunto facial de inicialización

## 4.2. Perturbaciones a analizar

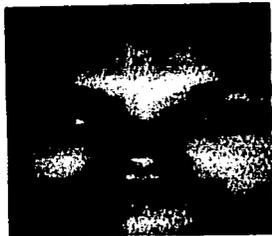
Para obtener imágenes de ruido se alteró una imagen con gris sólido (128 en una escala de 256 grises). La razón de elegir el gris sólido es la de no sesgar la norma de la imagen hacia cero o hacia el valor máximo.

Para obtener imágenes de prueba para el mejor caso, cuando la imagen pertenece al conjunto facial de inicialización, se aplicaron filtros y oclusiones a las imágenes en el conjunto, y se compara la reconstrucción obtenida contra la reconstrucción perfecta (excepto por errores de redondeo) obtenida por la imagen sin modificar.

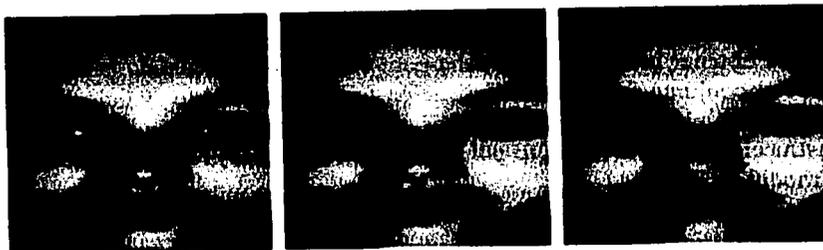
En la Figura 4.2 se muestra una de las imágenes de prueba antes de cualquier modificación.

Los filtros aplicados y sus parámetros son:

**Mediana.** Este filtro asigna a cada pixel la mediana de los valores de sus vecinos; sirve para eliminar pixeles cuyos valores difieren drásticamente de sus vecinos sin alterar demasiado el resto de la imagen. Para las pruebas se utilizan vecindades de tamaño 3, 5 y 7. La aplicación de este filtro se ilustra en la Figura 4.3.



**Figura 4.2: Imagen de prueba original**



**Figura 4.3: Filtro de medianas con aperturas de 3, 5 y 7 pixels**

**Ruido aleatorio.** Los pixels afectados son reemplazados por un valor aleatorio. El porcentaje de pixels afectados en las pruebas es de 5 % y 10 %. La aplicación de este filtro se ilustra en la Figura 4.4.



Figura 4.4: Ruido Aleatorio del 5 y 10 por ciento

**Ruido Uniforme.** Similar al ruido aleatorio, pero el valor de reemplazo es una perturbación del original para que la apariencia no cambie tanto. Se utilizaron porcentajes de 10 % y 20 % de afectación. La aplicación de este filtro se ilustra en la Figura 4.5.



Figura 4.5: Ruido Uniforme

**Promedio.** A cada pixel se le asigna el valor promedio de sus vecinos. Se utilizó una vecindad de tamaño 3. Uno de los efectos más notables de este filtro es la desaparición del brillo en los ojos aunque el resto de la imagen no cambia a simple vista. La aplicación de este filtro se ilustra en la Figura 4.6.

**Mosaico.** Este efecto es utilizado para censura, pero cuando el tamaño del mosaico es reducido puede apreciarse la imagen aunque la calidad disminuye y toma la apariencia característica de bloques que le da nombre. Se utilizaron mosaicos de  $2 \times 2$  píxeles y  $3 \times 3$  píxeles. La aplicación de este filtro se ilustra en la Figura 4.7.

**Oclusiones.** Para las oclusiones se utilizaron barras horizontales y verticales centradas en la imagen. Estas barras son de color blanco o negro sólidos, con 3 anchos: 6, 12 y 18 píxeles. La aplicación de las oclusiones se muestra en las Figuras 4.8, 4.9, 4.10, 4.11 y 4.12.

Las reconstrucciones generadas por estas entradas modificadas son comparadas contra la imagen real y la malla real para medir el error (que en este caso teóricamente es cero) y determinar el grado de corrupción que provoca cada alteración.

Además de estos resultados medibles, se realizan pruebas con fotos para las que no existe una malla 3D y por lo tanto la evaluación de los resultados será visual.

#### **Nombrado de imágenes de prueba**

Sea la imagen de prueba nombre.png. Cuando la imagen sea de 120 por 120 píxeles, se llamará nombre120.png. Conforme se apliquen las modificaciones las imágenes se nombran como dice el Cuadro 4.1.

### **4.3. Medidas**

En esta sección se describe la forma en que se midieron los errores en los rostros reconstruidos, tanto en 2D como en 3D.

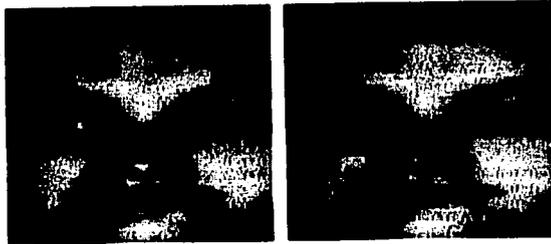
#### **Error en 2D y 3D**

Para medir el error en la reconstrucción se calcula la suma de cuadrados de diferencias entre los elementos correspondientes, valores de pixel no cubiertos por la máscara en el caso de imágenes, y la distancia geométrica o angular en el caso de vértices en los modelos 3D, según el tipo de vértice.

Además se registran las distancias mínima, máxima y promedio entre todos los pares de elementos y se calcula la desviación estándar de éstas.



**Figura 4.6: Filtro promedio con apertura de 3 pixels**



**Figura 4.7: Mosaico de 2 y 3 pixels de apertura**

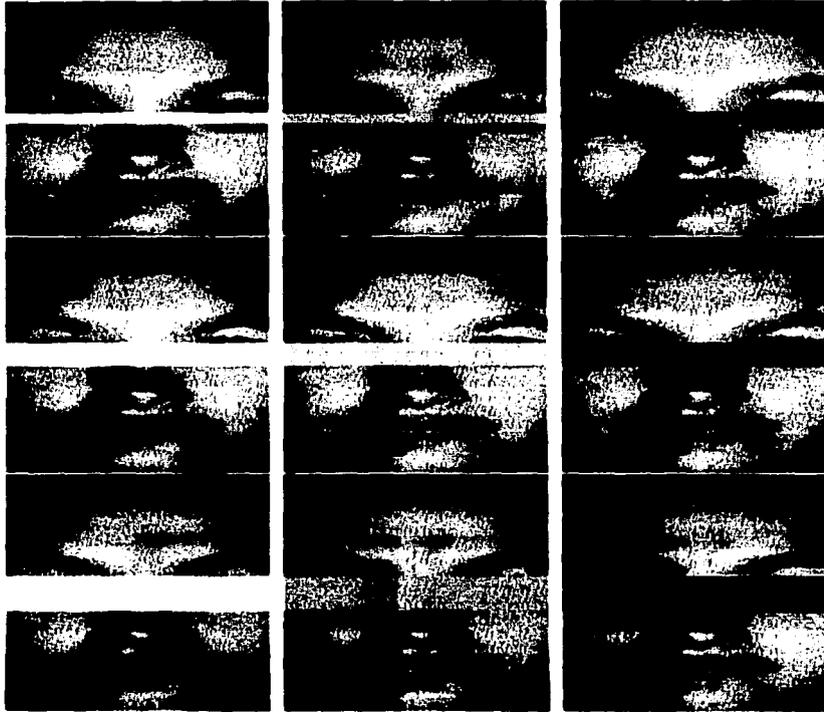
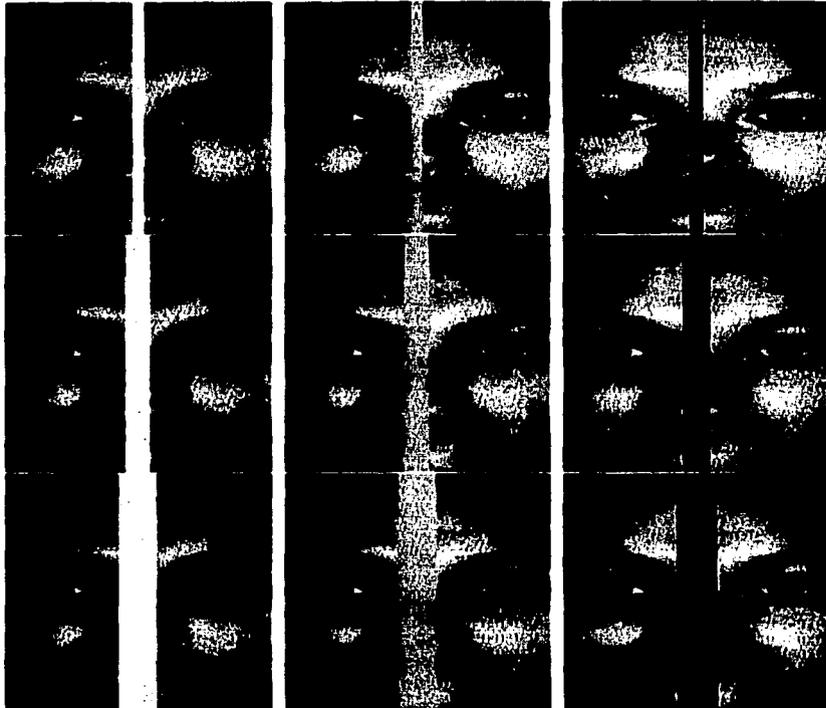
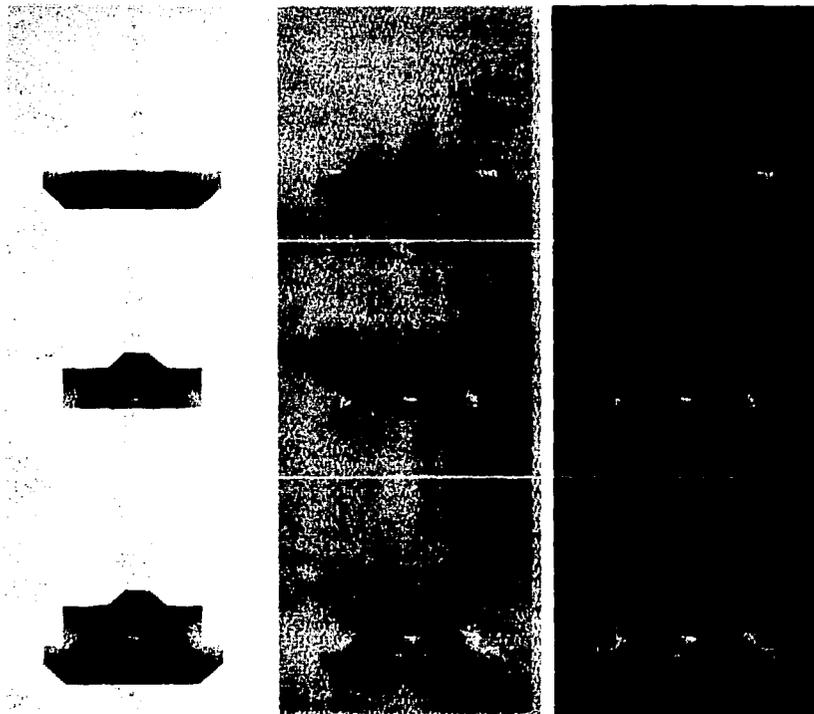


Figura 4.8: Oclusiones. Barras horizontales en tres anchos (6,12 y 18 pixels) en tres colores: blanco, gris y negro centradas en la imagen.



**Figura 4.9: Oclusiones: Barras verticales en tres anchos (6,12 y 18 pixels) en tres colores: blanco, gris y negro centradas en la imagen.**



**Figura 4.10: Máscaras en tres colores: blanco, gris y negro. De arriba a abajo, solamente son visibles la boca, la nariz y finalmente nariz y boca.**

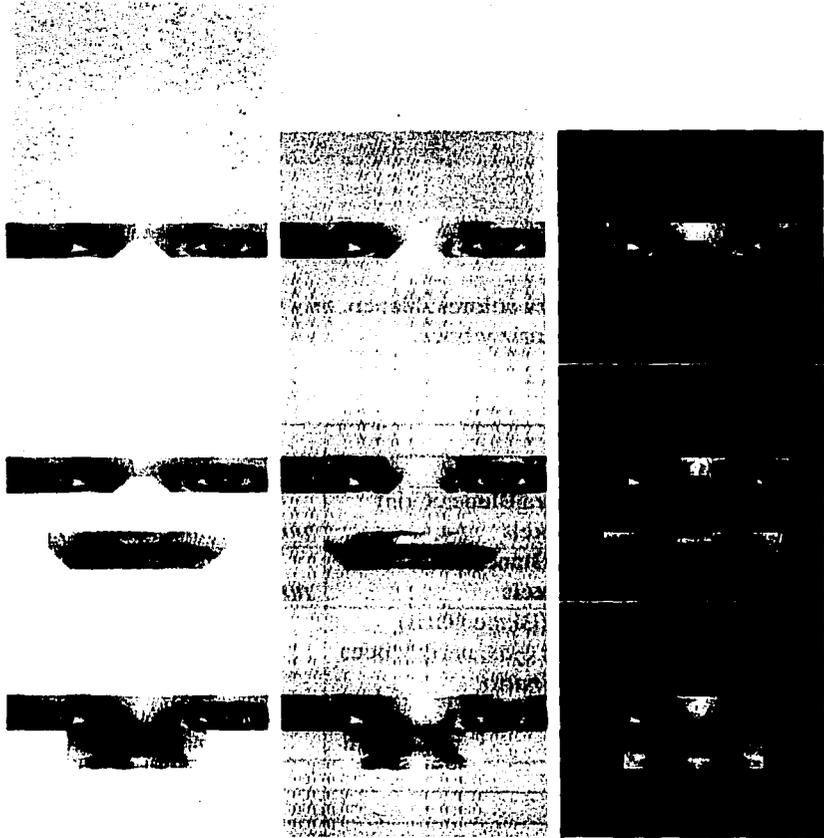


Figura 4.11: Máscaras en tres colores: blanco, gris y negro. De arriba a abajo, solamente son visibles ojos, ojos y boca, ojos y nariz



Figura 4.12: Máscaras en tres colores: blanco, gris y negro, mostrando solamente los rasgos del rostro: ojos, nariz y boca.

Descripción	Nombre de Imagen
Imagen Original	<i>nombre120.png</i>
Barra Horizontal (Negra Blanca Gris) de $pp = 06, 12$ o $18$ pixels	<i>nombre120bhpp(n b g).png</i>
Barra Vertical (Negra Blanca Gris) de $pp = 06, 12$ o $18$ pixels	<i>nombre120bvpp(n b g).png</i>
Pasamontañas (Negro Blanco Gris) donde la visibilidad de ojos, nariz y boca se indica con xyz en binario	<i>nombrexyz{n b g}.png</i>
Ruido Aleatorio 5%	<i>nombre120ra05.png</i>
Ruido Aleatorio 10%	<i>nombre120ra10.png</i>
Ruido Uniforme 10%	<i>nombre120ru10.png</i>
Ruido Uniforme 20%	<i>nombre120ru20.png</i>
Mosaico de $2 \times 2$	<i>nombre120p2.png</i>
Mosaico de $3 \times 3$	<i>nombre120p3.png</i>
Promedio de $3 \times 3$	<i>nombre120a3.png</i>
Mediana de $3 \times 3$	<i>nombre120m3.png</i>
Mediana de $5 \times 5$	<i>nombre120m5.png</i>
Mediana de $7 \times 7$	<i>nombre120m7.png</i>

Cuadro 4.1: Nombrado de las imágenes de prueba (de  $120 \times 120$  pixels)

RECEIVED  
MAY 10 2011

TESIS CON  
FALLA DE ORIGEN

### **Espacio facial: coeficientes**

El ángulo entre vectores de coeficientes indica la diferencia entre dos rostros. Además, el peso dado a cada componente principal también es útil para el análisis porque dice qué es lo que representa cada componente principal.

## **4.4. Conjuntos de prueba**

Como puede verse en el Cuadro 4.1, para cada imagen con modelo se generaron 10 perturbaciones y 39 oclusioniones, para un total de 50 imágenes de prueba. El objeto es medir la calidad de la reconstrucción en el caso ideal (sin perturbaciones) y cuando hay alteraciones (las otras 49 imágenes).

Se eligieron 6 imágenes de cada grupo de sexos: sólo hombres, sólo mujeres, hombres y mujeres juntos. Fueron las 3 más representativas, según el valor de su primer coeficiente en los rostros característicos correspondiente y las 3 menos representativas según el mismo criterio. A continuación se muestran en orden lexicográfico de más a menos representativas.

### **Conjunto de Hombres**

**Hombre1**



**Hombre2**



**Hombre3**



**Hombre4**



**Hombre5**



**Hombre6**



**Conjunto de Mujeres**

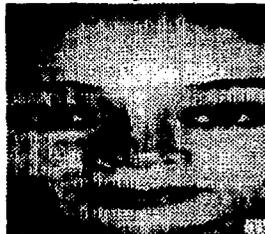
**Mujer1**



**Mujer2**



**Mujer3**



**Mujer4**



**Mujer5**



**Mujer6**



**Conjunto de Hombres y Mujeres**

**Hombre1**



**Mujer2**



**Mujer7**



**Hombre4**



**Hombre5**



**Hombre6**



**TESIS CON  
FALLA DE ORIGEN**

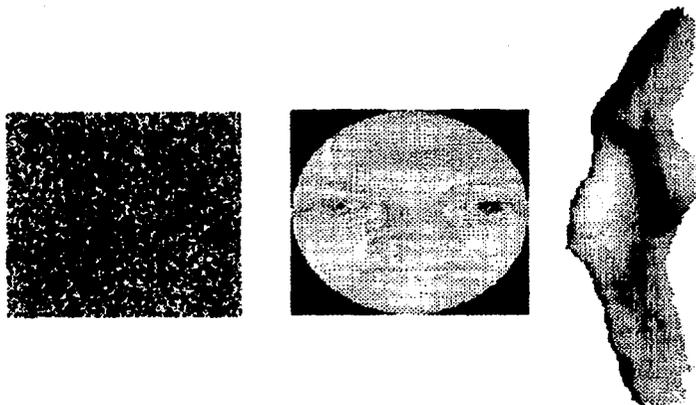


Figura 4.13: Resultados de usar una imagen espuria con el conjunto generador de hombres

## 4.5. Resultados

Se realizaron pruebas en dos dimensiones de imágenes de rostros:  $60 \times 60$  y  $120 \times 120$ , que condicionan el número de vértices en los modelos faciales. Exceptuando las diferencias debidas a un menor número de vértices, no se detectó influencia alguna en el desempeño de la técnica.

### 4.5.1. Imágenes espurias

La imagen generada es reconocible como un rostro, pero tiene una apariencia nevada claramente incorrecta. El modelo generado tiene una superficie rugosa, y aunque conserva la apariencia de un rostro, su forma no es correcta, sobre todo en el caso de los rostros generados obtenidos del conjunto facial de inicialización de hombres. En la Figura 4.13 puede verse el resultado de utilizar una imagen de ruido aleatorio con el conjunto generador obtenido a partir de los rostros de hombres. Es importante notar que el valor del primer coeficiente de la codificación nos indica que no se trata de un rostro, por lo que no es necesaria una verificación visual.

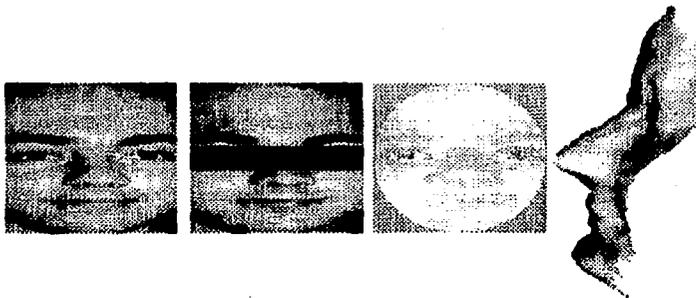


Figura 4.14: Resultados de usar una oclusión horizontal negra de 18 píxeles con el conjunto generador de hombres

#### 4.5.2. Imágenes perturbadas y ocluidas

Las imágenes perturbadas provocaron muy pocos errores, la mayoría apenas visibles, mientras que las oclusiones provocaron grandes errores, el negro más que el blanco, con el gris en tercer lugar. En el caso de las máscaras, los errores más grandes ocurrieron cuando sólo era visible la boca del sujeto.



Figura 4.15: Las dos imágenes de entrada y las generadas

#### 4.5.3. Rostros fuera del conjunto de inicialización

En este caso la valoración del resultado es solamente visual. Se muestra el resultado para un hombre y una mujer en el conjunto generador correspondientes. En la Figura 4.15 se muestran las imágenes de entrada y las generadas, mientras que en las Figuras 4.16 y 4.17 se muestran los modelos obtenidos, sin texturizar para poder apreciar la superficie.

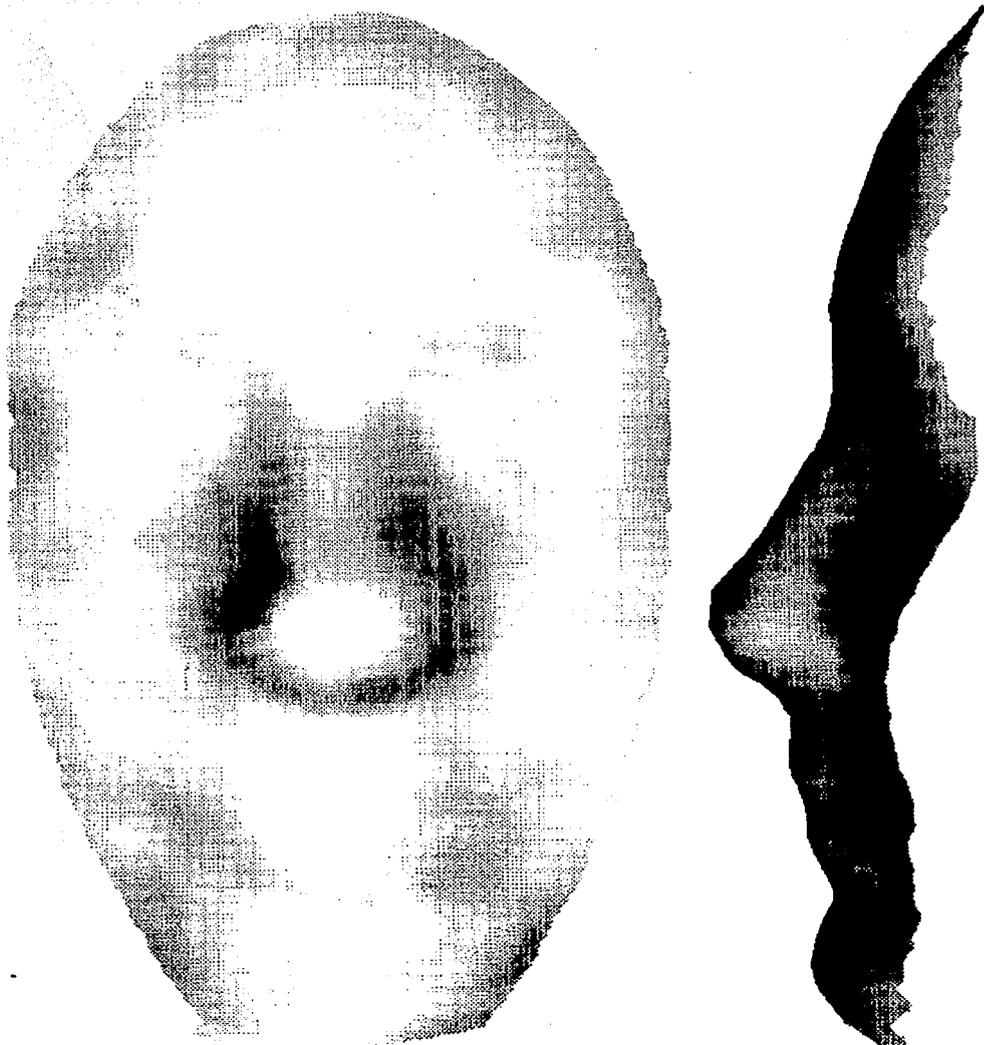
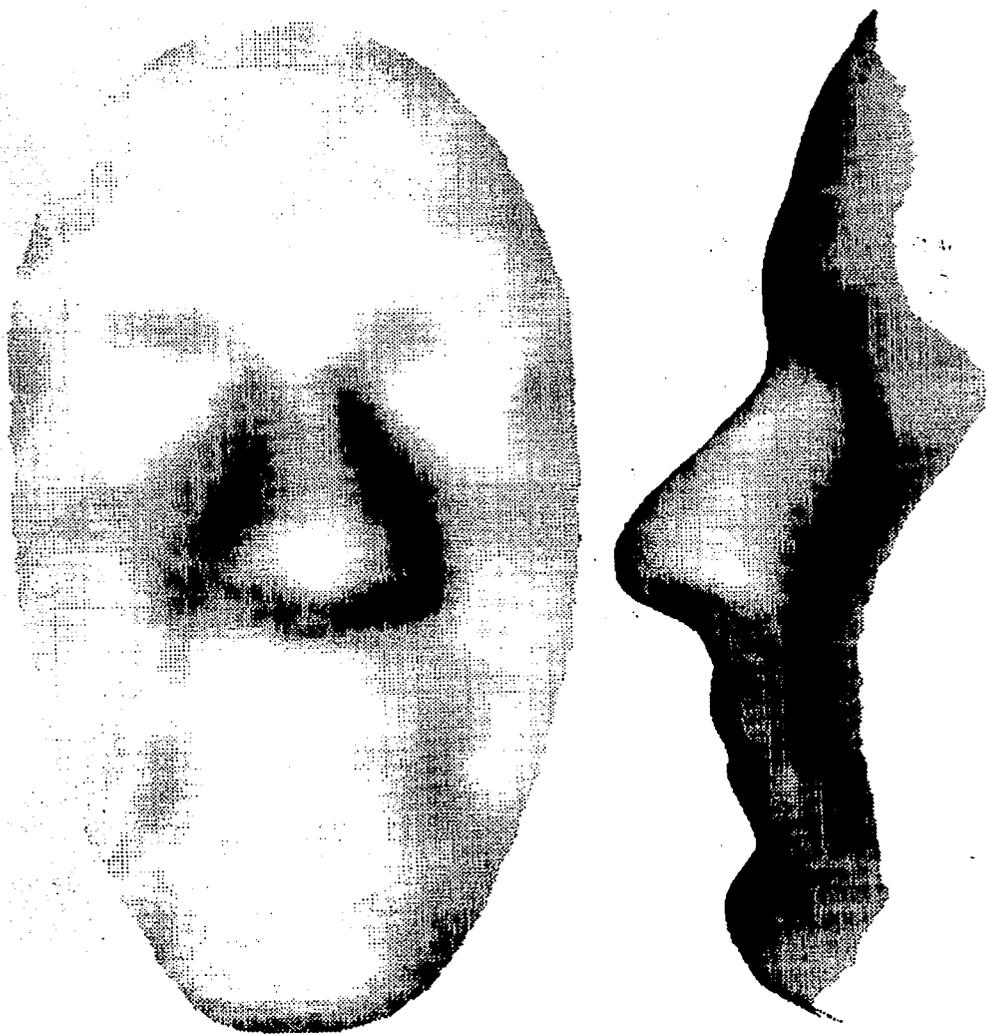


Figura 4.16: Generación de un rostro femenino fuera del conjunto de inicialización

UNIVERSIDAD  
NACIONAL AUTÓNOMA DE MÉXICO  
INSTITUTO DE INVESTIGACIONES EN ENGENNERIA Y CIENCIAS

TESIS CON  
FALLA DE ORIGEN



**Figura 4.17: Generación de un rostro masculino fuera del conjunto de inicialización**

TESIS CON  
FALLA DE ORIGEN

TESIS CON  
FALLA DE ORIGEN

## Capítulo 5

# Conclusiones y Trabajo Futuro

### 5.1. Conclusiones

La técnica de rostros característicos con datos mixtos fue bastante robusta a perturbaciones globales, mientras que fue particularmente sensible a las oclusiones. Contra lo que pudiera pensarse, el tamaño de la oclusión no es tan importante como su color: negro y blanco provocan los mayores errores. Esto se explica al recordar que para aplicar la técnica los valores de pixel de la imagen de entrada deben normalizarse a 1, por lo que un número artificialmente alto de pixeles en negro y en blanco sesga la normalización. En el caso de pixeles en gris, el gris se acerca al valor real de los pixeles ocluidos mucho mejor que el negro y el blanco, por lo que el sesgo a la normalización es menor.

Por otra parte, se descubrió que la propagación del error no es lineal, aunque sí es consistente: mayor error en la entrada genera un mayor error en la salida, recordando por supuesto que se están usando medidas globales por ser los rostros característicos una técnica global, aunque en términos generales la desviación estándar fue reducida.

Por esto se concluye que la técnica de los rostros característicos con datos mixtos es adecuada para la generación de modelos a partir de imágenes con errores comunes, no deliberados. Al igual que la técnica original, la calidad de la generación depende de la calidad de la base, aunque los errores son más notorios en la generación de imágenes que en la generación de modelos, probablemente porque los modelos generadores no son ortogonales mientras que los rostros característicos sí lo son. El primer coeficiente de la codificación es un buen discriminador para detectar automáticamente cuando la imagen no es la adecuada para generar un

buen modelo, permitiendo su uso no supervisado.

## 5.2. Trabajo Futuro

Existen muchos métodos de obtener bases generadoras para conjuntos de datos, a los cuales podrían aplicarse las ideas que dieron origen a la técnica de rostros característicos con datos mixtos propuesta en este trabajo.

Por ejemplo, el Análisis de Componentes Independientes (ACI), que trabaja en forma similar al ACP pero la base obtenida no es ortogonal. Esto es útil cuando los datos se alinean en ejes no ortogonales pero claramente diferenciados. En el caso de rostros esto ocurre cuando algunos pocos sujetos tienen una característica que los diferencia claramente de los demás pero no es anómala, como el uso de lentes.

La factorización de matrices no negativas es una técnica que genera imágenes únicamente por adición de los elementos en la base, a diferencia del ACP que permite coeficientes negativos en sus combinaciones lineales. Esta técnica enfatiza las partes de los objetos analizados.

Finalmente, los rostros característicos con datos mixtos requieren que se use la totalidad de rostros característicos obtenidos del conjunto de inicialización, lo cual limita en la práctica el número de rostros en éste. En el caso de imágenes, se sabe que no es necesario utilizar todos los rostros característicos al costo de introducir un error medible en la generación. Permitir la eliminación de algunos rostros característicos y medir el error que esto conlleva permitiría aplicar la técnica descrita en este trabajo a conjuntos de inicialización con cientos de individuos sin requerir grandes cantidades de memoria, almacenamiento y poder de cómputo.

Por otra parte, en este trabajo no se trataron las variaciones de iluminación y en la pose del sujeto, lo cual permitiría la aplicación de la técnica a un conjunto menos restringido de imágenes.

# Apéndice A

## Formato OBJ

Este formato ASCII desarrollado por Alias |Wavefront para el Advanced Visualizer se ha convertido en un formato común de transferencia aunque rara vez se utiliza todo su potencial -en este apéndice sólo se describen los detalles pertinentes a este trabajo.

Un archivo OBJ puede contener vértices de geometría (*v*), vértices de textura (*vt*) y normales de vértice (*vn*). Todos estos vértices se relacionan a través de caras poligonales, (*f*). Los vértices se representan por medio de listas para cada tipo de vértice y al momento de ser cargados los vértices de cada tipo se numeran consecutivamente a partir de uno. Los vértices se especifican en un sistema de coordenadas de mano derecha. Todos los valores son flotantes, mientras que *v*, *vn*, *vt* y *f* son etiquetas.

La sintaxis es la siguiente:

*v* *x* *y* *z* donde se da una posición en el espacio tridimensional.

*vn* *i* *j* *k* donde *i*, *j* y *k* son los componentes del vector especificado.

*vt* *u* *v* donde *u* es la dirección horizontal y *v* la vertical.

Para especificar caras poligonales se utiliza

```
f v/vn/vt v/vn/vt ...  
f v//vt v//vt ...
```

donde cada tripleta asigna una posición 3D, una normal y una coordenada de textura a cada vértice poligonal usando el número de vértice, aquí *f* es una etiqueta. La

cara poligonal puede tener cualquier número de vértices (mayor que 2) y el primero está unido al último.

Como ejemplo, sea un cuadrado orientado frente a la cámara donde los vértices se leen en orden contrario a las manecillas del reloj.

```
v 0.000000 2.000000 0.000000
v 0.000000 0.000000 0.000000
v 2.000000 0.000000 0.000000
v 2.000000 2.000000 0.000000
f 1 2 3 4
```

Para texturizar este cuadro con el archivo wood . jpg se definen los vértices de textura y se relacionan con los vértices de geometría.

```
v 0.000000 2.000000 0.000000
v 0.000000 0.000000 0.000000
v 2.000000 0.000000 0.000000
v 2.000000 2.000000 0.000000
vt 0.000000 1.000000 0.000000
vt 0.000000 0.000000 0.000000
vt 1.000000 0.000000 0.000000
vt 1.000000 1.000000 0.000000
# Este es un comentario
usemtl wood
f 1/1 2/2 3/3 4/4
```

No es necesario que los números de vértices coincidan en la especificación de caras poligonales. Una superficie poligonal se obtiene por simple enumeración de las caras que la componen. Para añadir normales, se escribe la lista y se modifica la especificación de caras poligonales. Las caras pueden compartir vértices y no es necesario que utilicen todos los que están disponibles, o que las listas sean del mismo tamaño, siempre y cuando todas las referencias en una cara poligonal sean válidas.

En su forma básica este formato es sencillo de leer y de escribir y su conversión a VRML para ver el modelo en un navegador tiene pocas dificultades.

## **Apéndice B**

# **Uso, diseño y programación de la implementación**

En este apéndice se describen las características de la implementación, que fue realizada en Java 1.4.1 corriendo sobre Windows 98/XP.

### **B.1. Uso**

Para facilitar la invocación correcta de los diferentes programas realizados se escribieron varios archivos de procesamiento por lotes (*batch*) que fueron llamados por línea de comandos. Aunque los archivos batch de Windows no son tan flexibles como los scripts de UNIX, fueron bastante adecuados.

La mayoría de los programas corren en modo texto no interactivo. Los únicos programas gráficos fueron los relacionados con el procesamiento de modelos faciales para formar el conjunto facial de inicialización.

A continuación se describen estos programas en el orden en que fueron utilizados junto con los detalles referentes a su implementación.

### **B.2. Programas implementados**

#### **B.2.1. Procesamiento de modelos faciales 3D**

La técnica de rostros característicos con datos mixtos requiere que los modelos faciales sean isomorfos topológicamente, es decir, que tengan el mismo número de vértices, que estén en posiciones correspondientes y las conectividades entre ellos

deben coincidir. Esto es necesario para poder combinarlos linealmente. Por ello se deben procesar los modelos faciales disponibles para que cumplan con estas condiciones y así formar el conjunto facial de inicialización.

## **IApp**

Este programa fue uno de los primeros en realizarse, debido a la importancia de procesar los rostros y a que las características de este procesamiento impactan el rendimiento posterior de la técnica.

Originalmente se había decidido que los modelos faciales estuvieran formados por un grupo reducido de vértices que correspondiera a puntos característicos del rostro, como el contorno de la boca, la quijada, etc. Esto reduciría el número de vértices y con ello el tiempo de procesamiento, además de que facilitaría crear una animación facial utilizando los modelos. En una primera implementación, llamada HApp, se cargaban el modelo y su textura, y una malla de puntos característicos era colocada manualmente en los puntos adecuados del rostro. Sin embargo, procesar manualmente los rostros tomaba una gran cantidad de tiempo y cualquier cambio en los puntos característicos usados requería repetir el trabajo, además de que los modelos tenían tan pocos puntos que la apariencia del modelo era claramente artificial, incluso con texturas.

Por ello se decidió realizar un muestreo denso, requiriendo solamente seleccionar la región cubierta por el rostro. Esto aceleró el procesamiento y aumentó el realismo de los rostros. Este muestreo denso acarrió problemas de almacenamiento en memoria, por lo que se tomó la decisión de leer los datos de disco conforme se necesitaran, alterando los formatos de almacenamiento adecuadamente. El programa resultante fue llamado IApp, invocado a través del siguiente archivo batch.

```
REM %1 el nombre sin extension del obj. El jpg tiene un 0 extra.  
REM %2 para indicar si se escriben conectividades o no  
REM El directorio de salida es outLFD  
java -cp Tenshi -ea IApp %10.jpg %10.jpg %1.obj outLFD  
120 120 %2 > err.out
```

Una invocación común es:

```
cd C:\hcvn\Avatar\programas  
procesa.bat ..\base\sujetoUno BANDERA
```

que se expande a

```
java -cp Tenshi -ea IApp ..\base\sujetoUno0.jpg
    ..\base\sujetoUno0.jpg
    ..\base\sujetoUno.obj outLFD
    120 120 BANDERA > err.out
```

Las dos imágenes son la que se muestra al usuario y la que se usa para realizar los cálculos que requieren conocer el tamaño de la textura —en este caso es la misma. Esta opción existe para cuando la textura es muy grande y queremos trabajar sobre una vista más pequeña. Luego sigue el archivo OBJ que sirve de entrada, el nombre del directorio al que se escribirán los resultados (usando el nombre de la entrada), las dimensiones deseadas para el resultado. BANDERA es un parámetro opcional que indica que se deben escribir las conectividades de la malla al archivo `c.alt` como un `java.util.ArrayList` serializado.

Los resultados se escriben como un `LinearizedFaceData` serializado, con extensión `lfd` y la textura como un archivo PNG. Los archivos `lfd` son un pequeño encabezado con información descriptiva seguidos por una secuencia de flotantes que conserva el orden del `ArrayList`, pero al ser solamente flotantes pueden deserializarse individualmente y no en conjunto como `ArrayList`.

### B.2.2. Obtención de rostros característicos y generadores

Se utilizaron tres programas: `EigenFaceCreator2D`, `LinearizedFaceData` y `StreamingEigenFaceCreator3D`.

`EigenFaceCreator2D` lee imágenes de un directorio y obtiene los rostros característicos comunes, que son almacenados como una matriz serializada, en un archivo con extensión `ef2d`. Al ser imágenes en tonos de gris de tamaños reducidos ( $60 \times 60$ ,  $120 \times 120$ ) es posible cargar la totalidad en memoria. Además se almacenan las codificaciones de los rostros en un archivo con extensión `enc2d`, que será utilizado al obtener los modelos 3D generadores.

`LinearizedFaceData` toma un conjunto de modelos faciales 3D almacenados como archivos `lfd` y los reordena y escribe en un solo archivo de extensión `sfd`, de modo que sea fácil obtener los modelos generadores 3D.

Teniendo los rostros característicos y las codificaciones, `StreamingEigenFaceCreator3D` obtiene los rostros generadores 3D mediante el método descrito en capítulos anteriores. El nombre no es exacto, pero sirve para recordar la relación entre ambos conjuntos de rostros, que son almacenados en bloques de ordenadas

correspondientes, para facilitar la lectura de disco y su combinación sin requerir cargar la totalidad de los datos en memoria.

Al igual que para el preprocesamiento se creó un archivo de procesamiento por lotes. `condensa.bat` llama a `LinearizedFaceData`.

```
REM %1 es el directorio sin diagonal de donde leo png's y lfd's.
REM %2 es el sufijo (antes de la extension) que se le da a todo
REM %3 es el directorio de salida SIN diagonal
REM %4 Archivo de conectividades para condensar
REM Debo revisar cuales parametros ya existen y cuales son creados
java -cp Tenshi -ea EigenFaceCreator2D 120 120
    %1 %3\%2.ef2d %3\%2.enc2d
call condensate.bat %1 %3\%2.sfd %4
java -cp Tenshi -ea StreamingEigenFaceCreator3D
    %3\%2.enc2d %3\%2.sfd %3\%2.ef3d
```

Ahora se leen todos los archivos PNG de un directorio y se obtienen los rostros característicos y las codificaciones, que son escritas en el directorio de salida.

Después se escriben las mallas serializadas en un solo archivo con extensión `sfd`, de `StreamingFaceData`. Esto para no tener que cargar en memoria simultáneamente todas las mallas en los siguientes pasos.

En la tercera línea se obtienen las mallas características con el método propuesto en este trabajo, utilizando los rostros característicos y las codificaciones. Las mallas características se escriben también en formato `sfd` pero con extensión `ef3d` para reconocerlas.

### **B.2.3. Generación de un modelo**

Para generar un modelo se usa el programa `Demo`, para el que también se realizó un archivo de procesamiento por lotes. Las imágenes que recibe deben estar escaladas y recortadas como las del conjunto de inicialización. El recortado se hace con cualquier aplicación de procesamiento de imágenes, mientras que el escalado se realiza con el programa `TexHalver`.

#### **TexHalver**

Como el sistema fue inicializado con rostros escalados de modo que la base de los ojos ocupe el centro de la fotografía es necesario procesar las imágenes de entrada como se muestra en la Figura B.1. El programa pide al usuario que defina

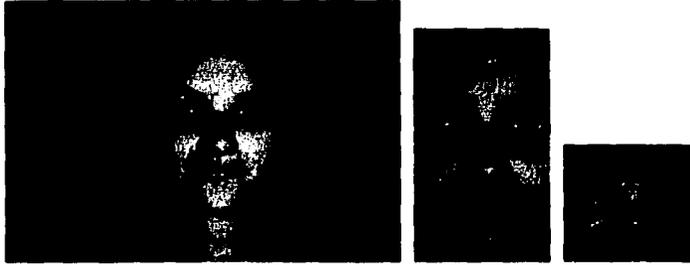


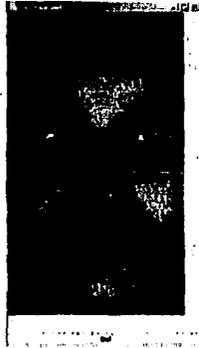
Figura B.1: La textura original es recortada y posteriormente escalada

la base de los ojos y entonces la escala para obtener una imagen lista para servir de entrada al sistema.

La invocación es la siguiente:

```
java -cp Tenshi -ea TexHalver cropped\foto.png  
croppedNscaled\foto.png
```

y se muestra en pantalla la aplicación:



La barra amarilla (en la ilustración se muestra a la mitad de la nariz) es colocada por el usuario en el lugar adecuado para dividir en dos la imagen. Cuando el usuario aprieta el botón en la parte inferior los pedazos son escalados y el resultado escrito a disco.

## Demo

REM %1 es la imagen de entrada  
REM %2 es la ruta hasta antes del punto de la extensión de las  
REM efases, que deben tener el mismo nombre y ruta.  
REM %3 es el sufijo de la malla y textura de salida  
java -cp Tenshi -ea Demo %1 %2.ef2d %2.ef3d %3

Así, para obtener el modelo correspondiente a rostro.png utilizando los rostros característicos dir/efases.ef2d y dir/efases.ef3d para escribirlos en salida/resultado.png y salida/resultado.wrl se escribe:

```
Demo.bat rostro.png dir/efases salida/resultado
```

## B.2.4. Realización de experimentos

Es indispensable que los experimentos y el reporte de sus resultados puedan hacerse sin supervisión, no sólo para facilitar el trabajo, también para evitar errores humanos.

### ExperimentoBatch

ExperimentoBatch lee todas las imágenes de un directorio y las toma como punto de comparación. En otro directorio están las imágenes ocluidas y perturbadas, que son agrupadas en base al esquema de nombrado. Entonces cada grupo de imágenes es codificado con los rostros característicos y se generan los rostros 3D. Se calculan todas las medidas de error y son escritas a disco en formato texto, del que posteriormente se obtendrán las gráficas usando Excel.

Una línea típica para realizar un bloque de experimentos es:

```
java -ea -cp Tenshi ExperimentoBatch  
in\60\perfectas6h6m  
ef\h31_60x60.ef2d ef\h31_60x60.ef3d  
in\60\errores in\60\oclu\6h6m  
out\60\oclu\6h6mConHom > 60oclulog6h6mConHom.txt
```

El primer parámetro es un directorio que contiene únicamente las imágenes y modelos originales, que sirven de punto de comparación con las demás y que por esto son llamadas perfectas. 6h6m quiere decir que las imágenes son de 6 hombres

y de 6 mujeres, seleccionados mediante su grado de representatividad en la base. En este caso las mujeres sirven para hacer pruebas con el sexo opuesto. Para las pruebas con una base mixta se utiliza 6hym que quiere decir 6 hombres y mujeres.

El segundo y tercer parámetro son los rostros característicos y generadores, en este caso los obtenidos a partir del conjunto facial de 31 hombres con imágenes de  $60 \times 60$  pixels.

Después se escriben dos directorios que contienen las imágenes modificadas. El directorio de errores contendría imágenes espurias, pero después de realizar los primeros experimentos se determinó que las pruebas con imágenes espurias deberían realizarse por separado. El segundo directorio contiene las oclusiones de las imágenes perfectas.

Finalmente el directorio donde son escritos los resultados. Sea la imagen perfecta nombre60.png junto con todas sus alteraciones. Se escribe a disco el archivo llamado nombre60.exp2, conteniendo los errores medidos para todo el bloque en formato texto. Para cada imagen se escriben dos archivos: el png de su reconstrucción como imagen y el lfd del modelo generado. Aunque estos últimos no se requieren para reportar los experimentos, si ya fueron calculados es preferible guardarlos dado el costo en tiempo que esto implica.

Cada bloque de experimentos se ejecuta en un promedio de dos horas en un procesador Pentium 4 a 2.0GHz.

En total hay 6 bloques de experimentos, ya que hay dos dimensiones de imagen ( $60 \times 60$  y  $120 \times 120$ ) con tres conjuntos de rostros generadores: el formado a partir de los 31 hombres, de las 23 mujeres y del total de 54 personas.

### **B.3. Tipos de Datos Abstractos**

El diseño de los tipos abstractos es particularmente importante ya que como son escritos a disco, cualquier cambio en ellos evita que puedan leerse nuevamente. Los siguientes son los más importantes:

**FaceData y LinearizedFaceData.** Se trata de contenedores de los vértices y las conectividades que definen completamente a una malla con funciones para normalización y comparación, entre otras.

**Todos los tipos de coordenadas.** ImageCoord, TextureCoord, NormalVec, y VertexCoord. Incluyen funciones para operaciones comunes, como normalizar y conversión de coordenadas de textura a coordenadas de imagen y viceversa, dado el tamaño de textura.

**StreamingFaceData.** Clase que se encarga de escribir y leer de disco las mallas generadoras de modo que no sea necesario cargarlas todas en memoria.

**Surel.** De *SURface ELeMent*. Clase que almacena un vértice, un vector normal y unas coordenadas de textura, además de proveer métodos de interpolación. Esta clase se usa al momento de procesar la malla original para usarla en el método.

**FaceScale.** Realiza escalado de los datos faciales, principalmente con fines de normalización geométrica.

**FaceDataComparer.** Realiza comparaciones entre rostros.

Además de clases para leer o escribir en los formatos OBJ y VRML. La lectura y escritura de imágenes se hace con las clases disponibles en Java 1.4, `java.imageio.*`.

## **B.4. Plataforma de Programación**

Hubo muchas razones para utilizar Java:

1. Es un lenguaje interpretado sobre una versión compilada. Esto facilita los cambios en los programas, ya que es posible recompilar un archivo y correr de inmediato sin ligar.
2. Es eficiente. A pesar de ser interpretado, la compilación del código fuente en bytecodes le da un buen rendimiento para muchas tareas.
3. Usa recolección de memoria. La liberación de memoria en lenguajes como C/C++ provoca errores y consume mucho tiempo de programación que no sirve para el propósito real del programa. La capacidad de los equipos actuales permite el uso de recolectores de memoria con bajo impacto en el rendimiento.
4. Incluye estructuras de datos y mecanismos para manipularlos. Iteradores, listas, conjuntos, etcétera, cuyo diseño permite cambiar la implementación usada con pocos o ningún cambio en el fuente.
5. Utiliza principalmente objetos, exceptuando los tipos primitivos. La encapsulación de código facilita la integración de todos los procesos diferentes que requiere la solución del problema.

6. Soporte nativo para serialización. Java permite serializar (escribir a disco) la mayor parte de los objetos, lo cual libera al programador de la necesidad de diseñar formatos de almacenamiento de datos en casos simples. Así una estructura de datos puede almacenarse y restaurarse fácilmente.
7. Capacidades gráficas. Un sistema moderno y completo para crear interfaces gráficas, con manejo de ventanas y eventos.
8. Capacidad de lectura y escritura de formatos gráficos a partir de Java 1.4. Esto representa un gran ahorro de tiempo al programador y fomenta el uso de formatos eficientes como JPG y PNG.

Todas estas capacidades lo hacen ideal para la programación rápida, y su portabilidad no limita el uso futuro del programa. La versión utilizada fue la 1.4.1, distribuida por Sun.

Siendo una máquina virtual, Java se ejecuta sobre una variedad de sistemas operativos. En este trabajo el desarrollo comenzó en Windows 98 y terminó en Windows XP Home.

#### **B.4.1. Rendimiento y uso de disco**

Los archivos de modelos en formato lfd ocupan alrededor de 354Kb, para texturas de  $120 \times 120$  píxeles. Se trata simplemente de un `ArrayList` serializado, una estructura de Java similar a un vector. Cuando el mismo modelo se escribe en formato VRML, ocupa alrededor de 2 megas.

La generación de un modelo 3D, desde la codificación hasta la escritura a disco del VRML y la textura ocupa unos segundos. Estos tiempos son para una PC Pentium 4 a 2Ghz, con 256 megas de RAM.

...the ... of ...  
...the ... of ...  
...the ... of ...

...the ... of ...  
...the ... of ...

...the ... of ...  
...the ... of ...

...the ... of ...  
...the ... of ...  
...the ... of ...  
...the ... of ...

...the ... of ...

...the ... of ...  
...the ... of ...  
...the ... of ...  
...the ... of ...  
...the ... of ...

## Apéndice C

### Gráficas de resultados

Aunque se realizaron pruebas con varios rostros (entre 6 y 12 según el experimento) por razones de espacio solamente se muestran los resultados para el hombre y la mujer más representativos de los conjuntos de hombres y mujeres bajo el conjunto facial de inicialización de sólo hombres. Para ambos se muestran los resultados de las pruebas de oclusión y perturbación, en resoluciones de  $60 \times 60$  y  $120 \times 120$ .

#### C.1. Descripción de las gráficas

A la imagen original en el conjunto de inicialización se le llama imagen perfecta.

Se muestran 6 gráficas para cada experimento. En las primeras 5 se muestra mediante una caja centrada en el promedio la desviación estándar, colocada sobre una línea que va del máximo al mínimo. En la última se muestran los valores obtenidos mediante puntos conectados por una línea.

En todos los experimentos se incluye la imagen perfecta (denotada *ori* por ser la original) para fines de verificación.

1. Error de la imagen reconstruida contra la imagen perfecta. La imagen reconstruida con los rostros característicos 2D se compara pixel a pixel contra la imagen perfecta. Por ello el eje de los errores va de 0 a 255.
2. Error de la imagen alterada contra la imagen perfecta. Similar a la anterior, pero la comparación se realiza con la imagen alterada (por filtros y oclusiones) que sirvió de entrada para la generación, con el objeto de analizar la propagación del error de la entrada en la salida.

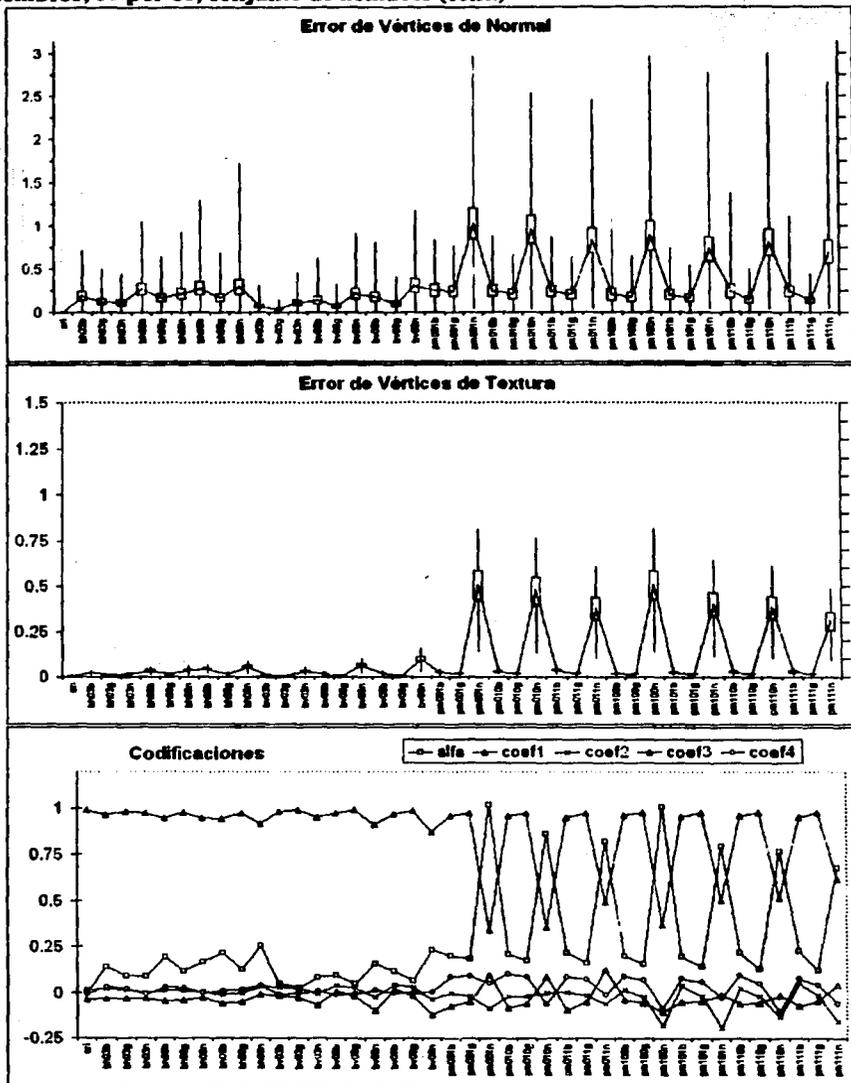
3. Error del modelo generado contra el perfecto para vértices de geometría. Después de normalizar a un cubo de  $1 \times 1 \times 1$  se calcula la distancia euclidiana entre ellos. El máximo de la escala es  $\sqrt{3}$ .
4. Error del modelo generado contra el perfecto para vértices de normal. Similar a la anterior, midiendo el ángulo entre las normales correspondientes, con un máximo de  $\pi$ .
5. Error del modelo generado contra el perfecto para vértices de textura. Midiendo las distancias euclidianas entre las coordenadas de textura  $(s, t)$  con un error máximo de  $\sqrt{2}$ .
6. Ángulo entre codificaciones ( $\alpha$ ) y los primeros cuatro coeficientes de cada codificación. El ángulo entre codificaciones indica la diferencia entre ambos rostros, al igual que el primer coeficiente, que para un rostro bien representado en el conjunto de inicialización debe valer cerca de 1. El resto de los coeficientes se incluye para mostrar su rango de valores y determinar si guardan alguna relación con las modificaciones aplicadas.

### C.1.1. Oclusiones

Para facilitar la consulta y comparación las gráficas se muestran en bloques de tres por página.

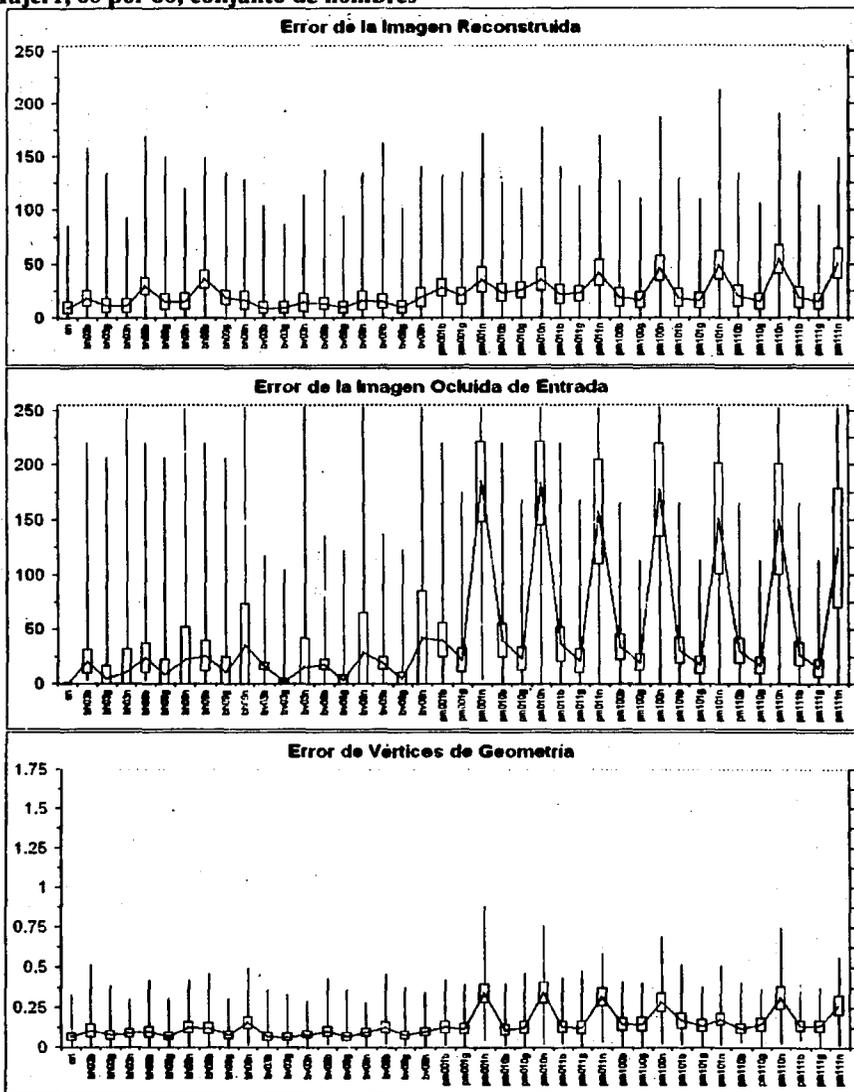


**Hombre1, 60 por 60, conjunto de hombres (cont.)**

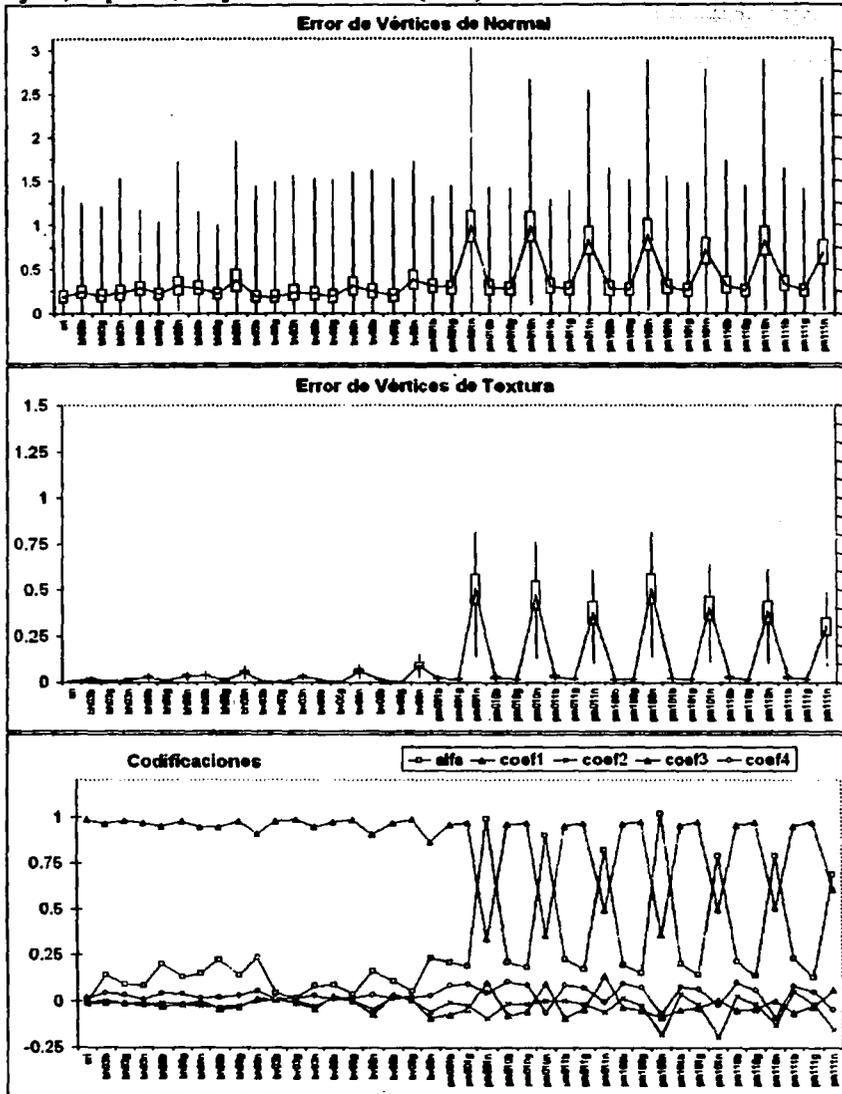


TESIS CON  
FALLA DE ORIGEN

**Mujer1, 60 por 60, conjunto de hombres**



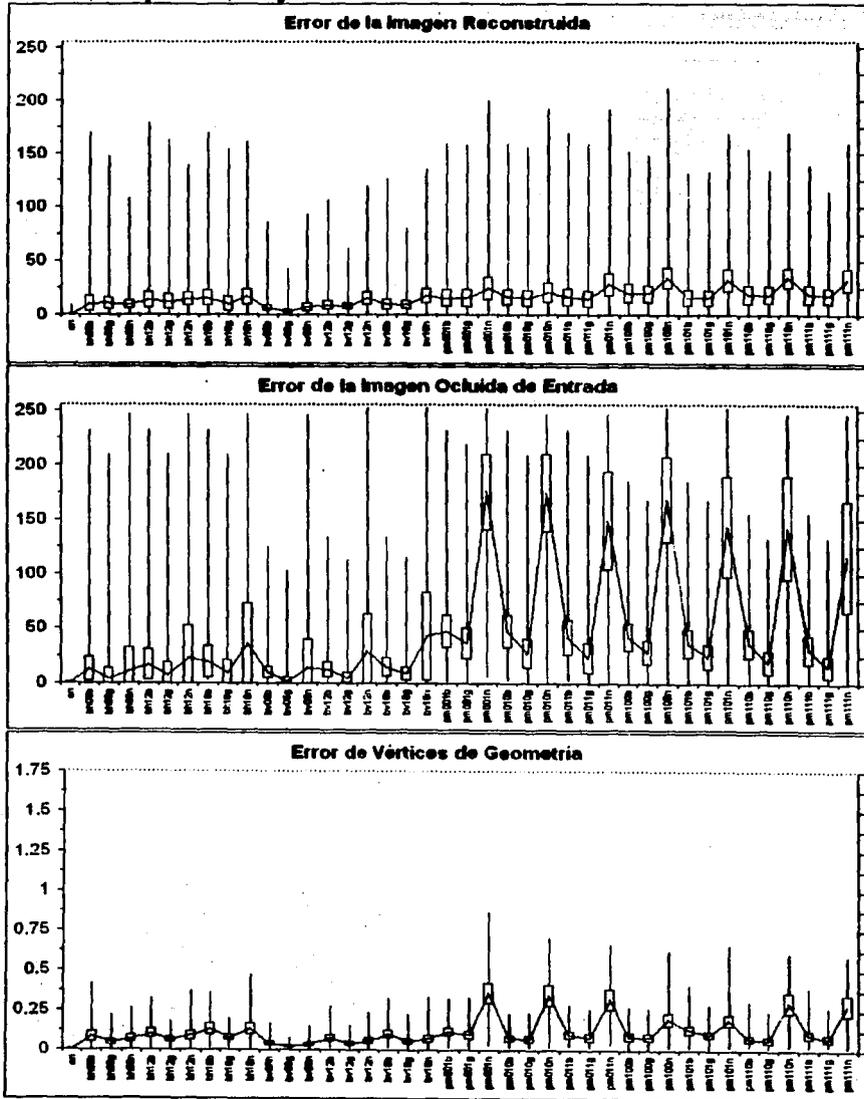
Mujer1, 60 por 60, conjunto de hombres (cont.)



10073827  
 10/10/2010

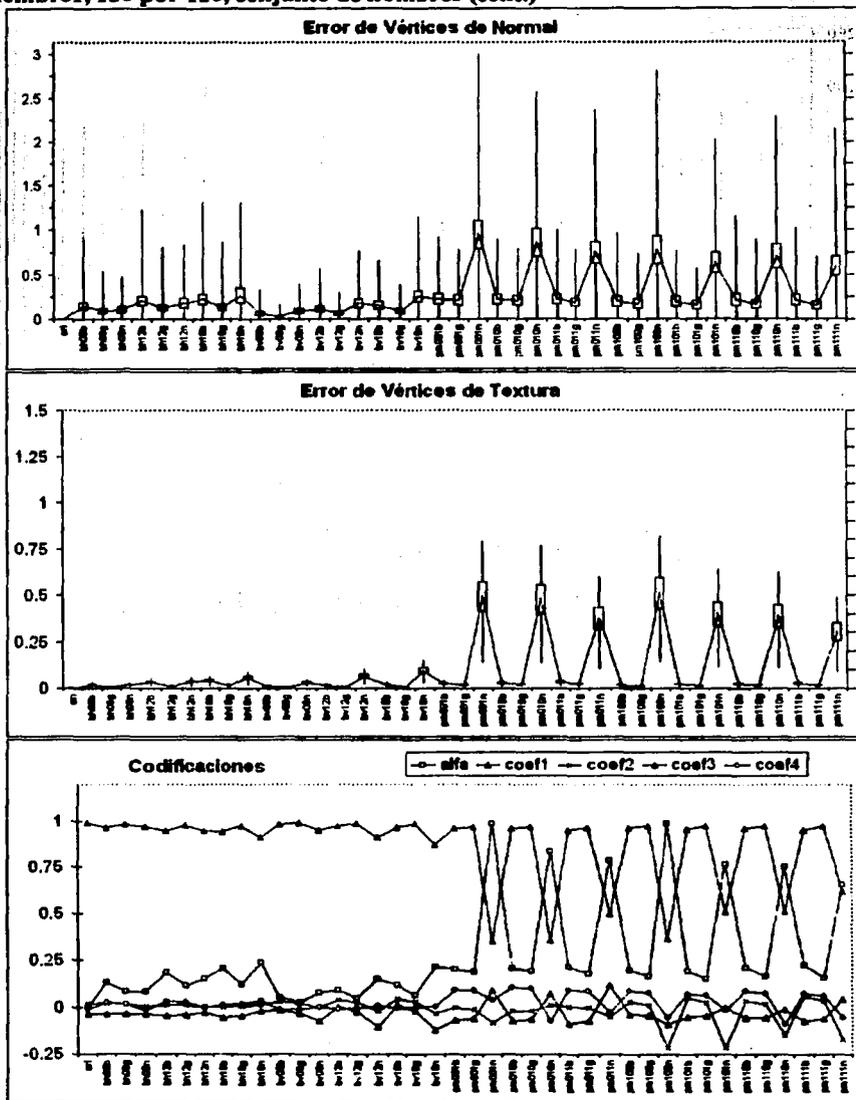
TESIS CON  
 FALLA DE ORIGEN

### Hombre1, 120 por 120, conjunto de hombres



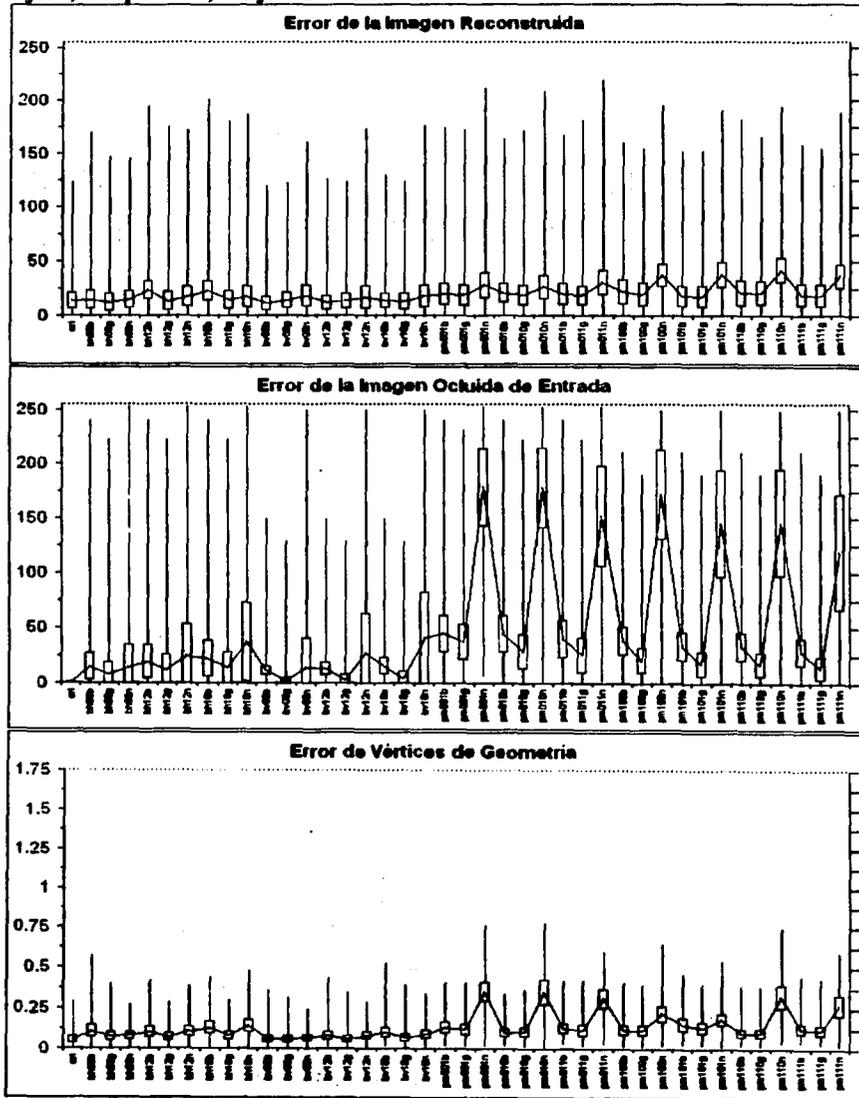
TESIS CON  
FALLA DE ORIGEN

Hombre1, 120 por 120, conjunto de hombres (cont.)



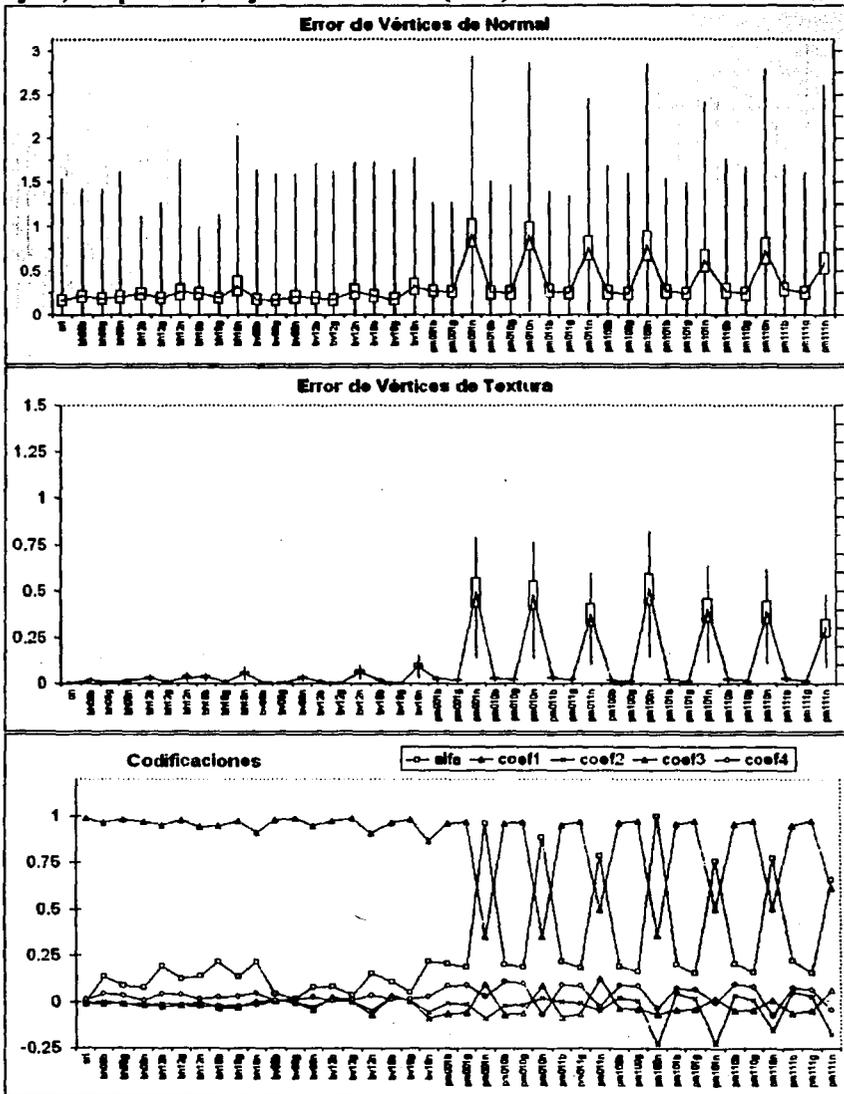
TESIS CON FALLA DE ORIGEN

**Mujer1, 120 por 120, conjunto de hombres**



**TESIS CON  
FALLA DE ORIGEN**

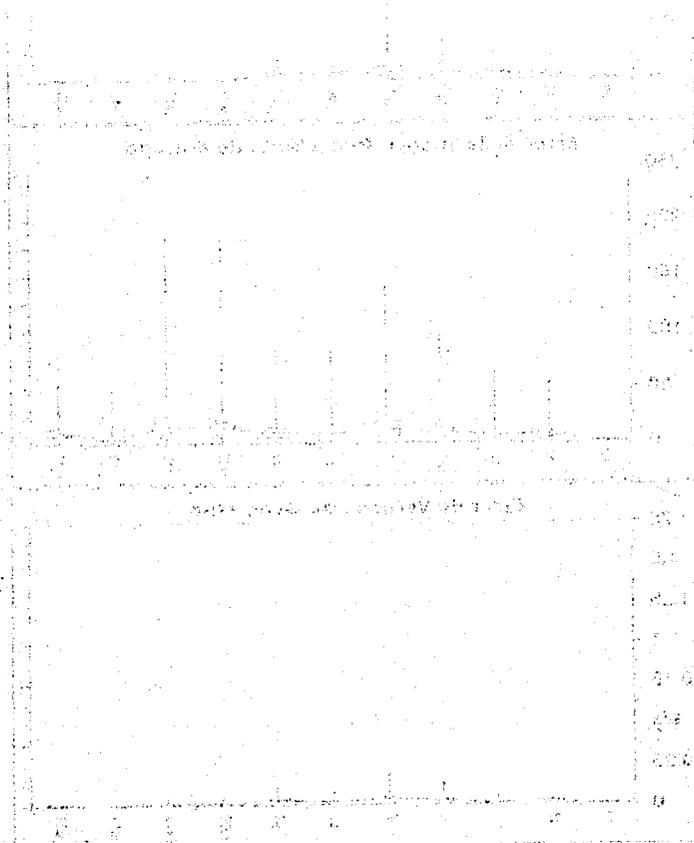
Mujer1, 120 por 120, conjunto de hombres (cont.)



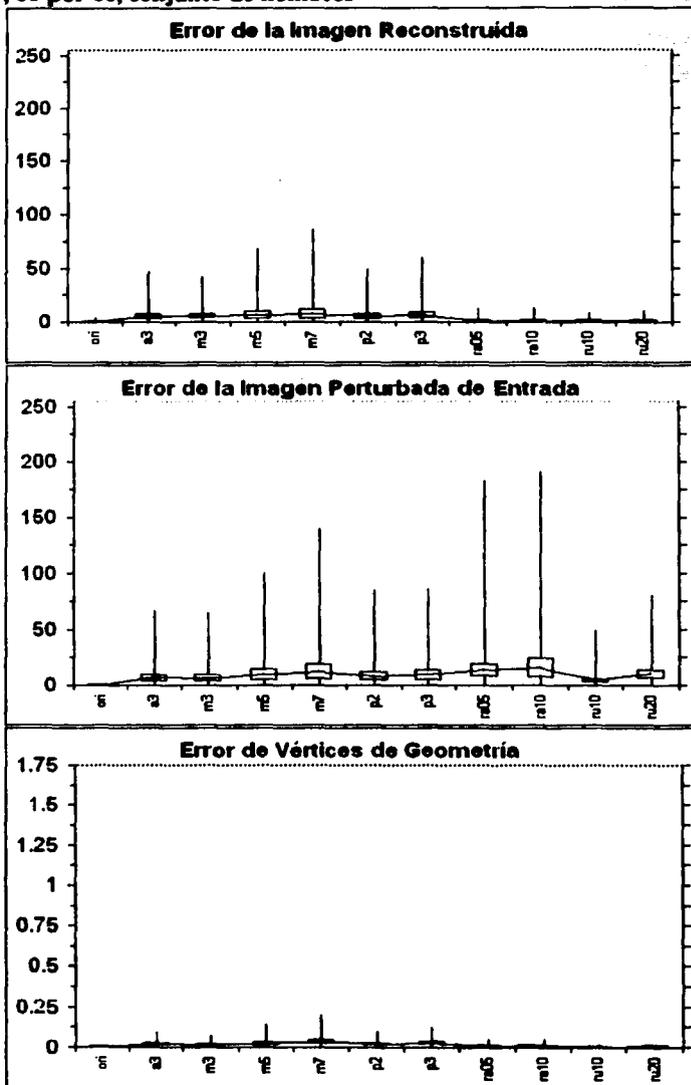
TESIS CON  
FALLA DE ORIGEN

### C.1.2. Perturbaciones

Para facilitar la consulta y comparación las gráficas se muestran en bloques de tres por página.

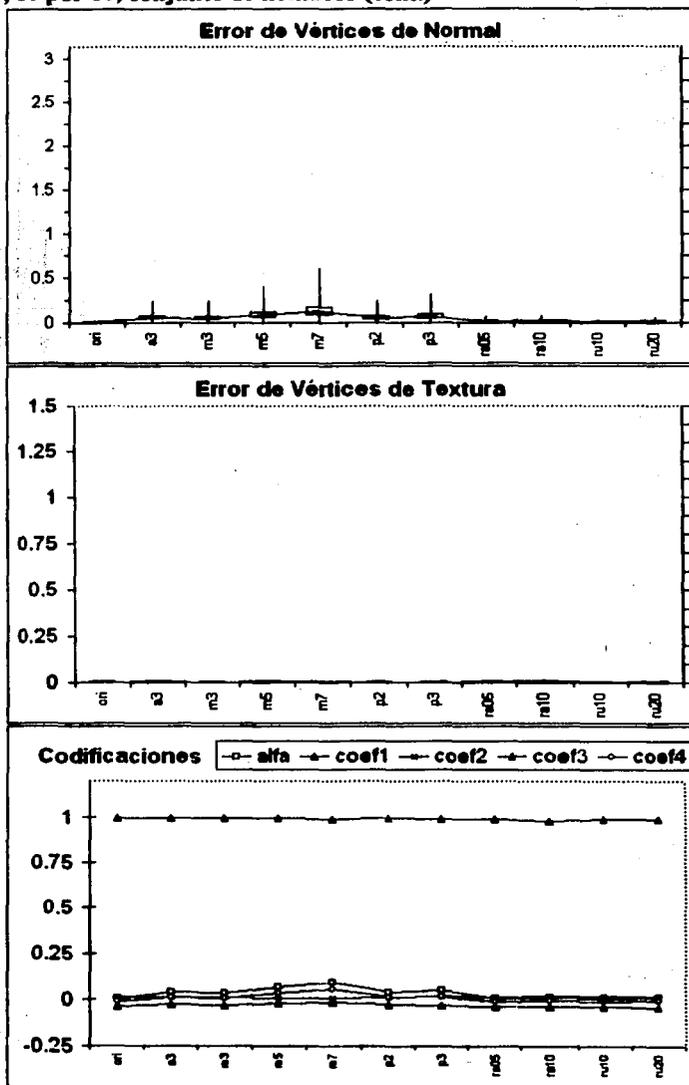


**Hombre1, 60 por 60, conjunto de hombres**



**TESIS CON  
FALLA DE ORIGEN**

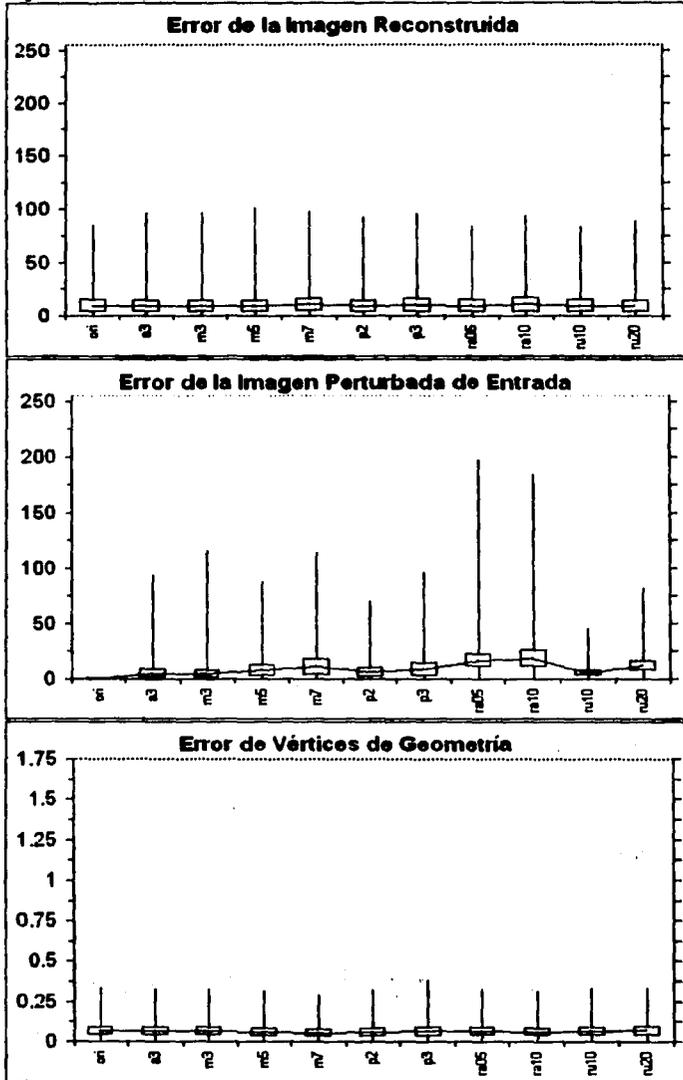
Hombre1, 60 por 60, conjunto de hombres (cont.)



TESIS CON FALLA DE ORIGEN

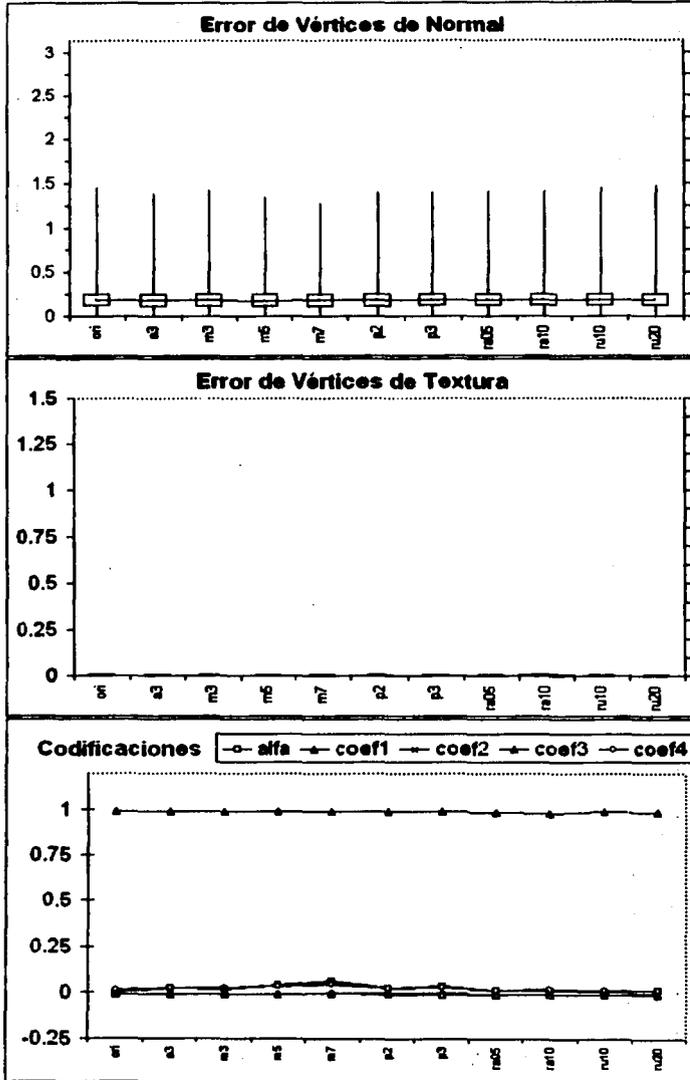
ESTA TESIS NO SALE DE LA BIBLIOTECA

**Mujer1, 60 por 60, conjunto de hombres**



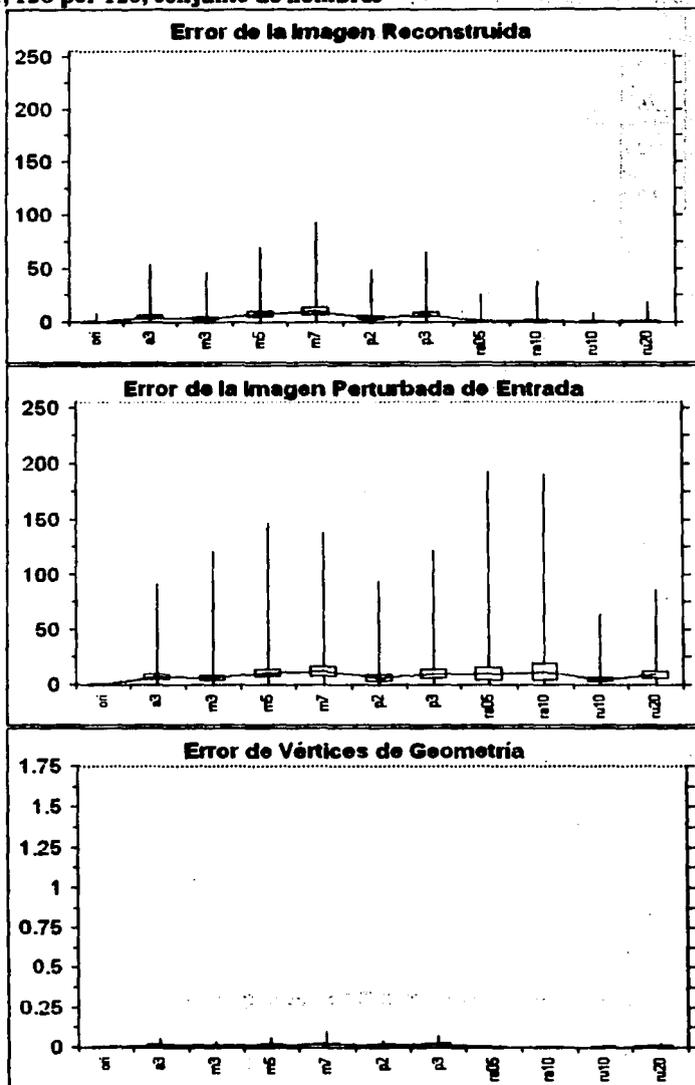
**TESIS CON FALLA DE ORIGEN**

Mujer1, 60 por 60, conjunto de hombres (cont.)

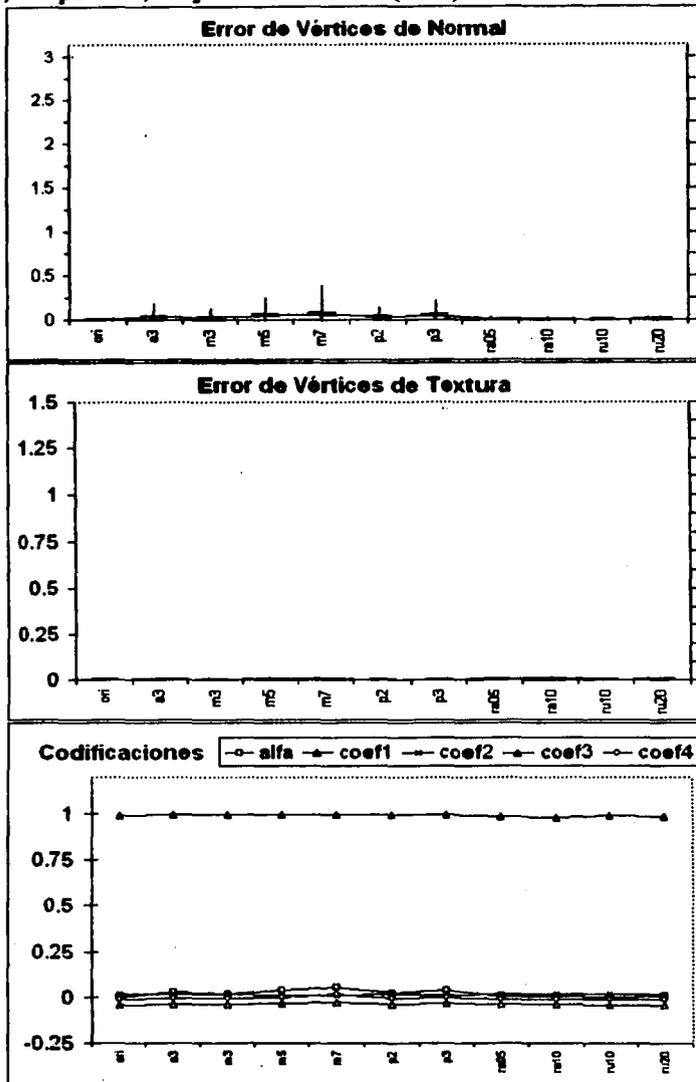


TESIS CON  
FALLA DE ORIGEN

**Hombrel, 120 por 120, conjunto de hombres**

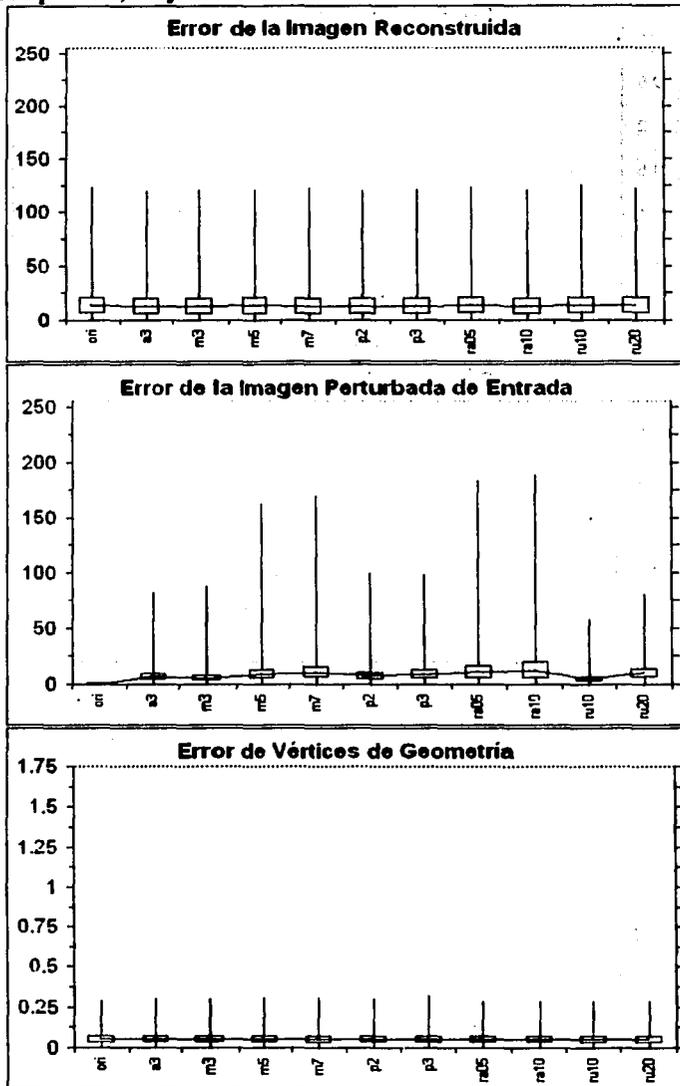


**Hombre1, 120 por 120, conjunto de hombres (cont.)**



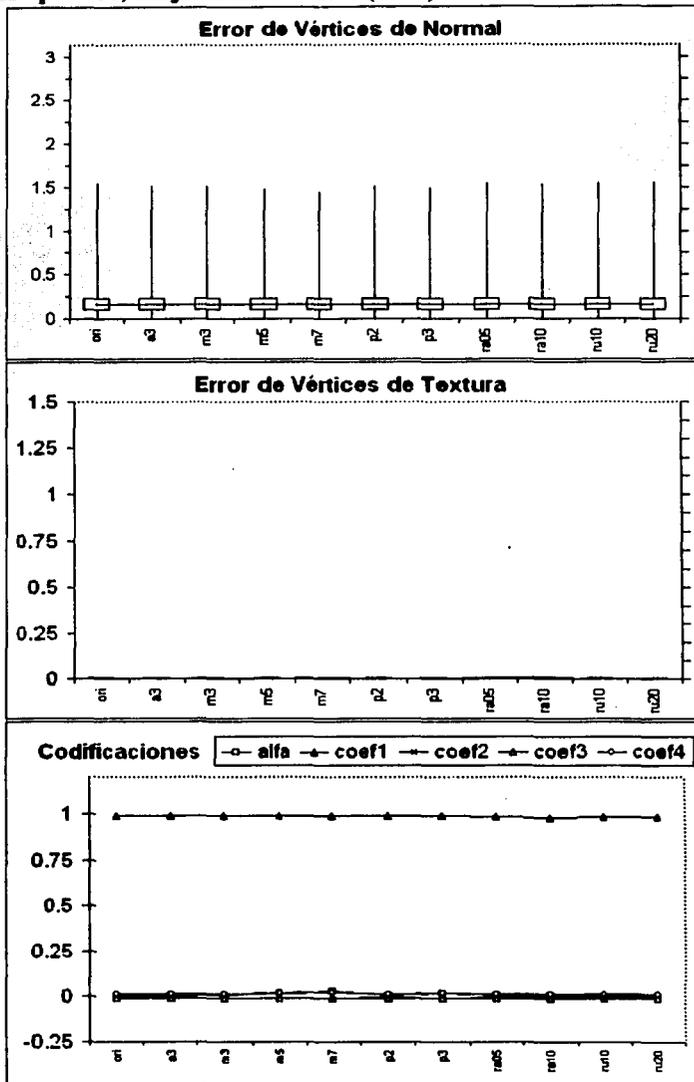
**TESIS CON  
FALLA DE ORIGEN**

**Mujer1, 120 por 120, conjunto de hombres**



**TESIS CON  
FALLA DE ORIGEN**

Mujer1, 120 por 120, conjunto de hombres (cont.)

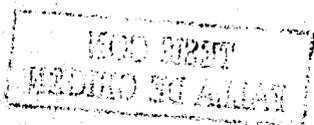




## Apéndice D

### Modelos Generados

Se muestran imágenes de los modelos generados para Hombre1 en el conjunto facial de inicialización de hombres, en ambas resoluciones:  $60 \times 60$  y  $120 \times 120$ . En cada página se muestra el rostro de frente sin textura y de perfil con textura. Se muestran el original, el obtenido a partir de una imagen ocluida con una barra horizontal en 9 y 18 pixels respectivamente en color gris y finalmente con ruido aleatorio de 10 por ciento.



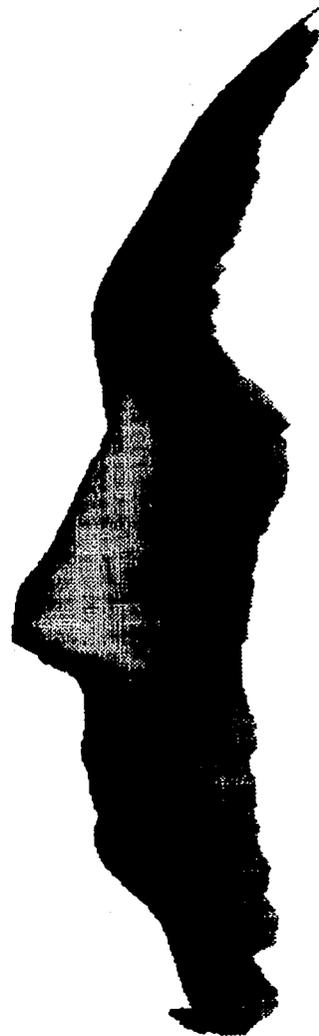
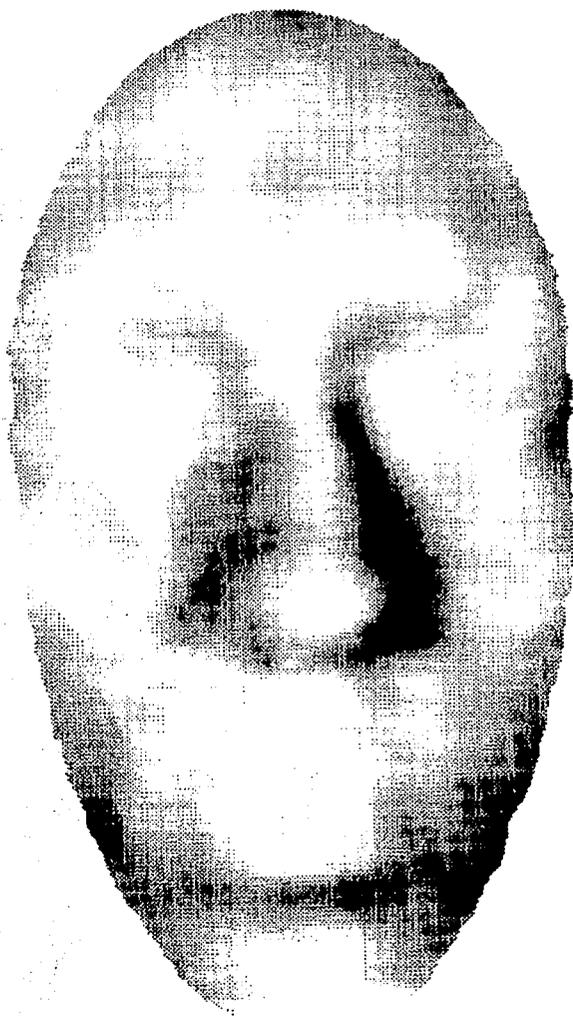
**D.1.** 120 × 120

**D.1.1. Original**



TESIS CON  
FALLA DE ORIGEN

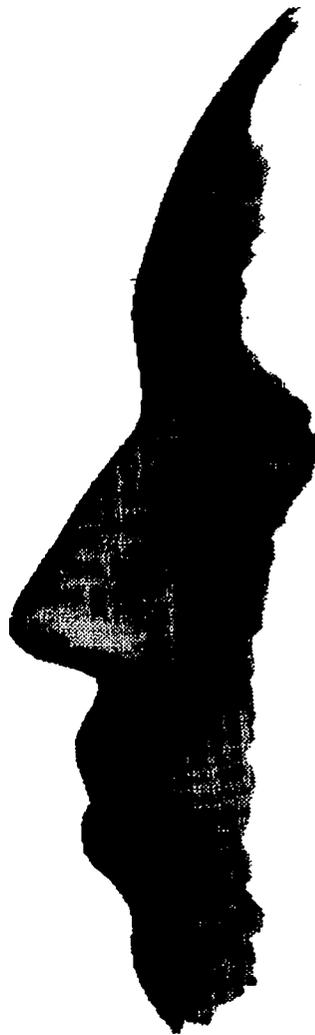
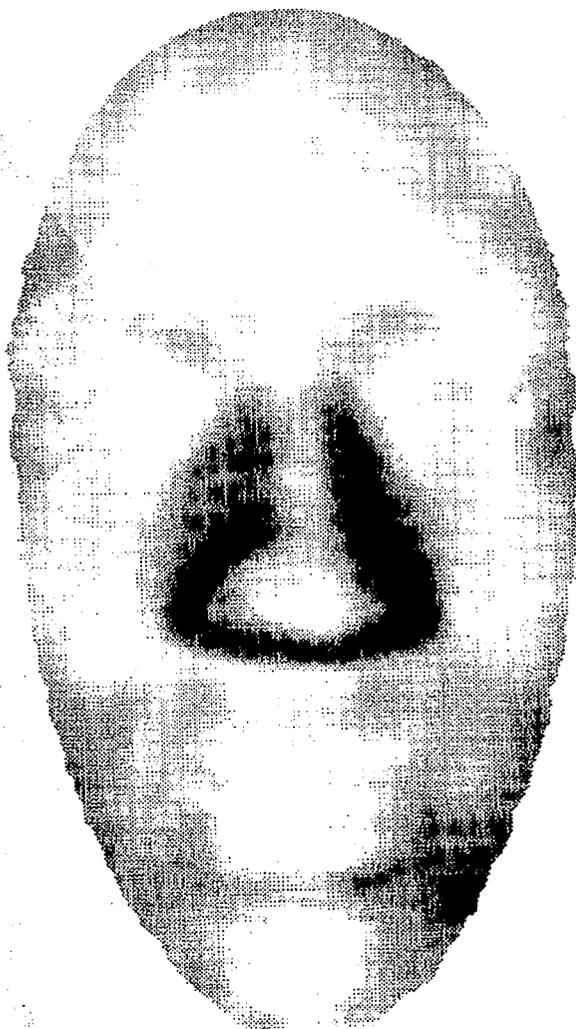
**D.1.2. Barra horizontal de 18 pixels en gris**



YCO ACUT  
WEDNESDAY

TESIS CON  
FALLA DE ORIGEN

**D.1.3. Ruido Aleatorio 10 por ciento**

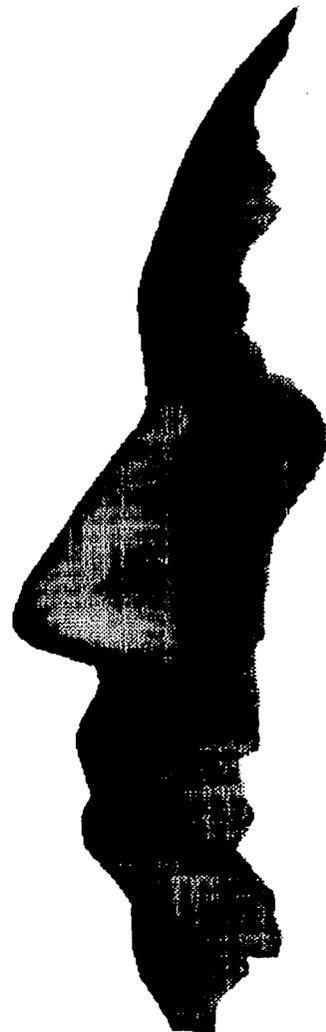
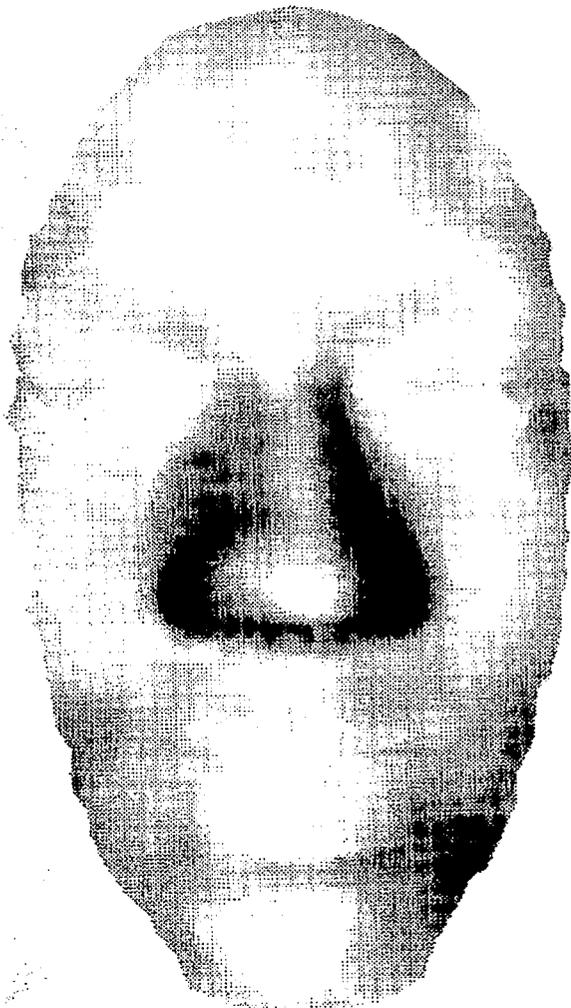


UNIVERSIDAD  
NACIONAL DEL PARAGUAY

TESIS CON  
FALLA DE ORIGEN

**D.2.** 60 × 60

**D.2.1. Original**

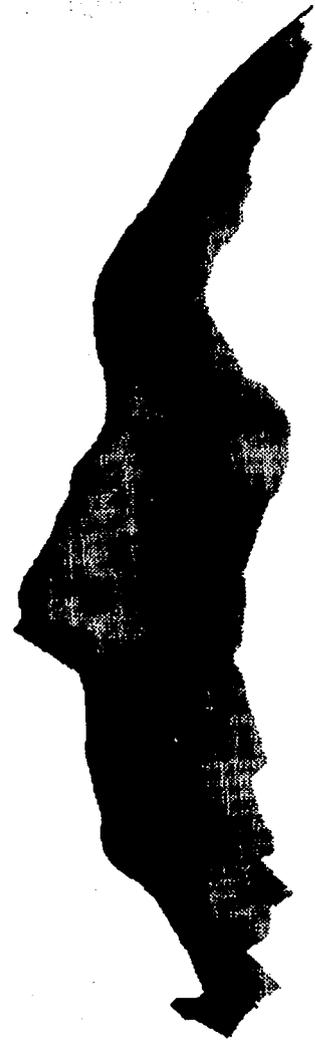


91

SECRET  
MEDIO DE AERIAL

TESIS CON  
FALLA DE ORIGEN

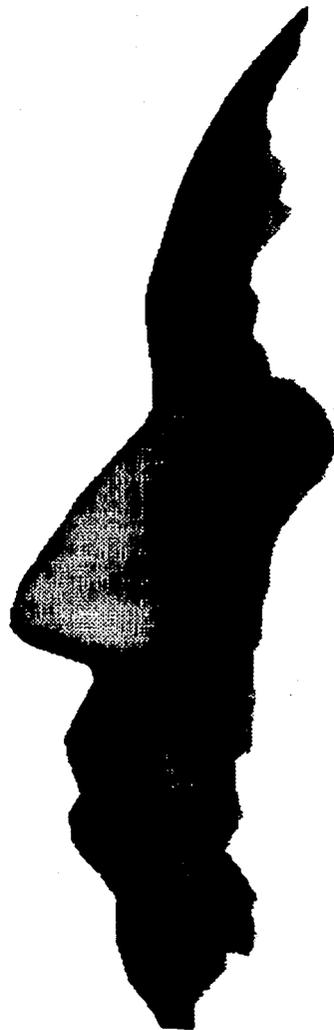
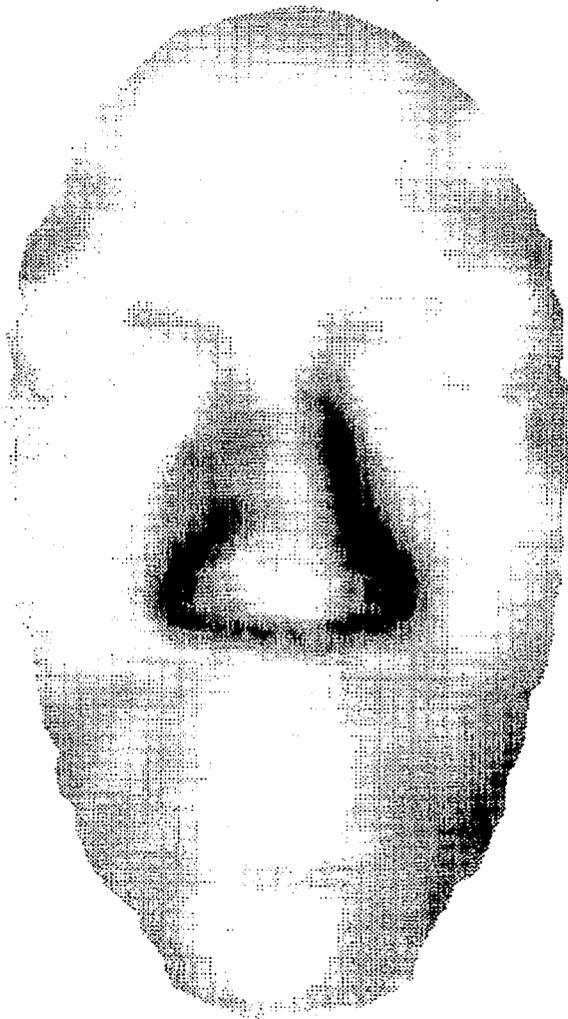
**D.2.2. Barra horizontal de 9 pixels en gris**



WFO BENT  
WFO BENT

TESIS CON  
FALLA DE ORIGEN

**D.2.3. Ruido Aleatorio 10 por ciento**



TESIS CON  
FALLA DE ORIGEN

Handwritten text at the top of the page, possibly a title or header, which is mostly illegible due to fading and noise.

NOV 28 1964  
WASHINGTON FIELD

## Literatura Citada

- Alexa, M. y Müller, W.: 2000, Representing Animations by Principal Components, *EUROGRAPHICS 2000*.
- Blanz, V. y Vetter, T.: 1999, A Morphable Model for the Synthesis of 3D Faces, *SIGGRAPH '99 Proceedings*, pp. 187–194.
- Cuevas, H. y Rudomin, I.: 2000, Generating a 3D Facial Model from a Single Image Using Principal Component Analysis, *Proceedings of the Third International Conference on Visual Computing*, pp. 137–140.
- Lee, W., Gu, J. y Magnenat-Thalmann, N.: 2000, Generating Animatable 3D Virtual Humans from Photographs, *EUROGRAPHICS 2000*.
- Lengagne, R., Tarel, J. P. y Monga, O.: 1996, From 2D Images to 3D Face Geometry, *Proceedings of IEEE Second International Conference on Automatic Face and Gesture Recognition*.
- Lyons, M., Plante, A., Jehan, S., Inone, S. y Akamatsu, S.: 1998, Avatar Creation using Automatic Face Processing, *ACM Multimedia '98*, pp. 427–434.
- Moghaddam, B. y Pentland, A.: 1995, Probabilistic Visual Learning for Object Detection, *Technical Report 326*, Media Laboratory, M.I.T.
- Moghaddam, B., Wahid, W. y Pentland, A.: 1998, Beyond Eigenfaces: Probabilistic Matching for Face Recognition, *Technical Report 443*, Media Laboratory, M.I.T.
- Pentland, A., Moghaddam, B. y Starner, T.: 1994, View-Based and Modular Eigenspaces for Face Recognition, *Technical Report 245*, Media Laboratory, M.I.T.

**Turk, M. y Pentland, A.: 1991, Eigenfaces for recognition, *Journal of Cognitive Neuroscience* 3(1), 71-86.**

Journal of Cognitive Neuroscience

Abstract: This paper presents a method for representing and recognizing human faces from a set of images. The method is based on the eigenfaces technique, which uses principal component analysis (PCA) to extract the most important features of the faces. The resulting eigenfaces are used to represent the faces in a low-dimensional space, which makes them easy to store and compare. The method is shown to be effective in recognizing faces from a set of images, even when the faces are rotated or scaled. The method is also shown to be robust to changes in lighting and background. The method is implemented on a personal computer, and the results are shown to be comparable to those of a human observer. The method is a simple and effective way to represent and recognize human faces.