

USS | UNIVERSIDAD SEÑOR DE SIPÁN

**FACULTAD DE INGENIERÍA, ARQUITECTURA Y
URBANISMO**

**ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA
DE SISTEMAS**

TESIS

**RECONOCIMIENTO FACIAL MEDIANTE IMÁGENES
ESTEREOSCÓPICAS**

**PARA OPTAR EL TÍTULO PROFESIONAL DE
INGENIERO DE SISTEMAS**

Autor:

Bach. Vejarano Campos Martín Desiderio

Asesor:

Mg. Zuloaga Cachay José Fortunato

Línea de Investigación:

Ciencias de la Computación

Visión por Computador

Pimentel – Perú

2018

Aprobación de la Tesis

Mg. Villegas Cubas Juan Elías
Presidente del jurado de tesis

Mg. Mejía Cabrera Heber Ivan
Secretario del jurado de tesis

Mg. Tuesta Monteza Víctor Alexci
Vocal del jurado de tesis



DEDICATORIA

Esta tesis se la dedico a nuestro Dios Padre quién supo guiarme por el buen camino dándome fuerzas necesarias para seguir adelante.

A mis padres, Angelita y Segundo, por su apoyo incondicional, consejos, valores, principios y perseverancia en los momentos difíciles para lograr los objetivos.

A mis hermanos Giovanna y Johnny por estar siempre apoyándome en seguir adelante. A mi hijo Johita por ser el sendero de esperanza a cumplir con mis metas.

Martín Vejarano Campos



AGRADECIMIENTO

En el presente proyecto de Tesis, en primer lugar me gustaría agradecer a Mi Dios Padre por bendecirme para llegar hasta donde he llegado y concretar este sueño anhelado.

A la Universidad Señor de Sipán por darme la oportunidad de estudiar y ser un buen Profesional.

A todos los docentes por su visión crítica, esfuerzo, dedicación y gracias a sus conocimientos impartidos en mi formación profesional para lograr culminar la presente Tesis.

A mi asesor metodológico Mg. José Zuloaga Cachay por su esfuerzo necesario quién con sus conocimientos, sus experiencias, motivación, se ha logrado que culmine con éxito la presente Tesis. Asimismo un agradecimiento especial a todos los miembros del jurado. Una mención honrosa a mi querido profesor Dr. Humberto Llampén Coronel por todo su apoyo incondicional y agradeciendo a todas las personas que ayudaron directa e indirectamente en el desarrollo de ésta Tesis.

Martín Vejarano Campos.



ÍNDICE

Contenido

DEDICATORIA	3
AGRADECIMIENTO.....	4
ÍNDICE.....	5
RESUMEN	9
ABSTRACT.....	10
INTRODUCCIÓN	11
CAPÍTULO I: PROBLEMA DE INVESTIGACIÓN.....	12
1.1 Situación Problemática.....	12
1.2. Formulación del Problema.....	13
1.3. Delimitación de la Investigación	13
1.4. Justificación e Importancia de la Investigación	14
1.4.1 Desde el punto de vista científica.....	14
1.4.2 Desde el punto de vista Social.....	14
1.5. Limitaciones de la Investigación.....	15
1.6. Objetivos de la Investigación	15
1.6.1 Objetivo general	15
1.6.2 Objetivos específicos	15
CAPÍTULO II: MARCO TEÓRICO	16
2.1. Antecedentes de Estudios.....	16
Nivel Internacional	16
Nivel Nacional.....	18
2.2. Estado del Arte.....	19
2.3. Bases teóricas científicas.	20
2.3.1. Visión Estereoscópica	20
2.3.2. Reconocimiento de imágenes 3D	31



2.3.3. Algoritmos para comparación de modelos faciales.....	39
2.4. Definición de la terminología.....	43
CAPÍTULO III: MARCO METODOLÓGICO.....	44
3.1. Tipo y Diseño de Investigación	44
3.2. Población y Muestra.....	44
3.3. Hipótesis.....	45
3.4. Operacionalización.....	45
3.5. Métodos, técnicas e instrumentos de recolección de datos	46
3.6. Procedimiento para la recolección de datos	47
3.7. Análisis Estadístico e Interpretación de los datos.....	48
3.8. Criterios éticos	49
3.9. Criterios de rigor científico	49
CAPÍTULO IV: ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS	50
4.1. Resultados en tablas y gráficos.	50
4.2. Discusión de resultados.	52
CAPÍTULO V: DESARROLLO DE LA PROPUESTA DE INVESTIGACIÓN.....	54
5.1. DISEÑO DEL SISTEMA.....	54
5.2. DESARROLLO DEL SOFTWARE.....	67
5.3. EVALUACIÓN DEL SISTEMA	73
CAPÍTULO VI: CONCLUSIONES Y RECOMENDACIONES.....	77
6.1. Conclusiones	77
6.2. Recomendaciones	77
REFERENCIAS BIBLIOGRÁFICAS:.....	78
ANEXOS.....	80



ÍNDICE DE ILUSTRACIÓN

<i>Ilustración 1 - (a) Sistema de visión estereoscópica biológico; (b) Superposición de las imágenes de ambos ojos, donde se observa la separación relativa de cada objeto de la escena (a menor distancia, más lejos se encuentra el objeto; a menor distancia, más cerca es.</i>	21
<i>Ilustración 2 Vista superior de un sistema de visión estereoscópica</i>	22
<i>Ilustración 3 - Pintura de cueva de Lascaux en Francia.</i>	22
<i>Ilustración 4 - Estereoscopio de Charles Wheatstone</i>	24
<i>Ilustración 5 - Ejemplo de visión Paralela.....</i>	24
<i>Ilustración 6 - Ejemplo de visión cruzada</i>	25
<i>Ilustración 7 - Ejemplo de Anáglifos</i>	26
<i>Ilustración 8 - Imagen de entrada de un tablero de ajedrez tomado en diferentes ángulos.....</i>	27
<i>Ilustración 9 - Resultado de la imagen después de aplicar el método " findChessBoardCornerGuesses ".....</i>	27
<i>Ilustración 10 - Resultado final de la calibración de imagen.....</i>	28
<i>Ilustración 11 - Proceso completo de visión estereoscópica</i>	31
<i>Ilustración 12 - La reducción en el promedio de error para los algoritmos de reconocimiento de rostros en base a las evaluaciones de FERET, FRVT 2002 y FRVT 2006.....</i>	32
<i>Ilustración 13 - Distancia geodésica azul vs distancia euclidiana (rojo)</i>	39
<i>Ilustración 14 - Diagrama a bloques del sistema de reconocimiento de rostros. Etapa de pre-reconocimiento de rostros reales.....</i>	54
<i>Ilustración 15 - Diagrama de la etapa de pre-reconocimiento</i>	55
<i>Ilustración 16 - Componentes izquierda y derecha de una imagen estereoscópica</i>	56
<i>Ilustración 17 - Traslape existente entre las imágenes tomadas de (a) Un rostro real (b) De una fotografía.....</i>	56
<i>Ilustración 18 - Valores obtenidos de D_p cuando se emplea la Luminosidad (Y) y la Intensidad (I); (a) empleando el formato YCbCr y (b) empleando el formato HSI</i>	58
<i>Ilustración 19 - Fusión de las imágenes izquierda y derecha</i>	64
<i>Ilustración 20 - Diagrama a bloques del nivel de fusión a nivel características</i>	66



Ilustración 21 Diagrama a bloques del sistema de reconocimiento de imágenes estereoscópicas (fase de entrenamiento y fase de identificación)..... 67

Ilustración 22 Interfaz principal del sistema..... 67

Ilustración 23 Interfaz de entrenamiento cuando la cámara izquierda y cámara derecha iniciadas correctamente..... 68

Ilustración 24 Imagen estereoscópica obtenida del sistema 69

Ilustración 25 Ilustración 23 Almacenando imagen la base de datos 70

Ilustración 26 Interfaz de reconocimiento cuando la cámara izquierda y cámara derecha iniciadas correctamente..... 70

Ilustración 27 Imagen segmentada..... 72

Ilustración 28 resultado si la persona fue identificada o no identificada 73

Ilustración 29 - Ejemplos de imágenes contenidas en la base de datos 74

ÍNDICE DE TABLA

Tabla 1: Algoritmos de reconocimiento facial usando 3D..... 35

Tabla 2: Algoritmos usando combinación de datos 3D y 2D 36

Tabla 3 Empresas que ofrecen aplicaciones de reconocimiento facial 42

Tabla 4 Operacionalización de variable Dependiente 45

Tabla 5 Operacionalización de variable Independiente 45

Tabla 6 Esquema de la hoja de verificación 47

Tabla 7 Matriz de Confusión..... 48

Tabla 8 Matriz de Confusión de Resultados obtenidos 50

Tabla 9 Porcentaje de identificación..... 75

Tabla 10 Porcentajes de verificación usando el método de PCA 75

Tabla 11 Porcentajes de verificación usando el método de LDA 76



RESUMEN

El reconocimiento facial para los seres humanos es sumamente sencillo. Sin embargo, no se puede decir lo mismo de los sistemas computacionales, y más aún cuando se habla de un reconocimiento facial en 3D. El problema que afronta este trabajo de investigación es cómo reconocer rostros en imágenes estereoscópicas, esto se hace con el fin de incentivar la mejora de aplicaciones o sistemas que emplean técnicas de reconocimiento facial. Es por ello, que el objetivo que sigue la presente tesis se centra en reconocer rostros en imágenes estereoscópicas mediante un sistema computacional que emplea algoritmos de modelos faciales con la finalidad de obtener mejores resultados de clasificación de rostros reales. La presente investigación será Tecnológica con una población compuesta por imágenes estereoscópicas, de las cuales se tomó una muestra de 348 imágenes. Después de analizar las características de los algoritmos de modelos faciales, se decidió utilizar el algoritmo de Análisis de Componentes Principales (PCA). Para la obtención de los datos se aplicaron técnicas de análisis documental y observación. Después de realizar las pruebas se obtuvieron resultados como nivel de eficacia del 94.54% en el reconocimiento de rostros, el sistema computacional presentó un nivel de especificidad del 93.78% y un nivel de sensibilidad del 95.32%.

Palabras clave:

Visión por computador, modelos faciales, reconocimiento facial en 3D, imágenes estereoscópicas.



ABSTRACT

Facial recognition for humans is extremely simple. However, the same can not be said of computer systems, and even more so faces in stereoscopic images, this is done in order to encourage the improvement of applications or systems that use facial recognition techniques. Therefore, the objective of when talking about facial recognition in 3D. The problem facing this research is how to recognize this thesis is to recognize faces in stereoscopic images by means of a computer system that uses algorithms of facial models in order to obtain better results of classification of real faces. The type of research that was carried out will be Technological with a population composed by stereoscopic images, of which a sample of 348 images was taken. After analyzing the characteristics of the algorithms of facial models, it was decided to choose the algorithm of Analysis of Principal Components (PCA). To obtain the data, techniques of documentary analysis and observation were applied. After performing all the tests, an efficiency level of 94.54% was obtained in the recognition, and the computer system had a specificity level of 93.78% and a sensitivity level of 95.32% .

Keywords: Computer vision, facial patterns, facial recognition 3D, stereoscopic images



INTRODUCCIÓN

En la actualidad, los sistemas de reconocimiento, especialmente en campo de reconocimiento facial ha sido y sigue siendo una de las áreas con mayores investigaciones en comparación con los demás sistemas de reconocimiento de patrones (Francesc Serratosa, 2011).

Las técnicas de reconocimiento facial en 2D, son numerosas que van desde al análisis del rostro como un todo hasta el análisis centrado en puntos de interés, sin embargo, la complejidad del rostro hace que estas técnicas presenten errores al clasificar rostros reales, y esto se debe a diferentes factores como ambientales, cambios de posición, variabilidad del rostro, presencia de oclusiones, etc. Para García et al 2014, afirma que estos sistemas disminuyen su funcionalidad drásticamente cuando operan en entornos no controlados.

Como solución a este problema se introduce otra técnica conocida como reconocimiento facial en 3D, razón por el cual, en el presente trabajo se opta por utilizar dicha técnica, teniendo como eje central el desarrollo un sistema basado en modelos faciales utilizando imágenes estereoscópicas a fin de obtener buenos resultados de clasificación de rostros reales.



CAPÍTULO I: PROBLEMA DE INVESTIGACIÓN

1.1 Situación Problemática

En la actualidad el estudio del Reconocimiento Facial se ha realizado no sólo en la Informática, sino también desde otras disciplinas como la Psicología, la Medicina Forense, la Antropometría, entre otros. (Moreno, 2004) Otra forma de emplear el reconocimiento facial es para la búsqueda de personas con antecedentes policiales, reconocer rostros en redes sociales. Por ejemplo, podemos mencionar la aplicación de FindFace (<https://findface.ru/>), que permite buscar el rostro de una persona dentro de su propia red social, logrando identificarla con un 70% de efectividad.

El reconocimiento facial, según explica Alberto Albiol, en un artículo publicado en (<http://www.agenciasinc.es/>) dependen de un gran número de circunstancias, tales como la iluminación de la cara, la resolución de la imagen, la expresión del rostro o el aspecto variable de cada individuo posee una mayor importancia dentro del campo de investigación. El principal reto según explica Cama (2015) es optimizar el reconocimiento cuando las caras no mantengan una posición frontal con respecto a la cámara. Los sistemas actuales, en su gran mayoría trabajan como un objeto bidimensional (Blázquez, 2013; Becerra ,2014; Zúñiga ,2015), permitiendo de esta manera que estos sistemas puedan ser vulnerados colocando imágenes de alta resolución o inclusive videos.

En el Perú la detección de rostros es cada vez más común y es usada en distintas aplicaciones por diferentes instituciones, por ejemplo el Registro de Identificación y Estado Civil, cuenta con un avanzado sistema biométrico de reconocimiento facial para la identificación de los ciudadanos peruanos (RENIEC, 2014).



Actualmente se busca optimizar la calidad del reconocimiento facial a través del uso de datos tridimensionales (3D), dado que los mismos son menos sensibles que los bidimensionales, a los cambios de luz, variaciones de postura en la imagen, maquillaje y expresiones faciales (Pereyra, 2014). Otro punto a tener en cuenta es que, al trabajar con imágenes de textura, obtenidas mediante imágenes estereoscópicas aportan información determinante de áreas de la cara donde no existe una gran variación en la estructura geométrica (García et al 2014). Por ejemplo, obtener puntos faciales de profundidad de la nariz o de la boca (Blázquez, 2013).

Es por ello que en el presente trabajo nos enfocamos por desarrollar un sistema de reconocimiento facial, que aproveche las ventajas de las imágenes estereoscópicas para lograr un reconocimiento facial en 3D, obteniendo información de profundidad, textura, y un sistema más robusto ante cambios de iluminación, pose y expresión.

1.2. Formulación del Problema

¿Cómo reconocer rostros en imágenes estereoscópicas?

1.3. Delimitación de la Investigación

El presente trabajo se desarrolló en la Universidad Privada Señor de Sipán, en el periodo de Setiembre a Diciembre para el cual se realiza el análisis comparativo de las técnicas de modelos faciales, empleando imágenes estereoscópicas para el reconocimiento facial en 3D.



1.4. Justificación e Importancia de la Investigación

1.4.1 Desde el punto de vista científica

Al utilizar algoritmos de modelos faciales y de extracción de características o patrones para compararlos y de esta manera obtener el de mejor eficacia, nos permitirá dar una mejor solución a la problemática en estudio, así mismo permite que el reconocimiento facial en 3D sea más eficiente en comparación con el reconocimiento facial en 2D. Dichos aporte de esta investigación, servirá de base para otros investigadores.

1.4.2 Desde el punto de vista Social

Vivimos en una sociedad ligada más a los cambios tecnológicos por lo que los sistemas de seguridad cada vez son más completos y complejos. Evolucionando desde el uso de usuario y contraseña, más tarde factores de doble autenticación basado en claves token o tarjetas de coordenadas y, por último, se está utilizando los sistemas biométricos, con el fin de optimizar la seguridad del usuario utilizando el reconocimiento facial en 3D.

Gracias a estos avances tecnológicos, hoy, contamos con eficientes equipos, que poseen un gran rendimiento y procesamiento. Estas características permitirán que, al utilizar algoritmos de modelos faciales, la seguridad y otras utilidades sean más eficientes. Algo que no podía hacerse en décadas anteriores, donde el reconocimiento facial estaba limitado por la tecnología de aquel entonces.



1.5. Limitaciones de la Investigación

- a. Escasa información referente a la visión estereoscópica, por lo que nuestro trabajo será de aporte a esta área.
- b. No se cuenta con un equipo especializado en la adquisición de imágenes estereoscópicas, por lo que se ha diseñado un sistema equivalente en base a cámaras de uso común.

1.6. Objetivos de la Investigación

1.6.1 Objetivo general

Reconocer rostros en imágenes estereoscópicas mediante un sistema computacional que emplea algoritmos de modelos faciales.

1.6.2 Objetivos específicos

1. Realizar una investigación documental del reconocimiento de rostros en imágenes estereoscópicas.
2. Seleccionar y evaluar algoritmos de modelos faciales.
3. Desarrollar un software con los algoritmos seleccionados cuya eficacia sea mayor a 90 % .
4. Evaluar el rendimiento del sistema desarrollado con los modelos faciales.



CAPÍTULO II: MARCO TEÓRICO

2.1. Antecedentes de Estudios

Nivel Internacional

Según **Pereyra** (2014), en su trabajo de investigación denominado **“Reconocimiento facial mediante imágenes estereoscópicas para control de ingreso”**, aborda el reconocimiento de las caras a partir de sus modelos numéricos en 3D, obtenidos por técnicas de estereoscopía. Estos modelos numéricos de una misma cara si bien deben ser propias e idénticos pueden haber sido obtenidos en cada caso con la cara ubicada en posición diferente respecto del referencial del mecanismo de toma. Para eso compara los modelos numéricos lo cual identifica sus propiedades propias, es decir propiedades independientes del referencial, y no por medio de una roto-traslación que superponga las caras que se quieren comparar. Siendo el resultado parcial obtenido es que la medición y comparación mediante la curvatura de gauss, por lo cual se aplicó la comparación mediante error cuadrático medio y se obtuvo un bajo nivel de precisión de 53%, observando del análisis de la curva ROC que su comportamiento es casi opuesto. El presente trabajo, se toma como referencia por el método del sistema y el uso del modelo LDA que será uno de los modelos evaluados en este trabajo.

Según **Ceron, Salazar, Prieto** (2013), en su trabajo investigación denominado **“Reconocimiento de rostros y gestos faciales mediante un análisis de relevancia con imágenes 3d”**, propone un sistema de reconocimiento de rostros, que utiliza un conjunto de descriptores de forma 3D, seleccionados a partir de un



análisis de relevancia mediante coeficientes de Fisher en diferentes regiones del rostro que hacen parte de un modelo antropométrico del rostro. Por lo cual se realizó un análisis de relevancia con un conjunto de experimentos para reconocer individuos e identificando su género, expresión. Mostrando como resultados que la elección de características utilizando un análisis de relevancia fueron una mejora de un 12.5% , para redes bayesianas 5% y 43.7% para dos casos de redes neuronales, y un 71% con el algoritmo de optimización mínima secuencial (SMO), en la verificación e identificación de la persona bajo análisis. El presente trabajo, se toma como referencia por el método de extracción de descriptores de forma 3D que nos ayudarán a optimizar la extracción de características.

Vásquez (2014), en su trabajo de investigación denominado “**Sistema de Reconocimiento Facial Mediante Técnicas de Visión Tridimensional**”, abordó el problema de identificación facial a partir del análisis de la información tridimensional del rostro con la finalidad de determinar la identidad de personas. Se usó el método de eigenfaces o Análisis de Componentes Principales (PCA), para obtener los descriptores del rostro, luego se realiza la clasificación e identificación con la técnica de Análisis Discriminante Lineal (LDA), permitiendo el cálculo de la nube de puntos relacionada al rostro, implementando un sistema que consta de cuatro módulos: reconstrucción tridimensional, alineación y segmentación, extracción de características, clasificación y reconocimiento. En el módulo de alineación del rostro se logró reducir el tiempo de procesamiento hasta un 99% y una identificación positiva de 97.87%. El presente trabajo, se toma como referencia por el método del sistema propuesto establecido en cuatro partes bien definidas y el uso del modelo PCA.



Castellani (2005), en su trabajo de investigación denominado “**3D Face Recognition Using Stereoscopic Vision**”, presenta un nuevo sistema completo para el reconocimiento facial 3D. El reconocimiento facial 3D presenta varias ventajas frente al reconocimiento facial 2D, como, por ejemplo, la invariancia a las condiciones de iluminación. Este sistema usa una metodología estéreo, que no requiere ningún sensor de alcance que sea muy costosa. Se usa máquinas de vectoriales de soporte (SVM) para la clasificación de imágenes en 3D del rostro usando coeficientes Multilevel B-Splines. El sistema alcanza en algunos casos una exactitud de clasificación perfecta, y en los otros sólo obtiene un error. El presente trabajo, se toma como referencia el uso de Máquinas Vectoriales de Soporte (SVM) para la clasificación de imágenes en 3D. Asimismo, se toma como referencia por el método estéreo que nos ayudará a optimizar el uso de 3D en imágenes estereoscópicas.

Nivel Nacional

Según **Villa** (2011), en su trabajo de investigación denominado “**SISTEMA DE RECONOCIMIENTO DE ROSTROS**”, propone un sistema computacional capaz de reconocer rostros a partir de imágenes faciales capturadas a través de una cámara web. El sistema compara paramétricamente la imagen adquirida con aquellas almacenadas en una base de datos (usuarios registrados). Para ello, se desarrolló una aplicación en MATLAB, utilizando la interfaz visual Guide. En esta aplicación se pueden controlar los parámetros utilizados para el proceso de reconocimiento. El método de reconocimiento empleado es basado en el procesamiento de imágenes eigenfaces o PCA. Los resultados parciales obtenidos en este trabajo de investigación fueron 92.26% de acierto. El presente



trabajo, se toma como referencia por el uso de imágenes “eigenfaces” o PCA y poder adaptarlo a imágenes estereoscópicas.

Según **Ocsa** (2008), en su trabajo de investigación denominado **“Reconocimiento de Rostros mediante Puntos Característicos Locales”**, presenta una comparación entre dos métodos de extracción de características locales. La propuesta para la identificación de puntos característicos locales en los rostros usa la técnica Active Shape Model (ASM). El primer método crea el vector característico considerando las distancias que forman en el grafo de Delaunay sobre los puntos localizados por ASM. El segundo método, con los mismos puntos se define una región localizada donde cada punto característico de 7x7 píxeles sirve para formar vectores características de componentes Wavelets en base a una familia de transformadas Wavelet en distintas escalas y orientaciones. Este enfoque permite que la propuesta sea invariante a rotación, traslación y escala. Con las características obtenidas se realiza una comparación paramétrica con los vectores de las imágenes almacenadas en la base de datos conformantes de un conjunto de entrenamiento. Los experimentos fueron hechos con la base de datos de BioID Face Data Base, con el cual se obtuvo un grado promedio de reconocimiento de 93 % (método 1) y 96 % (método 2). El presente trabajo, se toma como referencia por el uso de extracción de características locales o modelos faciales.

2.2. Estado del Arte.

Durante los últimos años se han realizado innovaciones tecnológicas en aspectos de hardware, creando equipos con alta capacidad de procesamiento, por lo que



se ha generado una gran innovación en el desarrollo de aplicaciones en aspectos de procesamiento de señales como el sonido, imágenes 2D, e imágenes 3D.

Se han realizado estudios sobre los modelos de reconocimiento de imágenes en dos dimensiones y en tres dimensiones de análisis de imágenes con variedades de técnicas y algoritmos biométricos, que requieren de imágenes con gran calidad de resolución, siendo de mucho interés en la seguridad debido a que los métodos convencionales, como solicitar, una contraseña o una tarjeta electrónica no son suficientemente seguros.

Podemos identificar dos familias de técnicas de reconocimiento facial:

- Técnicas lineales: PCA, DCT, LPP, LDA, y técnicas no lineales: KPCA, ISOMAP).
- Técnicas basadas en modelos, como técnicas 3D Morhable Model, etc.

Siendo el modelo 3D una de las técnicas con mayor interés basadas en modelos que construyen un modelo lo más descriptivo posible del rostro de la persona, capaz de detectar con precisión las variaciones faciales como distancia entre ojos, grosor de los labios, distancia entre ojos y labios, etc.

2.3. Bases teóricas científicas.

2.3.1. Visión Estereoscópica

2.3.1.1. Introducción

La visión estereoscópica es la técnica que permite recoger información de visión tridimensional o crear la ilusión de profundidad. Según Rodríguez (2009) indica que la visión estéreo hace referencia a la habilidad de inferir información de una estructura determinada en el espacio a partir de 2 o más imágenes tomadas desde diferentes puntos de vista. El modelo de dos cámaras se hace referencia al



propio modelo biológico de los seres vivos. Los dos ojos, situados en posiciones diferentes, recogen imágenes distintas en la retina, de las cosas que tienen delante. Estas diferencias son interpretadas en el cerebro para calcular la distancia a la que se encuentran los objetos mediante la técnica del paralaje. El cálculo de las distancias sitúa los objetos que estamos viendo en el espacio tridimensional, obteniendo una sensación de profundidad o volumen. Solo las variaciones horizontales, producidas por la diferente situación de los ojos, las que originan la sensación de profundidad.

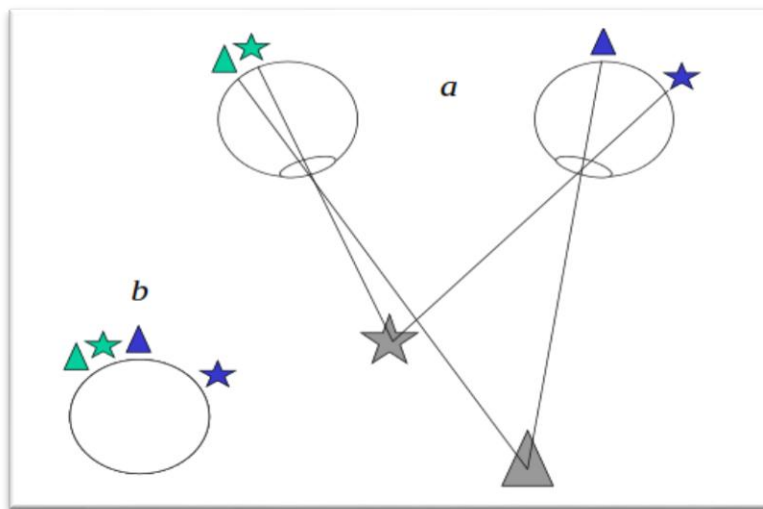


Ilustración 1 - (a) Sistema de visión estereoscópica biológico; (b) Superposición de las imágenes de ambos ojos, donde se observa la separación relativa de cada objeto de la escena (a menor distancia, más lejos se encuentra el objeto; a menor distancia, más cerca es.

Fuente: Guerrero et al. (2010)

Este modelo de la percepción 3d a través de nuestros ojos puede ser llevado a la parte computacional, así como podemos observar en la ilustración 2 la vista superior de un sistema de visión estéreo usado en un robot.



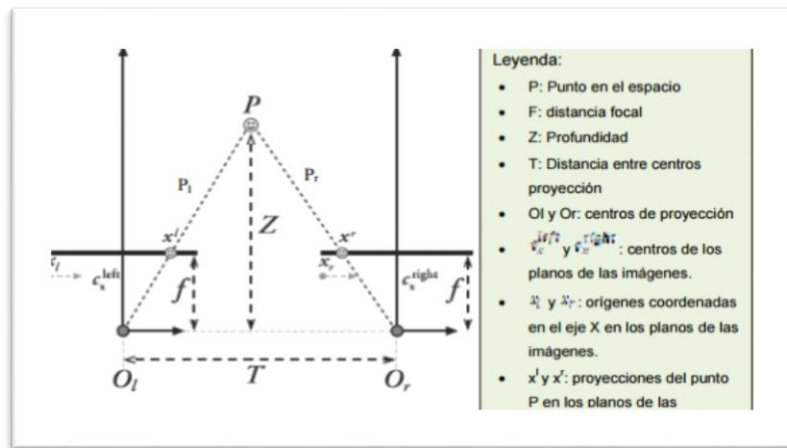


Ilustración 2 Vista superior de un sistema de visión estereoscópica

Fuente: Trabajo de investigación "Robot móvil con visión estereoscópica para la localización de objetos", autor Jorge Antonio Rodríguez Rodríguez

2.3.1.2. Historia

La humanidad, desde sus comienzos, ha pretendido representar las imágenes de su alrededor en forma tridimensional (Fernández Sánchez, 2000). Es el caso de los primeros pintores en las famosas cuevas de Lascaux, datadas de 16000 años. En la cueva francesa, bajo la luz tenue de la antorcha, pueden tomar una calidad tridimensional las representaciones de ciervos, búfalos u otras escenas de animales como se muestra en la ilustración 3.

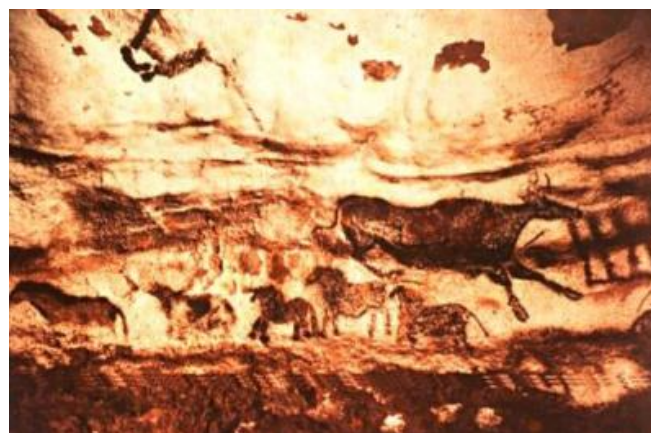


Ilustración 3 - Pintura de cueva de Lascaux en Francia.

Fuente: Cuevas de Lascaux – Wikipedia, the free encyclopedia



El interés por la tridimensionalidad visual se remonta a la Grecia clásica. En el siglo III A.C. el matemático griego Euclides, conocido como el padre de la geometría estudió en su tratado titulado "Óptica" la relación entre la visión en relieve y el hecho de que los seres humanos dispongan de dos ojos. Ya en el siglo II de nuestra era, el médico griego Galeno describiría también lo que supuso una profundización de una teoría razonada de la visión de la perspectiva a través de los dos ojos. Supo por la observación que si se tapaba un ojo, la imagen del entorno cambiaba ligeramente (Fernández Sánchez, 2000). Sería en 1833 cuando se descubre un dispositivo notablemente sencillo que reproduce la tridimensionalidad: El estereoscopio de Charles Wheatstone (ver ilustración 4), un físico británico que trabajó antes de que se inventara la fotografía, fue el primer y principal avance que conduciría a toda una larga cadena de aparatos e invenciones para ver en relieve. Dicho estereoscopio constaba de dos espejos colocados en ángulo con respecto al observador y una imagen situada frente a cada espejo. Wheatstone llegó a la conclusión de que las diferencias visuales, que presentan ambos ojos eran el origen de la percepción en relieve. Pocos años después, David Brewster, un científico e inventor escocés, diseñó y construyó la primera cámara fotográfica estereoscópica, con la que obtuvo las primeras fotografías en relieve. Y construyó también un visor con lentes para observarlas. (Fernández Sánchez, 2000).





Ilustración 4 - Estereoscopio de Charles Wheatstone

Fuente: La vida en 3D, Fernández Sánchez (2000)

2.3.1.3. Sistema de visión estereoscópica

- **Visión Paralela:** Es uno de los primeros y más antiguos sistema de visualización. El observador debe centrar su atención en un punto en el infinito, manteniendo los ejes de visión paralelos. Las imágenes no pueden tener un tamaño superior a 65 mm entre sus centros. Para facilitar la visión paralela, se utilizan lentes para acomodar la vista y evitar que cada ojo perciba lateralmente la imagen correspondiente al otro ojo.

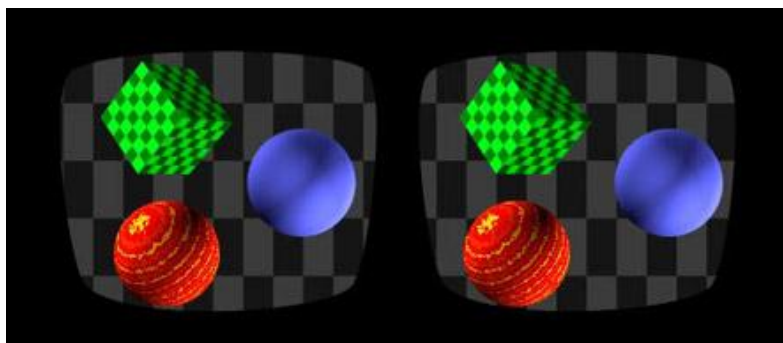


Ilustración 5 - Ejemplo de visión Paralela

Fuente: Jorge Westendarp Fotografía 3D estereoscópica real



- **Visión cruzada:** En este tipo de sistema, las líneas de visión de ambos ojos se cruzan, centrándose en un punto medio más cercano. De esta forma el ojo derecho observa la imagen izquierda y viceversa.



Ilustración 6 - Ejemplo de visión cruzada

Fuente: SABIA Dpto. Tecnologías de la Información y las Comunicaciones
Universidad de la Coruña

- **Anáglifos:** Son imágenes de dos dimensiones capaces de provocar un efecto tridimensional cuando se visualiza con lentes espaciales, utilizan filtros de colores, como rojo-azul, rojo-verde etc. La imagen presentada por ejemplo en rojo no es vista por el ojo que tiene un filtro del mismo color, pero sí que ve la otra imagen en azul o verde. Este sistema, por su bajo costo, se emplea sobre todo en publicaciones, así como también en monitores de ordenador y en el cine. Presenta el problema de la alteración de los colores, pérdida de luminosidad y cansancio visual después de un uso prolongado. Normalmente se sitúa el filtro rojo en el ojo izquierdo, y el azul en el ojo derecho.



Ilustración 7 - Ejemplo de Anáglifos

Fuente: SABIA Dpto. Tecnologías de la Información y las Comunicaciones
Universidad de la Coruña

- **Polarización:** Es un proceso por el cual un conjunto se establecen características que determinan de dos o más zonas, se usa luz polarizada para separar las imágenes izquierda y derecha. El sistema de polarización no altera los colores. Esta polarización funciona poniendo un filtro polarizado en cada objetivo.
- **Obturación:** En este caso, las dos imágenes son mostradas por un monitor de forma alternativa, con una frecuencia de refresco suficiente. El observador debe utilizar unas gafas dotadas de dos lentes de cristal líquido (LCD) que se oscurezcan alternativamente de forma sincronizada con el monitor.

2.3.1.4. Calibración del sistema estéreo

La calibración de cualquier sistema es la etapa más importante, debido a que, mediante esta etapa o procedimiento se va a obtener los parámetros de dicho sistema, Rodríguez (2009).



En primer lugar, se realizar una serie de fotos con la cámara que se va a utilizar. Para describir este proceso se trabajará con una imagen de un tablero de ajedrez. Las imágenes deben ser tomadas en diferentes ángulos con la finalidad de obtener un buen calibrado de la imagen.

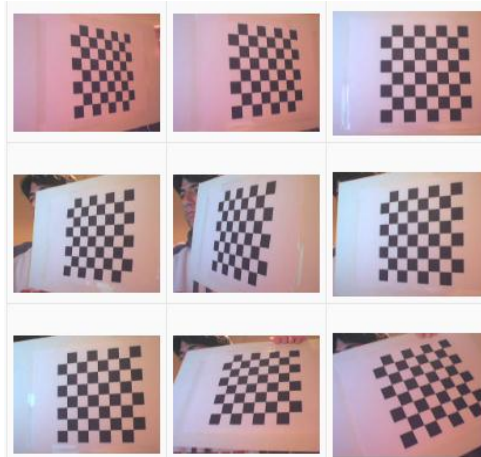


Ilustración 8 - Imagen de entrada de un tablero de ajedrez tomado en diferentes ángulos

Fuente: SABIA Dpto. Tecnologías de la Información y las Comunicaciones
Universidad de la Coruña

Después de obtener las imágenes, se crea una imagen para cada imagen original conocido como "findChessBoardCornerGuesses". Este proceso permite buscar los puntos interiores en el tablero y los marcará con colores.



Ilustración 9 - Resultado de la imagen después de aplicar el método "findChessBoardCornerGuesses"

Fuente: SABIA Dpto. Tecnologías de la Información y las Comunicaciones
Universidad de la Coruña

A continuación, se emplea otra técnica conocida como “Backprojecting”, donde devuelve toda la información de rotación, traslación, parámetros de la cámara fotográfica. Así como también la proyección de las coordenadas de la escena en 3D de la imagen. Para terminar con este proceso de calibración, se emplea la técnica de “Undistorting”, donde se combina la imagen de una manera que las líneas paralelas en la imagen sean paralelas a la imagen de la escena.

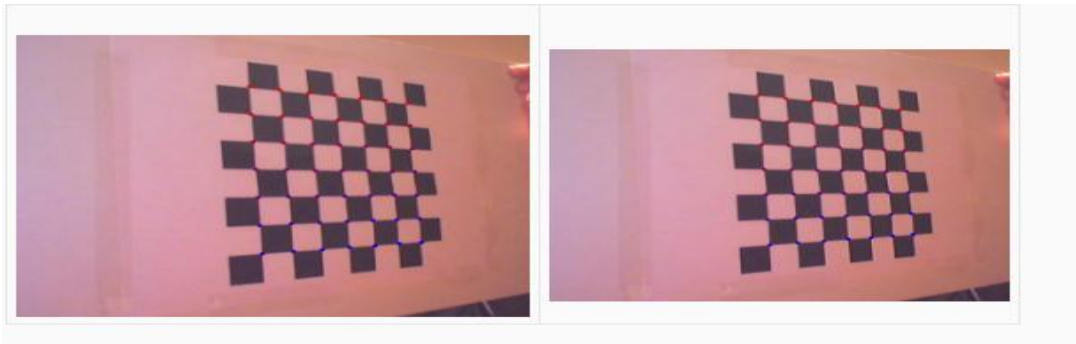


Ilustración 10 - Resultado final de la calibración de imagen

Fuente: SABIA Dpto. Tecnologías de la Información y las Comunicaciones
Universidad de la Coruña

2.3.1.5. Pasos en el proceso de la visión estereoscópica

El proceso completo de visión estereoscópica según Martínez (2010), contempla seis pasos importantes:

a. Adquisición de imágenes:

La adquisición de imágenes puede realizarse de muchas formas distintas. Por ejemplo, las imágenes pueden ser tomadas simultáneamente en el tiempo, o mediante intervalos de tiempo con una duración determinada. Pueden ser tomadas desde localizaciones y direcciones ligeramente distintas o radicalmente diferentes.

b. Geometría del sistema:



El modelo en el que nos centraremos es uno con dos cámaras que tienen sus ejes ópticos paralelos, siendo la distancia que los separa la línea base. Quedando sus ejes ópticos perpendiculares a la línea base, y sus líneas de exploración o líneas epipolares paralelas a la línea base.

Las líneas epipolares son líneas que unen la imagen izquierda y la imagen derecha de un mismo punto. Cualquier punto del espacio tridimensional unido a los dos centros de proyección de las cámaras define un plano, llamado plano epipolar. La intersección de un plano epipolar con el plano de proyección de una cámara define una línea epipolar.

c. Extracción de las características:

Para las técnicas de estereovisión basadas en las características se toman representaciones simbólicas obtenidas de las imágenes de intensidad en vez de utilizar las intensidades directamente. Algunas de las características cuyo uso está más extendido son: puntos de borde aislados, cadenas de puntos de bordes o regiones delimitadas por bordes.

En los tres tipos de características que se acaban de mencionar se suelen tener en cuenta los puntos de borde para su obtención. De lo anterior se deduce que los puntos de borde utilizados como primitivas son muy importantes en cualquier proceso de visión estereoscópica y en consecuencia suele ser habitual el hecho de extraer los puntos de borde de las imágenes del par estereoscópico.

Una vez que se han extraído los correspondientes puntos de borde algunos métodos utilizan cadenas de puntos de borde, dichas cadenas pueden ser segmentos rectos, segmentos no rectos, cadenas cerradas formando estructuras geométricas con forma definida o estructuras de formas geométricas desconocidas.



d. Correspondencia de las imágenes:

En el paso de la correspondencia de imágenes, bien utilizando píxeles o características, se debe determinar para un punto del espacio tridimensional, cuáles su proyección en cada imagen del par estereoscópico.

Al comienzo del proceso de correspondencia ya se tienen los vectores con los atributos de las características consideradas. Con estos vectores, comparando los valores que toman sus atributos se debe establecer una correspondencia local entre características. Esta correspondencia se determina mediante alguna métrica que proporcione cuál es el grado de similitud para dos vectores de atributos, como por ejemplo la distancia Euclídea y otras distancias, se debe comprobar su consistencia, para lo cual se comienza otro proceso de correspondencia, pero en este caso de naturaleza global.

e. Determinación de la distancia:

Culminando el proceso de correspondencia de las imágenes, la determinación de la profundidad es un proceso que reduce a una simple triangulación. En algunas ocasiones, cuando se intenta hallar la distancia a la que se encuentra una característica se presentan algunas dificultades debidas a una falta de precisión.

f. La interpolación en caso sea necesario:

Uno de los modos más sencillos de solucionar el problema de la interpolación es interpretar el mapa de disparidad como un muestreo de una función de profundidad continua, y utilizar un método de interpolación tradicional para hallar la función continua que la aproxime. Así se obtendría una función continua con la que se puede obtener la profundidad para cualquier punto del espacio tridimensional, consiguiendo transformar un mapa de profundidad disperso en denso. Si el mapa de profundidad disperso capturaba los principales cambios en



la profundidad, el mapa que se alcanza será válido. Métodos que se pueden utilizar para interpolar son: la interpolación de Lagrange, la interpolación de Hermite y la interpolación mediante Splines.

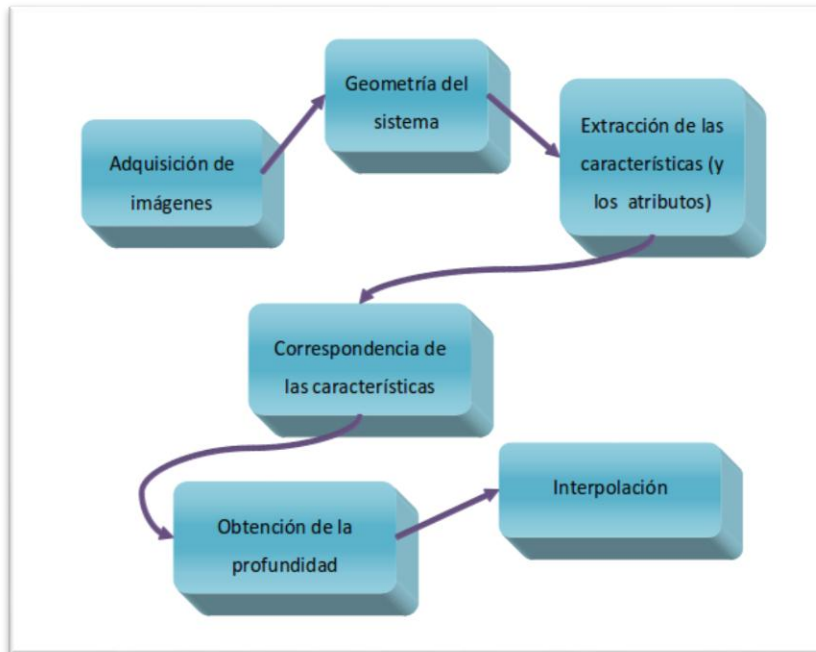


Ilustración 11 - Proceso completo de visión estereoscópica

Fuente: Martínez (2010).

2.3.2. Reconocimiento de imágenes 3D

2.3.2.1. Definición

Existió un test de evaluación para los algoritmos comerciales y prototipos de técnicas de reconocimiento facial, llamado Face Recognition Vendor Test (FRVT) realizado por The National Institute of Standards and Technology (NIST), cuyas evaluaciones fueron realizadas en el 2000, 2002 y 2006 por última vez. Mediante el análisis de estas pruebas realizadas se pudo obtener información para determinar cómo las técnicas de reconocimiento facial pueden ser optimizadas y



recomendar futuras direcciones de investigación. Como resultado de la última evaluación realizada en el 2006, como señala (Phillips et al., 2007), se obtuvo que desde 1993 al 2006 se ha optimizado en dos órdenes de magnitud el porcentaje de error (ver ilustración 12). Para realizar la comparación se utiliza el estudio realizado con eigenfaces de (Pentland et al., 1991). También se han tenido en cuenta para medir el avance de los algoritmos, los resultados de las evaluaciones del test Face Recognition Technology (FERET), que fuera realizado previamente en 1994, 1995 y 1996.

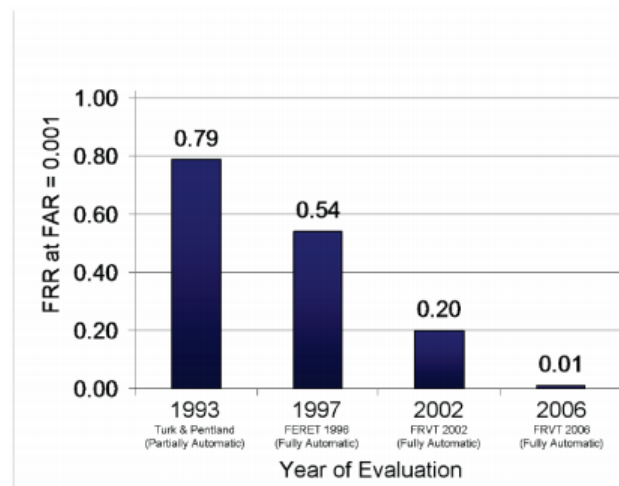


Ilustración 12 - La reducción en el promedio de error para los algoritmos de reconocimiento de rostros en base a las evaluaciones de FERET, FRVT 2002 y FRVT 2006.

Fuente: Phillips et al (2007)

En base al análisis, también se concluyó, como indica (Bowyer et al, 2005), que el actual estado del arte del reconocimiento facial 2D es insuficiente para la alta demanda biométrica de las aplicaciones, a partir de lo cual intentar usar información 3D se ha vuelto la dirección de búsqueda emergente con la esperanza de hacer que el reconocimiento facial sea más exacto y robusto. Por esta razón, a pesar de que históricamente es muy poco el trabajo realizado en esta área, en los últimos años han aumentado ampliamente las investigaciones



usando información 3D. Actualmente, la mayoría de los estudios de reconocimiento facial 3D se han orientado hacia el uso de técnicas holísticas o de técnicas de apareo de superficies 3D. (Gupta et al, 2007). Entre los métodos holísticos se encuentran por ejemplo PCA, LDA, que representan un avance de las técnicas que fueron bastante exitosas con imágenes 2D. Incluso el algoritmo 3D PCA es también, considerado como un referente de evaluación de la performance de otros algoritmos de reconocimiento facial 3D.

2.3.2.2. Técnicas de adquisición de imágenes 3D

Según Pereyra (2014), para obtener los modelos 3D faciales se pueden utilizar distintas técnicas entre ellas:

- **Escáner Facial (Laser Facial Scanning):** Este escáner facial genera un modelo 3D exacto, detallado y registrando la textura del rostro, pero requiere el apoyo de la persona bajo análisis por un tiempo en minutos mientras el escáner recorre alrededor de su cabeza. No es un método muy usado debido a su alto costo. Recientes investigaciones fueron realizadas sobre modelos 3D de las caras, obtenidos mediante escáneres de láser, como sucede en la toma realizada para la base de datos de caras BJUT-3D llevada a cabo por la universidad de Beijing.
- **Tomografía Computada (Computed Tomography Scan):** Esta tomografía posee un elevado costo, lo cual la persona debe exponerse a radiación y no es un método de mucho consumo.
- **Estereoscopía (Stereo-photogrammetry):** es una técnica pasiva donde mediante el uso de dos o más cámaras debidamente calibradas se obtiene la reconstrucción tridimensional de la superficie facial. Presenta dificultades por la



necesidad de una correcta calibración. Pero por otro lado tiene la ventaja de no ser intrusiva para el sujeto y no requiere la cooperación del mismo.

El uso de la estereoscopía para la obtención de un modelo tridimensional facial es más viable y sencillo, dado que se puede realizar a partir de dos imágenes normalizadas, la toma se hace casi instantáneamente y el equipo necesario a utilizar es de muy bajo costo comparado con un escáner tridimensional. Sin embargo, (Samani et al, 2006) apunta como desventaja que si bien varias de las técnicas estereoscópicas utilizadas reportaron una exactitud por encima del 90%, una precisa calibración de las cámaras fue necesaria.

2.3.2.3. Métodos de reconocimiento facial 3D

Entre los métodos de reconocimiento facial tridimensionales, podemos distinguir dos grandes grupos en base a la información que utilizan:

- 3D Puro

Estas técnicas solo utilizan información de imágenes tridimensionales. En la publicación de Bowyer et al (2005), describe las diferentes técnicas que forman parte de este método. En Tabla 1 se puede ver un resumen detallado de las técnicas de 3D puro.



Tabla 1: Algoritmos de reconocimiento facial usando 3D

Recognition algorithms using 3D shape alone						
Author, year, reference	Persons in dataset	Images in dataset	Image size	3D face data	Core matching algorithm	Reported performance
Cartoux, 1989 [12]	5	18	Not available	Profile, surface	Minimum distance	100%
Lee, 1990 [26]	6	6	256 × 150	EGI	Correlation	None
Gordon, 1992 [21]	26 train 8 test	26 train 24 test	Not available	Feature vector	Closest vector	100%
Nagamine, 1992 [39]	16	160	256 × 240	Multiple profiles	Closest vector	100%
Achermann, 1997 [3]	24	240	75 × 150	Range image	PCA, HMM	100%
Tanaka, 1998 [52]	37	37	256 × 256	EGI	Correlation	100%
Achermann, 2000 [2]	24	240	75 × 150	Point set	Hausdorff distance	100%
Chua, 2000 [17]	6	24	Not available	Point set	Point signature	100%
Hesher, 2003 [22]	37	222	242 × 347	Range image	PCA	97%
Lee, 2003 [27]	35	70	320 × 320	Feature vector	Closest vector	94% at rank 5
Medioni, 2003 [34]	100	700	Not available	Point set	ICP	98%
Moreno, 2003 [38]	60	420	2.2K points	Feature vector	Closest vector	78%
Pan, 2003 [42]	30	360	3K points	Point set, range image	Hausdorff and PCA	3–5% EER, 5–7% EER
Lee, 2004 [28]	42	84	240 × 320	Range, curvature	Weighted Hausdorff	98%
Lu, 2004 [30]	18	113	240 × 320	point set	ICP	96%
Russ, 2004 [49]	200 FRGC v1	468	480 × 640	Range image	Hausdorff distance	98% verification
Xu, 2004 [57]	120 (30)	720	Not available	Point set + feature vector	Minimum distance	96% on 30, 72% on 120
Bronstein, 2005 [11]	30	220	Not available	Point set	“canonical forms”	100%
Chang, 2005 [16]	466 FRGC v2	4007	480 × 640	Point set	multi-ICP	92%
Gökberk, 2005 [20]	106	579	Not available	Multiple	Multiple	99%
Lee, 2005 [29]	100	200	Various	Feature vector	SVM	96%
Lu, 2005 [31]	100	196 probes	240 × 320	Surface mesh	ICP, TPS	89%
Pan, 2005 [41]	276 FRGC v1	943	480 × 640	Range image	PCA	95%, 3% EER
Passalis, 2005 [44]	466 FRGC v2	4007	480 × 640	Surface mesh	Deformable model	90%
Russ, 2005 [50]	200 FRGC v1	398	480 × 640	Range image	Hausdorff distance	98.5%

Fuente: Bowyer et al. (2005)

- 3D Multimodal

Son técnicas que utilizan la información conjunta de imágenes bidimensionales y tridimensionales. Como se señala, en el estudio realizado por Bowyer et al. (2005) sobre métodos 3D multimodales, no fue hasta el año 2000 que se empezaron a evaluar algoritmos que combinan resultados de datos en 2D y 3D. La forma más adecuada es utilizar separadamente los datos 2D y 3D para luego combinar los resultados. En general, en base a los resultados reportados dicha combinación tiene una respuesta mucho mejor que cuando se utilizan las mismas técnicas 2D y 3D por separado. Ver tabla 2 para un detalle de las investigaciones sobre los métodos multimodales.



Tabla 2: Algoritmos usando combinación de datos 3D y 2D

Recognition algorithms combining use of 3D and 2D data						
Author, year, reference	Persons in dataset	Images in dataset	Image size	3D face data	Core matching algorithm	Reported performance
Lao, 2000 [25]	10	360	480 × 640	Surface mesh	Minimum distance	91%
Beumier, 2001 [4]	27 gallery 29 probes	81 gallery, 87 probes	Not available	Multiple profiles	Minimum distance	1.4% EER
Wang, 2002 [56]	50	300	128 × 512	Feature vector	SVM, DDAG	>90%
Bronstein, 2003 [10]	157	Not available	2250 points	Range, point set	PCA	Not reported
Chang, 2003 [14]	200 (275 train)	951	480 × 640	Range image	PCA	99% 3D + 2D, 93% 3D only
Tsalakanidou, 2003 [55]	40	80	100 × 80	Range image	PCA	99% 3D + 2D, 93% 3D only
Godil, 2004 [19]	200	400	128 × 128	Range image	PCA	82% rank 1
Papatheodorou, 2004 [43]	62	806	10,000 points	Point set	ICP	100–66%
Tsalakanidou, 2004 [54]	50	3000	571 × 752	Range image	EHHM per mode	4% EER
Hüsken, 2005 [23]	466	4,007 FRGC v.2	480 × 640	hier. graph	graph match	93% verification at 0.01 FAR
Lu, 2005 [32]	100	598	320 × 240	Point set	ICP, LDA	91%
Maurer, 2005 [33]	466	4007 FRGC v.2	480 × 640	Surface mesh	ICP, Neven	87% verification at 0.01 FAR

Fuente: Bowyer et al. (2005)

Otra posible clasificación de los sistemas de reconocimiento facial 3D se basa en la forma de representación de la cara teniendo en cuenta lo descrito por (GÖkberk, 2006), que está influenciado por lo especificado en (Akarun et al, 2005). Se pueden definir unos 5 grupos:

- Nube de puntos (Point Cloud)

La superficie facial es representada mediante una nube de puntos 3D, que son un conjunto de muestras en coordenadas (x,y,z). Es la forma más básica de representación, así como también la más difícil de trabajar. El método más utilizado con esta forma de representación es ICP (Iterative Closest Point) en el cual dadas dos nubes de puntos de distintas caras, el algoritmo busca mediante rotación y traslación en forma iterativa alinear un modelo al otro. Una de las desventajas de este algoritmo es el alto costo computacional que tiene comparar cada cara de prueba con la galería de caras a fin de realizar el reconocimiento.

- Imágenes de profundidad (Depth Images)

Se utiliza la intensidad de las imágenes en 2D para obtener la profundidad. Cada píxel en la imagen de profundidad representa la distancia del punto 3D a la cámara.



Con las imágenes de profundidad el problema de reconocimiento facial 3D puede ser tratado como un problema de 2D, por lo cual pueden utilizarse cualquiera de los algoritmos anteriormente mencionados para el reconocimiento 2D como han de ser PCA, LDA.

En este método además de la imagen 2D con sus valores de intensidad, se usa una imagen llamada 2½ D, donde la profundidad se codifica directamente en la imagen 2D, los valores de los píxeles se corresponden con la misma. Esta representación permite utilizar las técnicas de reconocimiento de imágenes 2D. La técnica eigenface se aplica en la imagen 2D, 2 ½ D y en el espacio compuesto (formado por ambas imágenes 2D y 2½ D), obteniéndose como resultado que la precisión del reconocimiento es menor que la citada en otros estudios, lo cual se aduce a que no se ha tenido extenso cuidado en el proceso de adquisición de datos. Los resultados obtenidos también evidencian que en el espacio compuesto, donde se combina información de intensidad y profundidad se obtienen mejores niveles de reconocimiento que en los casos que se utiliza la información por separado.

- Curvas Faciales (Facial Curves)

Se extraen curvas verticales, horizontales y del contorno de la superficie facial y se representa la misma utilizando características obtenidas de dichas curvas 2D.

Como primer trabajo en el campo tridimensional podemos ubicar el realizado por Cartoux en 1989, según Bowyer et al. (2005). En el mismo se realiza una segmentación de una imagen de rango basándose en la curvatura principal para encontrar un plano bilateral de simetría de la cara, el cual se utilizaba para normalizar la pose de la misma. Luego se utilizan métodos para comparar los perfiles desde el plano de simetría y con la superficie de la cara.



Gordon como se describe en (Bowyer et al, 2005), propuso un método basado en curvaturas, donde se extraen un conjunto de características que describen curvatura y mediciones de la cara. Cada cara se transforma en un punto en el espacio de características y se realiza una comparación de vecino cercano.

- Geometría Diferencial (Differential Geometry)

Se utiliza geometría diferencial para obtener características de la superficie que son invariantes ante transformaciones como rotación y traslación.

En el trabajo de (Gupta et al, 2007) con modelos 3D se localizan manualmente unos 25 puntos antropométricos y se toman distancias euclídeas y geodésicas entre los mismos. Estos puntos fueron seleccionados en base al trabajo realizado por Farkas, donde se observa que son los que presentan mayor variación entre los individuos con lo cual ayudan a realizar el reconocimiento de los mismos. Luego se clasifica el reconocimiento de una cara mediante la aplicación de LDA entre los modelos.

- Geometría basada en características faciales

Como en los inicios del reconocimiento facial 2D, se ubican varias características faciales como la punta de la nariz, ojos y boca y se obtienen distintas mediciones de las mismas como ángulos, distancias, longitudes.

- Distancias geodésicas

Según Berreti et al (2007), describe que el uso de las distancias para capturar la información es directamente motivado por la relevancia que la metrología tiene en la antropometría facial. Estas distancias geodésicas son utilizadas en varias aplicaciones gráficas, por ejemplo la interpolación sobre una superficie, cálculo de itinerarios de rutas mediante mediciones de distancias entre otras.



En general, la distancia geodésica es la longitud de una curva que corresponde al camino más corto entre dos puntos recorrido sobre una superficie que contiene ambos puntos.

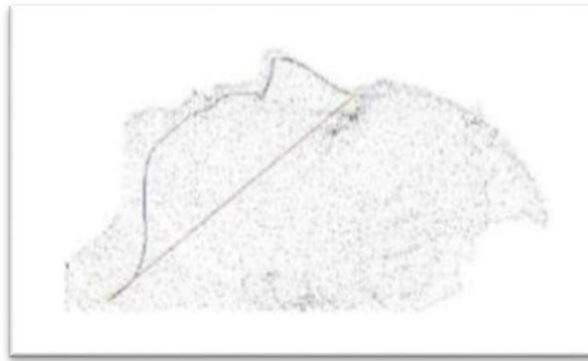


Ilustración 13 - Distancia geodésica azul vs distancia euclidiana (rojo)

Fuente: Pamela Pereyra (2014)

2.3.3. Algoritmos para comparación de modelos faciales

2.3.3.1. PCA (Principal Component Analysis)

Este método permite transformar un número de variables posiblemente correlacionadas en un pequeño número de variables incorrelacionadas denominado componentes principales, es decir es un algoritmo de reducción dimensional que permite encontrar los vectores que mejor representan la distribución de un grupo de imágenes. Este método está basado en la Transformada de Karhunen-Loeve, que consiste en la representación de un proceso estocástico, no periódico a través de una base de vectores, es decir permite representar una imagen de una cara usando una base que se ha conseguido a partir de muchas observaciones de diferentes caras.

La fórmula del análisis de este método se describe a continuación:



a. Se calcula el número total de vectores:

$$N = \sum_{clase=1}^c n_{clase}$$

b. Se calcula el vector medio de todos los vectores:

$$m = \frac{1}{N} \sum_{clase=1}^c \sum_{j=1}^n V_{clase}$$

c. Se calcula la matriz de covarianza, donde T significa transpuesto

$$c = \frac{1}{N} \sum_{i=1}^N (V_i - m)(V_i - m)^T$$

- d. Se diagonaliza la matriz de covarianza y se obtiene P auto vectores
- e. Se ordenan los P auto vectores, en función de su auto valor asociado (de mayor a menor). De esta forma el primero es el más importante y despreciar los últimos suelen suponer un error pequeño.
- f. Los auto vectores proporcionan los ejes de un nuevo sub espacio vectorial. Para representar los vectores dato en el nuevo sub espacio hay que proyectarlos, sobre los auto vectores:

$$V_j = a_j^T v_i, \quad j = 1, \dots, P \quad i = 1, \dots, N$$

Y se obtiene las coordenadas del vector en la nueva base.

$$v_{PCA} = \begin{bmatrix} V_1 \\ \dots \\ V_j \\ \dots \\ V_p \end{bmatrix}$$

Dónde:

- P = dimensión de los vectores



- $v_{PCA} = \begin{bmatrix} V_1 \\ \dots \\ V_j \\ \dots \\ V_p \end{bmatrix}$ = vector P-dimensional
- C= número de clases distintas:
- V_{clasej} = el vector j-ésimo de la clase, donde clase= 1,...c

2.3.3.2. LDA (Linear Discriminant Analysis)

LDA tiene como objetivo convertir un problema de alta dimensionalidad en uno de baja. El LDA proyecta las imágenes en un espacio vectorial de baja dimensionalidad de manera que la ratio entre la distancia entre clases y la distancia dentro de la clase se maximiza. De este modo se garantiza una máxima discriminación entre las clases (Gimeno, 2010).

A diferencia de la transformación anterior no busca la transformación eficiente de los datos, sino una transformación que separe lo máximo posible las diferentes clases entre sí. Matemáticamente se representa de la siguiente manera:

- Se calcula la media de los vectores de cada clase:

Clase =1,..c

$$\rightarrow m_{clase} = \frac{1}{n_{clase}} \sum_{j=1}^{n_{clase}} V_{clasej}$$

- Se calcula la matriz de dispersión interclase:

$$SW = \sum_{clase=1}^c \sum_{j=1}^{n_{clase}} (V_{clasej} - m_{clase})(V_{clasej} - m_{clase})^T$$



c. Se calcula la matriz de dispersión entre clase:

$$B = \sum_{\text{clase}=1}^c (m_{\text{clase}} - m)(m_{\text{clase}} - m)^T$$

c. Se calcula los auto vectores. A partir de aquí se sigue los mismos pasos (e) y (f) del método anterior.

2.3.3.4. Aplicaciones

En la tabla 3 se detallan algunas empresas que ofrecen aplicaciones de reconocimiento facial:

Tabla 3 Empresas que ofrecen aplicaciones de reconocimiento facial

Empresa	Aplicación
Visaje Technology http://www.l1id.com/pages/18	Reconocimiento facial. Identificación de individuos en checkpoints: reconocimiento de personas buscadas, entre otros objetivos
FaceKey Corp. http://www.facekey.com/	Reconocimiento facial para control de acceso
Cognitec Systems http://www.cognitecsystems.de/	Reconocimiento facial en video, control de frontera, identificación de sospechosos personas perdidas, control de acceso.
Passfaces from ID-arts http://www.id-arts.com/	Control de acceso complementario al existente o para su reemplazo.
ImageWare Software http://www.iwsinc.com/	Control de acceso mediante una combinación de varios métodos biométricos entre ellos reconocimiento facial
BioID sensor fusion http://www.bioid.com/	Reconocimiento biométrico mediante cara, iris, voz o combinación de las tres para accesos web o dispositivos móviles.
Biometric Systems, Inc. http://www.biometrica.com/	Dispone de un módulo de reconocimiento facial para inspeccionar similitud con sospechosos.
SpotIt for face composite http://spotit.fbk.eu/SpotIt.html	Permite buscar similitudes entre caras, crear identikit en base a las caras, entre otras opciones.



FaceSDK	La empresa de biometría Luxand tiene disponible un SDK de reconocimiento facial. El cual será utilizado en este trabajo.
---------	--

Fuente: Pamela Pereyra (2014)

2.4. Definición de la terminología

2.4.1. Estereoscopia: Es la capacidad del sistema visual del hombre de observar objetos en tres dimensiones. De esta forma somos capaces de crear dos imágenes con un ángulo ligeramente distinto y se las mostramos a cada ojo por separado, el cerebro podrá reconstruir la distancia y por lo tanto dicha sensación de profundidad. Recuperado del artículo “*Estereoscopia*” publicado en www.upf.edu/

2.4.2. Ambiente no controlado: Según Blázquez (2013), describe como un espacio de trabajo donde no se tiene control de ninguna variable del entorno. Por ejemplo, se puede presentar mucha luminosidad, o puede existir ausencia de luz en un espacio determinado de trabajo.

2.4.3. Ambiente controlado: Entorno cerrado donde parámetros como luz, humedad relativa y algunas veces presión parcial de los gases están completamente controlados. Recuperado de <http://www.profepa.gob.mx/>

2.4.4. Reconocimiento De Patrones: Según Julius T. Tou y Rafael C. González en su libro *Pattern Recognition Principles*, define al reconocimiento de patrones como la "categorización" de los datos de entrada en clases o patrones mediante la extracción de propiedades significativas que permiten discriminar entre las clases en estudio.



CAPÍTULO III: MARCO METODOLÓGICO

3.1. Tipo y Diseño de Investigación

3.1.1. Tipo de investigación

La presente investigación es del tipo Tecnológica

3.1.2. Diseño de Investigación

Su diseño Cuasi-Experimental.

3.2. Población y Muestra

3.2.1. Población

La población está compuesta por imágenes estereoscópicas en RGB, que cuenten con una resolución de 720 x 576 píxeles.

Para el presente trabajo se tiene una población indeterminada

3.2.2. Muestra

$$n = \frac{Z^2 \cdot p \cdot q}{e^2}$$

Dónde:

n: es el tamaño de la muestra;

Z²: nivel de confianza (1.96)².

p: probabilidad de que ocurra (0.5)

q: probabilidad de que no ocurra (0.5).

e²: error de estimación (0.05)².

$$n = \frac{(1.96)^2 (0.5)(0.5)}{(0.05)^2}$$



n = 348,16 imágenes.

3.3. Hipótesis

Mediante un sistema computacional que emplea algoritmos de modelos faciales podemos reconocer rostros en imágenes estereoscópicas.

3.4. Operacionalización

Tabla 4 Operacionalización de variable Dependiente

VARIABLE DEPENDIENTE	DIMENSIONES	INDICADORES	FÓRMULA	TÉCNICA E INSTRUMENTOS DE RECOLECCIÓN DE DATOS
Reconocimiento de rostros en imágenes estereoscópicas	Correctitud	Nivel de Eficacia	$\frac{(VP + VN)}{Total\ de\ imágenes}$	Observación
				Hoja de verificación

Fuente: Creación Propia

Tabla 5 Operacionalización de variable Independiente

VARIABLE INDEPENDIENTE	DIMENSIONES	INDICADORES	FÓRMULA	TÉCNICA E INSTRUMENTOS DE RECOLECCIÓN DE DATOS
Sistema computacional que emplea algoritmos de modelos faciales	Confiabilidad	Nivel de Sensibilidad	$\frac{VP}{VP + FN}$	Observación
				Hoja de verificación
		Nivel de Especificidad	$\frac{VN}{VN + FP}$	Observación
				Hoja de verificación

Fuente: Creación Propia



3.5. Métodos, técnicas e instrumentos de recolección de datos

3.5.1. Métodos de investigación

Los métodos adecuados para el desarrollo de la investigación son: La Observación, y captura de Imágenes.

a. Observación: Este método nos ayudara hacer un análisis mediante las pruebas que realicemos.

b. Captura de imágenes: Este método nos ayudará a obtener las imágenes de los rostros de las personas, con las cuales realizaremos las pruebas.

3.5.2. Técnicas de recolección de datos

Las técnicas empleadas en la investigación ha sido la siguiente: La observación, análisis documental.

a) Análisis documental: Consiste en extraer la información de los distintos documentos consultados, donde presentan teorías, fórmulas, técnicas, métodos para dar solución al problema planteado. Esto Servirá para determinar los métodos requeridos y así analizar los resultados obtenidos después de su implementación.

b) La observación: Porque se realizará un registro visual para captar la realidad sin distorsionar información sobre los acontecimientos de la situación en la que se encuentra el objeto de estudio, captando hechos relevantes para plasmarlo y tomarlos en cuenta en los resultados de la investigación a realizar.

3.5.3. Instrumentos de recolección de datos.

El instrumento que se usará en esta investigación será una hoja de verificación en la cual se llevará un registro de todas las incidencias ocurridas.



3.6. Procedimiento para la recolección de datos

Para la captura digital de los rostros se ha implementado un mecanismo basado en visión estereoscópica, compuesto por dos cámaras a una distancia de unos 5 cm en dirección horizontal generando una visión paralela, para luego con el procesamiento conseguir la percepción de profundidad que percibimos las personas.

Con ayuda del análisis documental hecho previamente escogeremos algoritmo para la clasificación del modelo facial, el algoritmo que aplicaremos en este caso será PCA, que es más eficiente para la extracción de características, aplicaremos este modelo a las imágenes obtenidas anteriormente y observaremos los resultados para posteriormente plasmarlos en una hoja de verificación, la hoja de verificación contendrá el registro de todos los verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos (ver en tabla 6) que obtengamos del reconocimiento en las imágenes estereoscópicas.

Tabla 6 Esquema de la hoja de verificación

	VP	FP	VN	FN
1	1	0	0	0
2	0	1	0	0
3	0	0	0	1
...
174	1	0	0	0
175	1	0	0	0
176	0	0	1	0
...				
346	1	0	0	0
347	0	0	0	1
348	0	0	0	1

Fuente: Creación Propia



3.7. Análisis Estadístico e Interpretación de los datos

La información obtenida mediante la observación, será plasmada en una hoja de hoja de verificación, la información de esta hoja será organizada en una matriz (ver tabla 7) .

Tabla 7 Matriz de Confusión

	Tiene Rostro	No tiene Rostro
Reconoce	VP	FP
No Reconoce	VN	FN

Fuente: Creación Propia

Dónde:

- **VP (Verdadero Positivo)** = Imágenes con rostro reconocidas.
- **VN (Verdadero Negativo)** = Imágenes con rostro no reconocidas.
- **FP (Falso Positivo)** = Imágenes sin rostro reconocidas.
- **FN (Falso Negativo)** = Imágenes sin rostro no reconocidas.

Para encontrar el valor de cada uno de nuestros indicadores usaremos las siguientes fórmulas:

$$\text{Nivel de Eficacia} = \frac{(VP + VN)}{\text{Total de imágenes}}$$

$$\text{Nivel de Sensibilidad} = \frac{VP}{VP + FN}$$

$$\text{Nivel de Especificidad} = \frac{VN}{VN + FP}$$



3.8. Criterios éticos

En el presente trabajo de investigación denominado: Reconocimiento facial mediante imágenes estereoscópicas, se desarrolló un sistema computacional empleando algoritmos de modelos faciales para el reconocimiento de rostros en imágenes estereoscópicas, se realiza siguiendo criterios éticos que todo investigador debe tener muy esclarecido durante cada etapa de desarrollo del proyecto. Entre los principios básicos seguidos en el desarrollo de nuestra propuesta de investigación tenemos:

Autonomía: Exige que las personas que estén en condiciones de deliberar acerca de sus metas personales se les respete su capacidad de autodeterminación.

Igualdad: Todos los seres humanos son iguales tienen los mismos derechos y obligaciones y forman un todo que podemos llamar familia humana.

Justicia: Que los casos considerados similares se traten de manera similar, y que los casos considerados diferentes se trate de tal forma que se reconozca la diferencia.

3.9. Criterios de rigor científico

En el presente trabajo de investigación se realiza siguiendo los juicios científicos establecidos, estos permiten garantizar la calidad de la propuesta de investigación.

Así, el presente trabajo se desarrolla bajo una estructura metodológica establecida por el centro de estudios. Realización apropiada del muestreo de datos, los cuales son al azar para ser totalmente imparcial en el recojo de datos.



CAPÍTULO IV: ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS

4.1. Resultados en tablas y gráficos.

Para realizar el análisis de los resultados de las pruebas de identificación se utiliza una matriz de confusión (ver tabla 8), presentada en las secciones anteriores, de manera a obtener la sensibilidad, la especificidad y la eficacia del sistema.

Tabla 8 Matriz de Confusión de Resultados obtenidos

	Tiene Rostro	No tiene Rostro
Reconoce	163	11
No Reconoce	166	8

Fuente: creación propia

A partir de dicha matriz son calculados el número total de verdaderos positivos, falsos negativos, falsos positivos y verdaderos negativos, los cuales son necesarios para realizar la evaluación de los indicadores, dichos resultados se encuentran en la tabla “**Ficha de Verificación**” Anexo 1.

Indicador 1: Nivel de Eficacia

Para poder obtener los datos referentes al indicador se calculó con la siguiente fórmula:

$$\text{Nivel de eficacia} = \frac{(VP + VN) * 100\%}{\text{Total de imágenes}}$$

Entonces reemplazando con los datos de la tabla 8

$$\text{Nivel de eficacia} = \frac{(163 + 166) * 100\%}{348}$$

Nivel de eficacia = 94.54 %



Indicador 2: Nivel de Sensibilidad

Para poder obtener los datos referentes al indicador se calculó con la siguiente fórmula:

$$\text{Nivel de sensibilidad} = \frac{VP * 100\%}{VP + FN}$$

Entonces reemplazando con los datos de la tabla 8

$$\text{Nivel de sensibilidad} = \frac{163 * 100\%}{163 + 8}$$

Nivel de sensibilidad = 95.32 %

Indicador 3: Nivel de Especificidad

Para poder obtener los datos referentes al indicador se calculó con la siguiente fórmula:

$$\text{Nivel de Especificidad} = \frac{VN * 100\%}{VN + FP}$$

Entonces reemplazando con los datos de la tabla 8

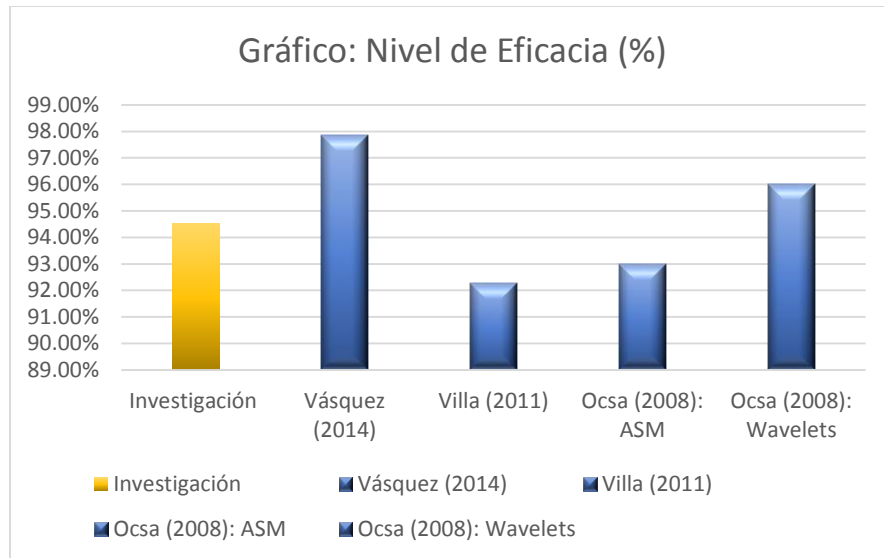
$$\text{Nivel de Especificidad} = \frac{166 * 100\%}{166 + 11}$$

Nivel de Especificidad = 93.78 %



4.2. Discusión de resultados.

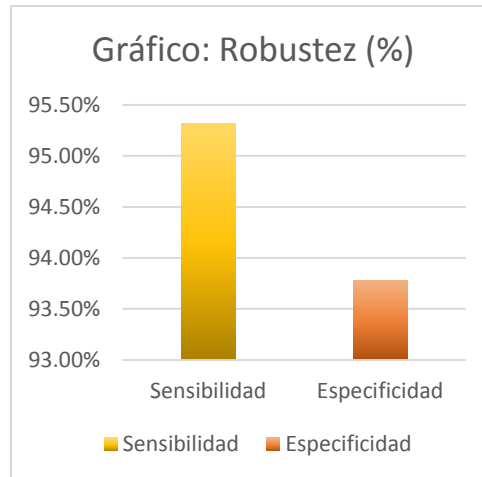
De los resultados anteriores podemos observar:



Fuente: Creación Propia

El nivel de eficacia obtenido en la presente investigación es de 94.54%, interpretándose como la capacidad que el sistema computacional implementado posee de reconocer el rostro de una persona, es una tasa suficientemente aceptable considerando que en otros trabajos como Vásquez (2014) obtienen una eficacia de 97.87% en ambiente controlado con equipos especiales para la captura de imágenes estereoscópicas, en el trabajo de investigación denominado “Sistema de Reconocimiento de Rostros” de Villa (2011), llegan a obtener una eficacia del 92.26% habiendo utilizado algoritmos eigenfaces” o PCA . Y en el trabajo “Reconocimiento de Rostros mediante Puntos Característicos Locales”, propuesto por Ocsa (2008) llegan a obtener un grado de reconocimiento del 93% con algoritmos Active Shape Model, y 96% con wavelets.





Fuente: Creación Propia

El Indicador de sensibilidad nos plasma la capacidad de nuestro sistema computacional al momento de reconocer como verdadero a un rostro que se encuentra en nuestra base de datos, esta característica es importante ya que en nuestro trabajo se ha obtenido un nivel de 95.32%, plasmándonos que la probabilidad de analizar correctamente a un rostro es elevada.

El indicador de especificidad nos plasma la capacidad de nuestro sistema computacional de detectar a una imagen que está fuera de nuestra base de datos, con la finalidad de no generar errores en el análisis y detección de un rostro, este indicador está en un 93.78%, siendo al igual que el indicador anterior dos características que brindan una robustez de nuestra investigación ante posibles casos no esperados.



CAPÍTULO V: DESARROLLO DE LA PROPUESTA DE INVESTIGACIÓN

5.1. DISEÑO DEL SISTEMA

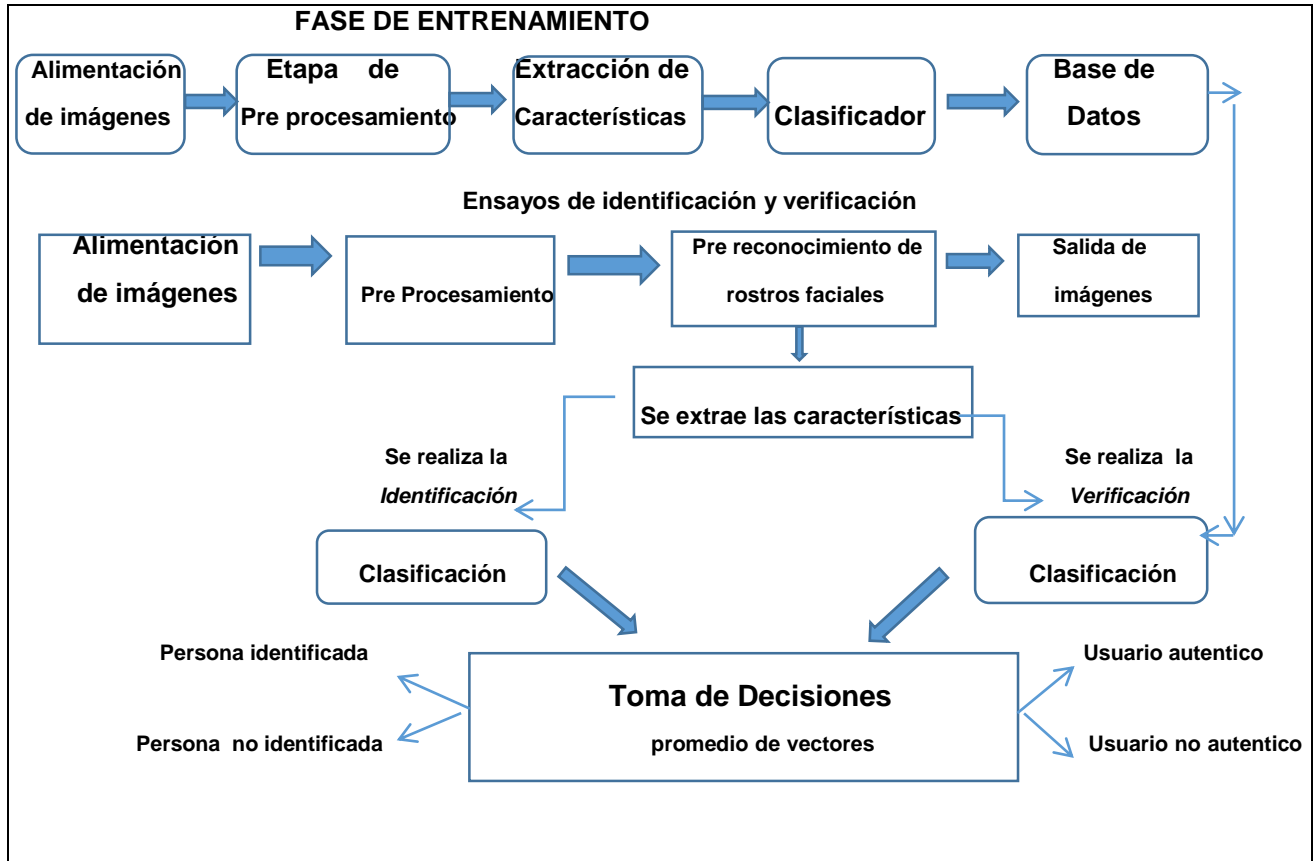


Ilustración 14 - Diagrama de bloques del sistema de reconocimiento de rostros. Etapa de pre-reconocimiento de rostros reales.

Fuente: Creación Propia

Etapa de pre procesamiento

En esta etapa inicial se elimina el ruido aditivo que puede llegar a contaminar la imagen capturada, se segmenta la imagen y eventualmente, dependiendo del tamaño de la imagen capturada, se realiza una reducción de la dimensión de la imagen de entrada mediante una operación de filtrado pasa bajas, para facilitar el análisis de la misma. En caso de ser necesario, en esta etapa se podría también llevar a cabo la restauración de la imagen de entrada empleando esquemas de



igualación del histograma para mejorar su calidad cuando ésta ha sido alterada debido a cambios en la iluminación (Olivares-Mercado y col., 2007).

La etapa de pre-reconocimiento, mostrada en la ilustración 15, intenta evitar la suplantación de identidad mediante el uso de las dos imágenes estereoscópicas presentes a la entrada del sistema. A fin de determinar si las imágenes estereoscópicas que se presentan a la entrada del sistema corresponden a un rostro real, tridimensional o una imagen en 2 dimensiones obtenida a partir de una fotografía, de alta resolución, perteneciente a una de las personas registradas en el sistema, se calcula la disparidad existente entre la imagen derecha e izquierda del rostro de entrada obtenidas usando una cámara estero. La idea principal detrás de este procedimiento es que la diferencia existente entre la imagen izquierda y derecha de un rostro real, ilustración 16, al ser un volumen, es mayor que la diferencia obtenida cuando las imágenes de entrada corresponden de un objeto plano tal como se muestra en la ilustración 17.

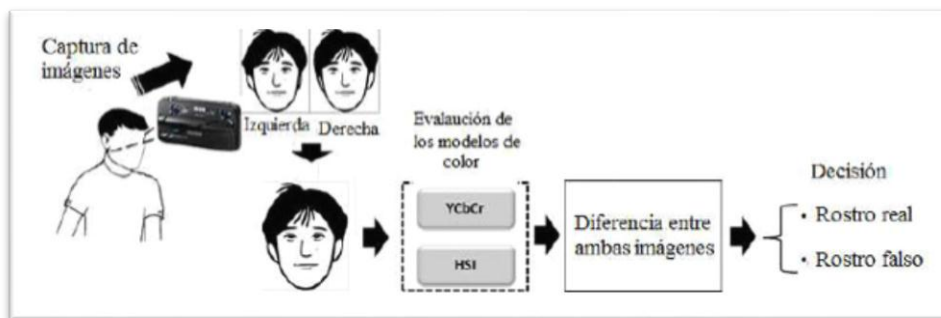


Ilustración 15 - Diagrama de la etapa de pre-reconocimiento

Fuente: Información tecnológica Vol. 25 N° 62014



Ilustración 16 - Componentes izquierda y derecha de una imagen estereoscópica

Fuente: Información tecnológica – Vol. 25 N° 62014

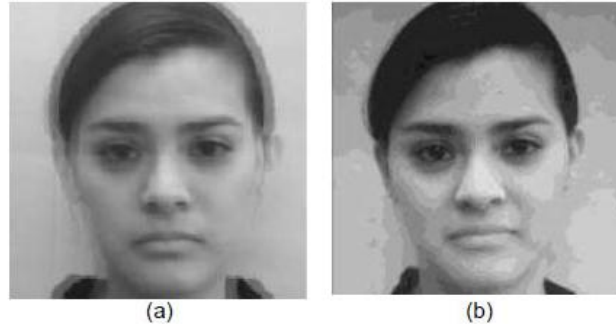


Ilustración 17 - Traslape existente entre las imágenes tomadas de (a) Un rostro real (b) De una fotografía

Fuente: Información tecnológica Vol. 25 N° 62014

A continuación, se detalla el procedimiento realizado en esta etapa. Cada una de las imágenes estereoscópicas de entrada se pre-procesan como se mencionó anteriormente, se realiza la conversión de las mismas a los modelos de color YCbCr o HSI (Trussell y col., 2005; Marcial-Basilio y col., 2012) separándose en las componentes de intensidad o luminancia y cromas, que comprenden los modelos de color, considerándose únicamente las componentes de intensidad (I), cuando se emplea el formato HSI, o las componentes de luminancia (Y) cuando se emplea el formato YCbCr. Una vez adquiridas las componentes ya sea I o Y, de imágenes de entrada las cuales son imágenes en escala de grises, se obtiene la diferencia existente entre las imágenes estereoscópica izquierda y derecha del rostro. Si las imágenes derecha e izquierda corresponden a un volumen, ambas imágenes presentan diferencias significativas entre sí y no se empalman perfectamente, como se muestra en la ilustración 17 (a), dando como resultado una diferencia significativa en el valor de un número considerable de píxeles. Esta diferencia es más evidente en los extremos de imagen del rostro como se puede



ver en la ilustración 17 (a). Por otro lado cuando ambas imágenes corresponden a un objeto en 2-dimensiones, el empalme es casi perfecto, como se muestra en la ilustración 17 (b). Tomando en cuenta lo anterior, para determinar si se trata de una imagen obtenida de un rostro real, se calcula el promedio de la diferencia absoluta entre la luminancia o intensidad de la imagen izquierda $I_i(m,n)$ y la imagen derecha $I_d(m,n)$, pixel a pixel, esto es:

$$D_p = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |I_i(m,n) - I_d(m,n)|. \quad (1)$$

Aquí si D_p es mayor que un umbral (Th) dado de antemano, se considera que la imagen corresponde a un rostro real, en cuyo caso ambas imágenes se insertan en la etapa de extracción de características (García y col., 2013). Por otro lado, si D_p es menor que el umbral dado de antemano, el sistema determina que la imagen corresponde a un rostro falso, concluyendo en este momento el proceso de reconocimiento o verificación de identidad. Esta decisión se basa en el hecho de que en un par de imágenes estereoscópica tomadas directamente de un rostro real, la diferencia entre los valores de Luminancia o Intensidad de ambas imágenes es alto ya que existe diferencia relativamente grande entre el valor de los píxeles de ambas imágenes ubicados en la misma posición debido a la disparidad que presentan. Por otro lado, cuando las imágenes son tomadas de fotografía en 2-dimensiones la diferencia entre la Luminancia o Intensidad existente entre las imágenes derecha e izquierda es bajo, ya que no existe una diferencia apreciable entre el valor de los píxeles de ambas imágenes localizados en la misma posición. Con el fin de determinar un valor adecuado para el umbral se realizó una evaluación experimental en la cual se capturaron imágenes estereoscópicas de personas diferentes. Los resultados experimentales obtenidos



muestran que una posible selección de los umbrales, Th , apropiados serían: $Th=0.05$ cuando se emplea la intensidad y $Th=0.025$ cuando se emplea la Luminosidad. (García y col., 2013). La ilustración 18 muestra las distribuciones de los valores de D_p , junto con los valores de Th .

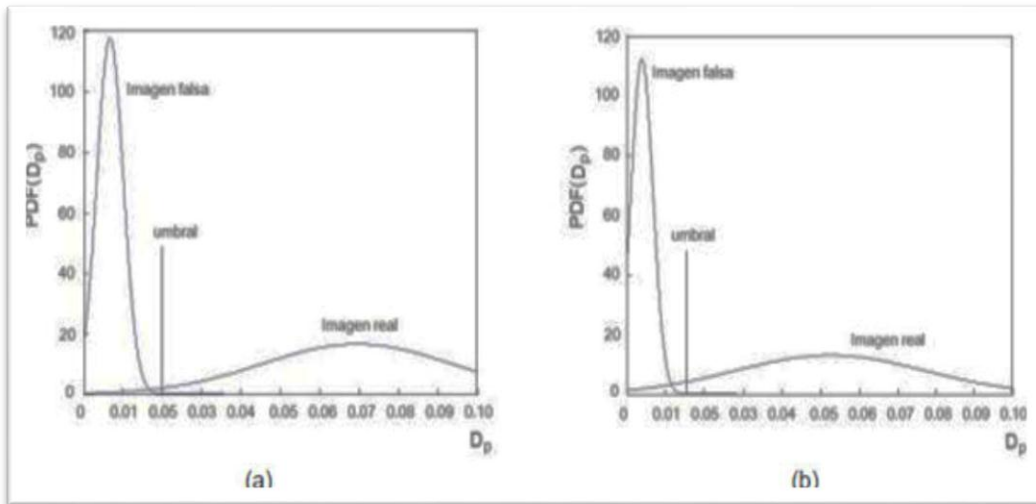


Ilustración 18 - Valores obtenidos de D_p cuando se emplea la Luminosidad (Y) y la Intensidad (I); (a) empleando el formato YCbCr y (b) empleando el formato HSI

Fuente: Información tecnológica Vol. 25 N° 62014

Extracción de características faciales

Una vez que la etapa de pre-reconocimiento determina que la imagen bajo análisis corresponde a un rostro real, se procede a la extracción de características usando ya sea: el Análisis de Componentes Principales (PCA), las funciones de Gabor (GF) o la DWT los cuales han sido empleados con éxito en diversos esquemas de reconocimiento de rostros. (Corcoran, 2011).

Extracción de características usando Análisis de Componentes Principales (PCA)

Uno de los métodos más usados para reducir la dimensión de un problema de reconocimiento de patrones es el PCA (Aguilar y col., 2009; Faruque y Hassan,



2009; Kshirsagar y col., 2011; Turk, y Pentland, 1991), el cual permite representar las imágenes bajo análisis, en una forma extremadamente compacta en comparación con el tamaño de las imágenes mismas. Con el fin de establecer las ecuaciones principales del PCA para llevar a cabo la caracterización de las imágenes de los rostros, considere un conjunto de M imágenes en niveles de gris, cada una de $L=N_1 \times N_2$ pixeles, las cuales se arreglan en una matriz de dimensión $L \times M$. Seguidamente, con el fin de encontrar los vectores que mejor representan la distribución de imágenes del rostro bajo análisis, a cada imagen de entrenamiento, $\Gamma_i, i=1,2,\dots,M$, se le resta la imagen Ψ la cual se obtiene promediando las M imágenes de entrenamiento. De manera que se obtiene el conjunto de imágenes dado por (Aguilar y col., 2009; Faruqe y Hassan, 2009).

$$\Phi_i = \Gamma_i - \Psi \quad i=1,2,\dots,M. \quad (2)$$

Para reducir la dimensión de los vectores involucrados se puede emplear el método de descomposición en valores singulares (SVD) el cual establece que si U_k y λ_k son el k -ésimo vector propio y k -ésimo valor propio, respectivamente, de la matriz de covarianza C , dada por

$$C = \frac{1}{M} \sum_{k=1}^M \Phi_k \Phi_k^T = AA^T \quad (3)$$

Entonces

$$U_k^T C U_k = \lambda_k \quad (4)$$

Multiplicando ambos lados por U_k^T y usando la propiedad de ortonormalidad de los vectores propios, de la ecuación (3) obtenemos (Faruqe y Hassan, 2009; Kshirsagar y Col., 2011).



$$AA^T U_k = \lambda_k U_k, \quad k=1,2,3,\dots,M \quad (5)$$

Seguidamente, multiplicando en ambos lados por la izquierda por A^T se obtiene:

$$A^T A(A^T U_k) = \lambda_k (A^T U_k), \quad k=1,2,3,\dots,M \quad (6)$$

$$A^T A Y_k = \lambda_k Y_k, \quad k=1,2,3,\dots,M \quad (7)$$

Seguidamente, de la ecuación (5) y (7) se tiene que si U_k es el k-ésimo vector propio de AA^T entonces $Y_k=A^T U_k$ es el k-ésimo vector propio de $A^T A$, estando ambos relacionados a través de la expresión

$$U_k = \frac{A Y_k}{\lambda_k}, \quad k=1,2,3,\dots,M \quad (8)$$

La importancia de esta reducción radica en el hecho de que mientras AA^T puede llegar a tener hasta $L=N_1 \times N_2$ valores propios, donde L es el número de píxeles de la imagen bajo análisis, $A^T A$ puede tener hasta M valores propios, donde M es el número de imágenes de entrenamiento, correspondiendo los M valores propios de $A^T A$ a los M valores propios, mayores, de AA^T . Dado que en general $M \ll L$, la reducción obtenida es significativa. Una vez que los M vectores propios han sido calculados, la imagen bajo análisis es proyectada en el espacio de rostro mediante una simple operación dada por (Turk y Pentland, 1991; Faruge y Hassan, 2009):

$$w_k = U_k^T (\Gamma - \Psi), \quad k=1,2,\dots,M. \quad (9)$$

Finalmente, los pesos que forman el vector $\Omega=[w_1, w_2, w_3, \dots, w_M]$



describen la contribución de cada componente principal en el espacio de la imagen del rostro de entrada.

Extracción de características usando Análisis de Discriminantes Lineales (LDA)

Se busca maximizar la separación en la proyección del grupo conformado por las imágenes diferentes de los mismos sujetos y el grupo conformado por las imágenes de diferentes sujetos.

En esta técnica consideramos las imágenes correspondientes a un mismo sujeto como una clase, de tal forma que se tendrán tantas clases como sujetos haya. Con estas clases definidas, se procede a obtener dos matrices de dispersión diferentes, la matriz de dispersión interna de las clases (S_W) y la matriz de dispersión entre clases (S_B), que se calculan de acuerdo a las ecuaciones (10) y (11) respectivamente.

$$S_W = \sum_{i=1}^P \sum_{k=1}^{N_i} (I_k^i - \mu_i)(I_k^i - \mu_i)^T \quad (10)$$

$$S_B = \sum_{i=1}^P (\mu_i - \mu)(\mu_i - \mu)^T \quad (11)$$

Se trata de obtener un vector de proyección w , que haga que la razón entre la dispersión intra-clase y la dispersión inter-clase sea máxima. Habrá que maximizar la siguiente función objetivo:

$$J(w) = \frac{w^t * S_b w}{w^t * S_w w}$$



El vector w que maximiza esta función será aquel que cumpla la siguiente ecuación:

$$S_b \cdot w = \lambda \cdot S_w \cdot w$$

Si la matriz S_w es no singular (tiene inversa) se tiene un problema de valores propios para la matriz $S_w^{-1} \cdot S_b$, que puede sustituirse en la función objetivo:

$$S_w^{-1} \cdot S_b \cdot w = \lambda \cdot w \rightarrow J(w) = \frac{w^t \cdot S_b w}{w^t \cdot S_w w} = \lambda_k \frac{w_k^t \cdot S_b w_k}{w_k^t \cdot S_w w_k}$$

Con $k=1, \dots, d$ (siendo d la dimensión de las imágenes).

En la expresión anterior se puede ver que el vector propio que maximiza la función es aquel que tenga un mayor valor propio (mayor λ_k).

Etapas de clasificación.

Después de obtener los vectores característicos en la etapa anterior, se lleva a cabo la identificación o verificación de la identidad de la persona bajo análisis, empleando una Máquina de Soporte Vectorial (SVM) (Yu y col., 2011; Chang y Lin, 2011) la cual, durante el entrenamiento, es responsable de generar un modelo para cada persona. Dado que el sistema propuesto puede ser empleado para verificar la identidad de una persona o para identificar a la misma, la SVM puede ser entrenada de dos formas distintas. Si se requiere que el sistema verifique la identidad de una persona, la SVM es entrenada en forma tal que su salida se aproxime a uno cuando la persona bajo análisis es quien él o ella afirman ser, mientras su salida debe aproximarse a cero en caso contrario. Este trabajo se lleva a cabo usando solamente un modelo de la base de datos junto con un umbral dado. Por otro lado, cuando la operación a realizar es la



identificación, la SVM es entrenada en forma tal que su valor de salida mayor corresponde al modelo de la persona más probable entre los que se encuentran en la base de datos, usando todos los modelos contenidos en ésta. En ambos casos, para entrenar la SVM se empleó el programa *LIBSVM library* escrito en matlab(Chang y Lin, 2011) .Una vez obtenidos los modelos de las personas durante el entrenamiento, estos son almacenados en la base de datos para su uso durante los procesos ya sea la identificación o verificación.

Por su parte para llevar a cabo ya sea la verificación o identificación, los modelos almacenados en la base de datos se cargan en la SVM, obteniéndose el vector de salida de la misma cuando la entrada es el vector característico de la imagen bajo análisis. En el primer caso, el vector de salida que consiste de un solo elemento, se compara con un umbral para determinar si el vector de entrada corresponde o no, a la persona que dice ser. Por su parte, cuando el sistema opera como identificador, se obtiene el índice del valor máximo del vector de salida, el cual corresponderá a la persona identificada.

Implementación de distintos niveles de fusión.

Una vez que se lleva a cabo la caracterización de las imágenes derecha e izquierda de la persona bajo análisis, se requiere determinar la mejor manera de combinar o fusionar la información extraída de ellas, para lo cual se analizaron tres niveles diferentes para fusionar esta información. Los niveles de fusión analizados, son: fusión a nivel sensor, fusión a nivel características y fusión a nivel decisión. Cada nivel de fusión puede considerarse como un sistema constituido por los diferentes módulos que componen el sistema biométrico, descritos anteriormente. A continuación, se describe cada uno de estos niveles de fusión.



Fusión a nivel sensor

La fusión a nivel sensor se lleva a cabo después de que son obtenidas las imágenes de entrada, y antes de proceder a la caracterización de las mismas, mediante la combinación de la mitad izquierda de la imagen izquierda del rostro con la mitad derecha de la imagen del mismo lado; complementando de esta manera la información proporcionada por ambas imágenes estereoscópicas.

Con esta finalidad, primeramente las imágenes son sometidas a la etapa de pre procesamiento para la segmentación del rostro, seguidamente se verifica que estas correspondan a un rostro real por medio de la etapa de pre-reconocimiento.

Seguidamente se segmentan las partes derecha e izquierda de ambas imágenes estereoscópica del rostro bajo análisis, llevándose a cabo finalmente la concatenación de la parte derecha de una con la izquierda de la otra, como se mencionó anteriormente, formándose así una nueva imagen del rostro conteniendo la información más relevante de cada imagen. Este proceso se muestra en la ilustración 19.



Ilustración 19 - Fusión de las imágenes izquierda y derecha

Fuente: Información tecnológica Vol. 25 N° 62014



Una vez obtenida la imagen resultante se procede a insertarla en las etapas de extracción de características, clasificación y toma de decisión, mostrados en la ilustración 14, para su reconocimiento o verificación de identidad.

Fusión a nivel de características

En el procedimiento de fusión a nivel de características, la fusión de los datos se efectúa después de llevar a cabo la caracterización de las imágenes estereoscópica bajo análisis (García y col., 2013). A diferencia del nivel de fusión anterior, en este se caracteriza el par de imágenes estereoscópicas, imagen izquierda e imagen derecha, por separado. Con esta finalidad, inicialmente introducen ambas imágenes a la etapa de pre procesamiento para su segmentación, antes de pasarlas a la etapa de pre-reconocimiento para determinar si corresponden o no a un rostro real en 3D. En caso afirmativo, ambas imágenes se insertan en la etapa de extracción de características en la cual se obtienen los vectores característicos de cada imagen. Una vez caracterizada cada una de las imágenes se procede a la fusión de los datos obtenidos, la cual se realizó en dos formas distintas: La primera mediante el cálculo de la media aritmética, elemento por elemento, de los vectores característicos de ambas imágenes, dando como resultado un vector característico cuya longitud es igual a cada uno de los vectores característicos individuales. La segunda opción consiste en la concatenación de ambos vectores característicos, lo que da como resultado un vector característico cuya dimensión es el doble de la dimensión de cada uno de los vectores característicos de cada una de las imágenes. La realización de cada uno de estos esquemas de fusión se muestra en la ilustración 20.



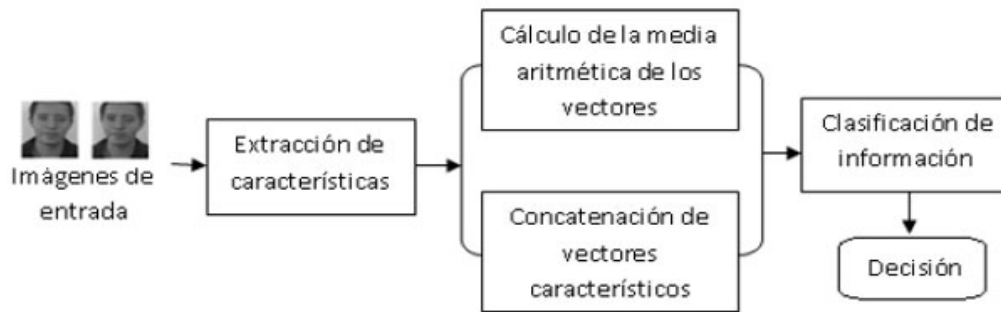


Ilustración 20 - Diagrama a bloques del nivel de fusión a nivel características

Fuente: Información tecnológica Vol. 25 N° 62014

Fusión a nivel decisión

El último mecanismo de fusión es a partir del nivel de decisión, donde cada una de las imágenes que componen la imagen estereoscópica es procesada de manera independiente, mediante las etapas descritas con anterioridad. Inicialmente pasan a la etapa de pre procesamiento, pasando seguidamente a la etapa de pre-reconocimiento para determinar si corresponden a un rostro real. Seguidamente se procede a la extracción de los vectores característicos de cada una de las imágenes, los cuales se insertan separadamente en dos etapas de identificación o verificación de identidad, independientes entre sí, en donde cada una de las cuales toma una decisión acerca de la identidad de la persona bajo análisis. Finalmente se aplica la fusión de los dos resultados obtenidos. Aquí el sistema determina que la verificación de identidad es positiva solamente si ambas etapas determinan que la persona bajo análisis es quien dice ser. Por otro lado, si se requiere que el sistema realice una operación de identificación, se promedian elemento a elemento los vectores de salida de ambas etapas de clasificación y se considera como la persona identificada aquella que corresponde al índice del vector de salida resultante cuyo valor promedio es máximo.



5.2. DESARROLLO DEL SOFTWARE

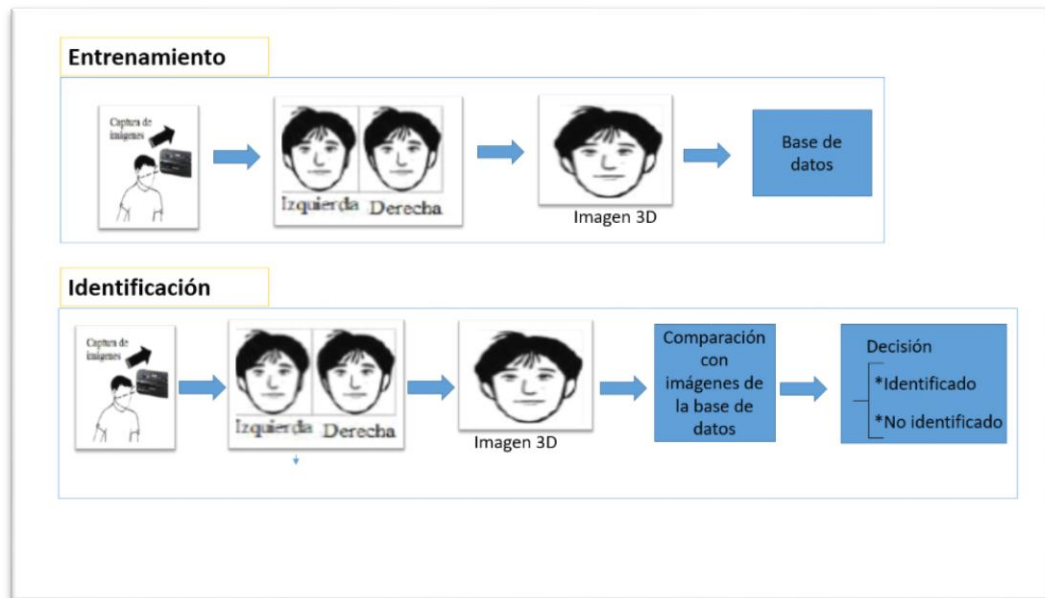


Ilustración 21 Diagrama a bloques del sistema de reconocimiento de imágenes estereoscópicas (fase de entrenamiento y fase de identificación).

Fuente: Creación Propia

5.2.1 Presentación Inicial de la Aplicación : interfaz principal de la aplicación.

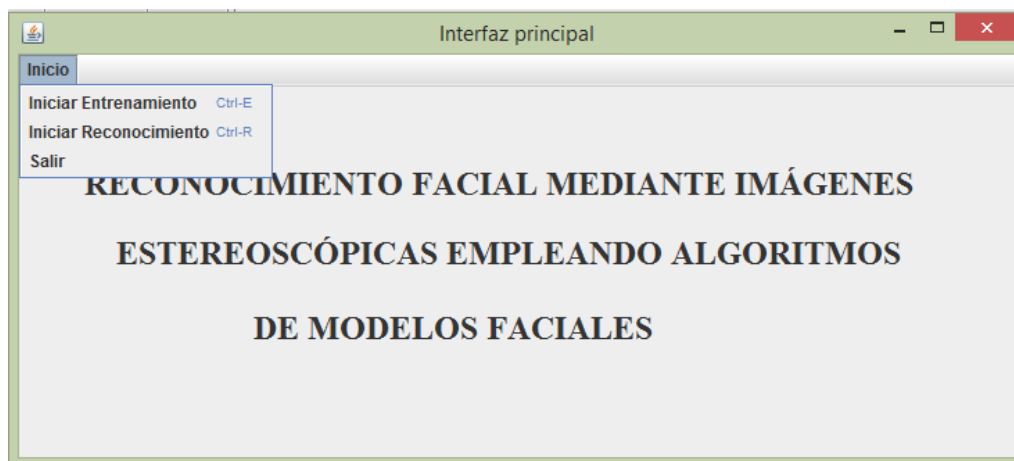


Ilustración 22 Interfaz principal del sistema

Fuente: Creación Propia

5.2.1 Fase entrenamiento



- **Iniciar entrenamiento.-** Esta interfaz nos muestra como al dar clic en iniciar entrenamiento, se activan las dos cámaras: izquierda y derecha lista para capturar la imagen.



Ilustración 23 Interfaz de entrenamiento cuando la cámara izquierda y cámara derecha iniciadas correctamente

Fuente: Creación Propia

Imagen 3D capturada.- En esta interfaz nos muestra que al presionar la tecla espacio se obtienen una imagen de la cámara izquierda a la cual se le hace un filtro rojo y una imagen derecha a la cual se le hace un filtro verde-azul, para luego sumar estas dos imágenes filtradas y así obtener una imagen estereoscópica. Este procesamiento se puede apreciar en los anexos 2.4, 2.5, 2.6 donde se muestran los códigos para poder obtener una imagen estereoscópica.





Ilustración 24 Imagen estereoscópica obtenida del sistema

Fuente: Creación Propia

Almacenamiento en la Base de datos.- Una vez obtenida la imagen estereoscópica, se procede a guardar en nuestra base de datos, por lo que al momento de dar clic en el botón guardar nos aparecerá una ventana indicando que debemos ingresar un nombre y un apellido (estos datos serán identificadores de la imagen) y así posteriormente guardar la imagen estereoscópica en nuestra base de datos la imagen con su respectivo identificador.





Ilustración 25 Ilustración 23 Almacenando imagen la base de datos

Fuente: Creación Propia

5.2.1 Fase Reconocimiento

Inicialización de la interfaz de reconocimiento.- Esta interfaz nos muestra como al dar clic en iniciar reconocimiento, se activan las dos cámaras: izquierda y derecha lista para capturar la imagen e iniciar el reconocimiento.



Ilustración 26 Interfaz de reconocimiento cuando la cámara izquierda y cámara derecha iniciadas correctamente

Fuente: Creación Propia



Obtención de la imagen.- Extraemos las dos imágenes de las dos cámaras, y al igual que en la fase de entrenamiento de la suma de ambas obtenemos una imagen estereoscópica como muestra el anexo 2.6, procedemos aplicar filtro de limpieza de la imagen para realizar la posterior detección del rostro.

Detección del rostro.- Para la detección del rostro utilizamos la ayuda de la libreta Opencv y Haar-cascade como podemos ver en el anexo 2.1, que al enviarle la imagen hará una comparación de patrones de toda la imagen para detectar el rostro, una vez que detecta el rostro esta nos devuelve las coordenadas del rostro a detectar o caso contrario unas coordenadas negativas. Posterior a la detección se dibuja un rectángulo con las coordenadas obtenidas caso el reconocimiento haya sido exitoso, caso contrario mostrara un mensaje informando que no se detectó el rostro.

Segmentación de la Imagen.- Luego de la detección y con las coordenadas dadas procedemos hacer la segmentación o recorte de la imagen (anexo 2.2), para quedarnos solo con la imagen del rostro. Esto nos ayudara aprovechar más cada una de las características del rostro.





Ilustración 27 Imagen segmentada

Fuente: Creación Propia

Carga de las imágenes de la base de datos y entrenamiento.- De nuestra base de datos obtenemos todas las imágenes y procedemos hacer el pre procesamiento, y extraemos los nombres de las imágenes y sus respectivos índices (anexo 2.3). Una vez cargadas todas las imágenes con ayuda de la librería Open cv y la clase Face Recognition procedemos a realizar el entrenamiento donde extrae todas las características de las imágenes contenidas en nuestra base de datos.

-Muestra el resultado.- Con el índice obtenido (anexo 2.3), Realizamos la comparación con los nombres obtenidos de las imágenes de nuestra base de datos, y al encontrar una coincidencia muestra en pantalla que se reconoció y el nombre de la persona reconocida.



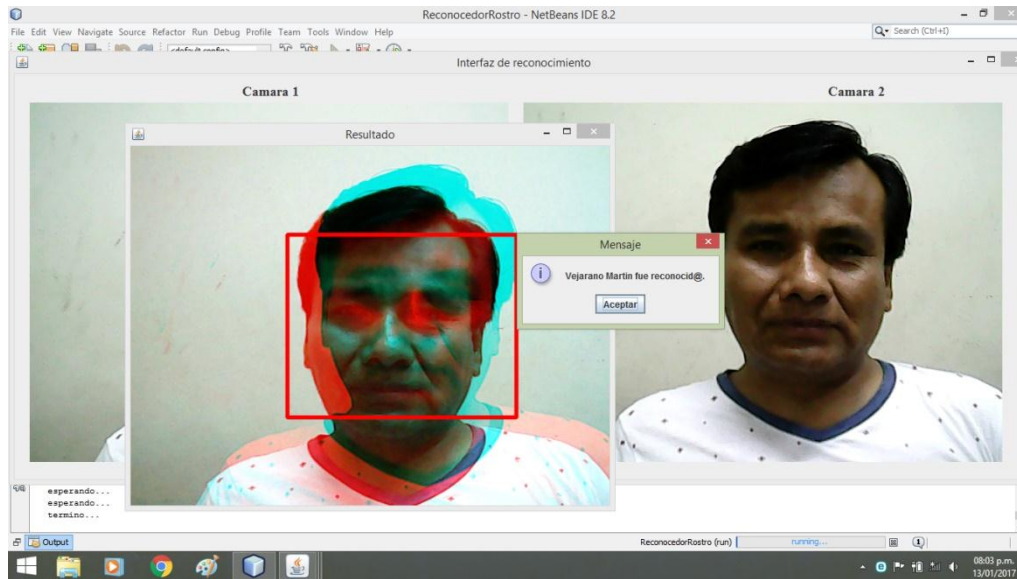


Ilustración 28 resultado si la persona fue identificada o no identificada

Fuente: Creación Propia

5.3. EVALUACIÓN DEL SISTEMA

La base construida consiste en imágenes de 40 personas diferentes con 10 imágenes estereoscópicas por cada persona, de las cuales 6 imágenes de cada persona se usaron para el entrenamiento del sistema y 4 para la realización de pruebas de identificación y verificación. Así, el número total de imágenes contenidas en esta base de datos es de 400 imágenes de las cuales solo se utilizó 348 imágenes, capturadas en un ambiente con iluminación y fondo controlado. La ilustración 21 muestra algunas de las imágenes que conforman esta base de datos.





Ilustración 29 - Ejemplos de imágenes contenidas en la base de datos

Fuente: Creación Propia

Se empleó el método de PCA para caracterizar las imágenes de entrada, la dimensión del vector característico depende del número de imágenes de entrenamiento. Por su parte para llevar a cabo la evaluación de la etapa de pre-reconocimiento usada para determinar si el rostro es real o falso, se construyó una base de datos con 50 imágenes, 35 fotografías tomadas directamente del rostro, correspondientes a la base de datos mencionada anteriormente y 15 fotografías tomadas de otras fotografías de las mismas personas.

La Tabla 9 muestra el comportamiento del sistema cuando se efectúa el proceso de identificación usando los 3 métodos de fusión descritos anteriormente. En todos los casos, los vectores característicos son obtenidos mediante el PCA, así como el LDA. Los resultados obtenidos muestran que los 2 esquemas se comportan de manera similar, difiriendo en menos de un punto porcentual entre ellos. Es importante hacer notar que en todos los casos los resultados obtenidos, cuando el sistema se evalúa usando solo una componente de las imágenes estereoscópicas, son bastante cercanos a aquellos obtenidos en los cuales se emplean ambas imágenes.



Tabla 9 Porcentaje de identificación

Niveles de fusión		LDA	PCA
Fusión a nivel sensor		99.22	99.50
Fusión a nivel características	Media aritmética	99.59	99.76
	Concatenación	99.78	99.82
Fusión a nivel decisión		99.40	99.50
Solo parte izquierda		99.46	99.51
Solo parte derecha		99.50	99.52

Fuente: Creación Propia

La Tabla 10 muestran los resultados obtenidos cuando se requiere que el sistema verifique la identidad de las personas cuyas imágenes reales se presentan a su entrada. En la Tabla 10 muestra el funcionamiento del sistema propuesto cuando se emplea el PCA. En la Tabla 11 se muestran los resultados obtenidos cuando se emplea la LDA. En los dos casos se presentan, además del porcentaje de verificación, el porcentaje de falsa aceptación (FAR) y el porcentaje de falso rechazo (FRR) proporcionado por el sistema propuesto.

Tabla 10 Porcentajes de verificación usando el método de PCA

Niveles de fusión		PCA		
		%FAR	%FFR	%Verificación
Fusión a nivel sensor		2.429	0.786	98.20
Fusión a nivel características	Media aritmética	2.357	0.714	97.76
	concatenación	2.000	0.786	96.820
Fusión a nivel decisión		2.5	0.714	96.071
Solo parte izquierda		2.429	0.714	97.143
Solo parte derecha		2.142	0.643	97.56

Fuente: Creación Propia



Tabla 11 Porcentajes de verificación usando el método de LDA

Niveles de fusión		LDA		
		%FAR	%FFR	%Verificación
Fusión a nivel sensor		0.98	0.78	93.20
Fusión a nivel características	Media aritmética	0.60	0.31	96.76
	concatenación	0.82	0.27	94.82
Fusión a nivel decisión		0.69	0.29	96.50
Solo parte izquierda		0.74	0.17	92.10
Solo parte derecha		0.65	0.23	95.5

Fuente: Creación Propia



CAPÍTULO VI: CONCLUSIONES Y RECOMENDACIONES

6.1. Conclusiones

Luego de culminada la investigación podemos concluir en:

- Se logró reconocer rostros en imágenes estereoscópicas mediante un sistema computacional que emplea algoritmos de modelos faciales (PCA).
- Para desarrollar el presente trabajo se realizó una investigación documental, la cual nos permitió identificar el estado actual del reconocimiento de rostros en imágenes estereoscópicas, así como los avances que se han venido experimentando desde la adquisición hasta el reconocimiento de rostros.
- Se consiguió evaluar y seleccionar los algoritmos de modelos faciales que permitieron la implementación del sistema computacional de reconocimiento de rostros.
- Se desarrolló un software base en algoritmo de modelos faciales como PCA cuya eficacia fue 94.35% siendo mayor a la planteada 90%.
- Se evaluó el rendimiento del sistema computacional obteniendo una sensibilidad de 95.35% y una especificidad de 93.78%.

6.2. Recomendaciones

Luego de desarrollar la presente investigación podemos recomendar continuar el desarrollo de sistemas computacionales de reconocimiento de rostros, haciendo algunas variaciones, como:

- Evaluar otros algoritmos para la etapa de extracción de características, de tal manera que se pueda optimizar la caracterización del modelo.
- En la etapa de comparación e interpretación se pueden evaluar otras técnicas como redes neuronales o redes bayesianas, que permitan aumentar la tasa de efectividad.



REFERENCIAS BIBLIOGRÁFICAS:

Aguilar Torres G., Toscano Medina K., Sánchez Pérez G., Nakano Miyatake M. y Pérez Meana H. (2009). *Eigenface-Gabor algorithm for feature extraction in face recognition*, International Journal of Computers, 3(1), pp. 20-30.

Alexander Ocsa M., Alexander Rodríguez V., Herbert Chuctaya H., Gabriel Humpire M. (2008). *Reconocimiento de Rostros mediante Puntos Característicos Locales*. Universidad San Agustín, Arequipa-Perú.

Berreti, S., Del Bimbo, A., Pala, P., Silva Mata, F. J. *Geodesic distances for 3D-3D and 2D-3D face recognition*. Multimedia and Expo, 2007. IEEE International Conference on Volume, Issue, 2-5. July 2007. Page(s):1515 – 1518.

Blázquez Pérez, L. (2013). *Reconocimiento facial basado en puntos característicos de la cara en entorno no controlado*. Universidad Autónoma de Madrid, España.

Bowyer, K. W., Chang, K., Flynn, P. A. (2005). *Survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition*, Department of Computer Science and Engineering, University of Notre Dame, IN 46556, USA.

Cama Castillo, Y. A. (2015). *Prototipo computacional para la detección y clasificación de expresiones faciales mediante la extracción de patrones binarios locales*. Pontificia Universidad Católica del Perú.

Castellani. (2005). *3D Face Recognition Using Stereoscopic Vision*. Departamento de Informática, Universidad de Verona, Italia.

Ceron, Salazar, Prieto. (2013). *Reconocimiento de Rostros y Gestos Faciales Mediante un Análisis de relevancia con imágenes 3D*. Revista de Investigación, desarrollo e innovación, vol. 4, no. 1, 2014.

Chang, C. y Lin, C., *LIBSVM: A Library for Support Vector Machines*. ACM T Intel Syst Tec, (2011).

Chun-yang L., Dao-zheng H., y Chang-an L. (2013). *Study on moving object detection based on RGB color model*, Advanced Materials Research 710, pp. 700 - 703.

Corcoran P. (2011). *Reviews, Refinements and New Ideas in Face Recognition*, Intech Publisher, Croatia.

Faruqe O. y Hasan A. (2009). *Face Recognition Using PCA and SVM, Anti-counterfeiting, Security and Identification in Communication*, pp. 97-101.

Fernández Sánchez, M. C. (2000). *ÁMBITOS*. Revista Andaluza de Comunicación. Número 3-4 - Universidad de Sevilla.

Francesc Serratos. (2011). *La biometría para la identificación de las personas*. Universidad Abierta de Cataluña.

García Ríos, E., Escamilla Hernández, E., Nakano Miyatake, M., & Pérez Meana, H. (2014). *Sistema de Reconocimiento de Rostros usando Visión Estéreo*. Instituto Politécnico Nacional. México.



- García E., Escamilla E. Aguilar Torres G., Jacobo O., Nakano M. y Perez-Meana H. (2013). *Multibiometric face recognition using level fusión*, NAUN International Journal of Computers, 7(3) pp. 99-108.
- Guerrero Hernández, Gonzalo Pajares Martinsanz, María Guijarro Mata-García. (2010).
- Gökberk, B. *Three dimensional face recognition* – Thesis. 2006. En línea. Accedido 15 de de julio del 2013.
- Gupta, S., Aggarwal, J. K., Markey, M. K., Bovik, A. C. *3D Face recognition founded on the structural Diversity of human faces*. IEEE CVPR. 2007. Minneapolis.MN, USA.
- Marcial-Basilio J., Aguilar-Torres G., Sánchez-Pérez G., Toscano-Medina K. y Pérez-Meana H., *Novedosa técnica para la detección de imágenes pornográficas empleando modelos de color HSV y YCbCr*, Revista Facultad de Ingeniería, 64, pp. 79 - 90, (2012).
- Montalvo Martínez, M. (2010). *Técnicas de visión estereoscópica para determinar la estructura tridimensional de la escena*. Universidad Complutense de Madrid, España.
- Moreno Díaz, A. B. (2004). *Reconocimiento Facial Automático mediante Técnicas de visión tridimensional*. Universidad Politécnica de Madrid, España.
- Kshirsagar V., Baviskar M. y Gaikwad M., *Face Recognition Using Eigenfaces*, International Conference on Computer Research and Development, pp. 302-306 (2011).
- Olivares-Mercado J., Sánchez-Pérez G., Nakano-Miyatake M. y Pérez-Meana H. (2007). *Feature Extraction and Face Verification Using Gabor and Gaussian Mixture Models*, *Lecture Notes in Artificial Intelligence LNAI* pp. 48-27,
- Pereyra, P. A. (2014). *Reconocimiento Facial Mediante Imágenes Estereoscópicas Para Control de Ingreso*. Universidad de Buenos Aires, Argentina.
- Phillips, P. J., Todd Scruggs, W., O'Toole, A. J., Flynn, P. J., Bowyer, K. W., Schott, C. L., Sharpe (2007). *M: FRVT 2006 and ICE 2006 Large-Scale Results*.
- Rodríguez Rodríguez J. A. (2009). *Robot Móvil con visión estereoscópica para la localización de objetos*. Pontificia Universidad Católica del Perú.
- Trussell, H.J., Saber E. y Vrhel M. (2005). *Color image processing: basics and special issue overview*, IEEE Signal Processing Magazine, ISSN 1053-5888, 22(1), pp. 14 - 22.
- Turk M. y Pentland A. (1991). *Eigenfaces for Recongnition*, Journal of Cognitive Neuroscience, 3(1) 71-86,
- Vásquez López, M. A. (2014.). *Sistema de Reconocimiento Facial Mediante Técnicas de Visión Tridimensional*. Centro de investigaciones en óptica, A.C. Ciudad de León, México.
- Villa (2011). *Sistema de Reconocimiento de Rostros*. Universidad Peruana de Ciencias Aplicadas UPC.
- Yu X., Wang X., y Chen J. (2011). *Support Vector Machine Regression for Reactivity Parameters of Vinyl monomers*, *Journal of the Chilean Chemical Society*, 56(3), pp. 746-751.
- Zúñiga P., F. (2015). *Reconocimiento frontal de rostros en base a imágenes de alta resolución*. Universidad de Chile.



ANEXOS

ANEXO 01: Ficha de Verificación

HOJA DE VERIFICACIÓN				
N° Imagen	VP	FP	VN	FN
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	1	0	0	0
5	1	0	0	0
6	1	0	0	0
7	0	1	0	0
8	1	0	0	0
9	1	0	0	0
10	1	0	0	0
11	0	0	0	1
12	0	0	1	0
13	0	0	0	1
14	0	0	0	1
15	0	0	0	1
16	0	0	0	1
17	0	0	0	1
18	0	0	0	1
19	0	0	0	1
20	0	0	0	1
21	1	0	0	0
22	1	0	0	0
23	1	0	0	0
24	1	0	0	0
25	1	0	0	0
26	1	0	0	0
27	1	0	0	0
28	1	0	0	0
29	1	0	0	0
30	1	0	0	0
31	0	0	0	1
32	0	0	0	1
33	0	0	0	1
34	0	0	0	1
35	0	0	0	1
36	0	0	0	1



37	0	0	0	1
38	0	0	0	1
39	0	0	1	0
40	0	0	0	1
41	0	0	0	1
42	0	0	0	1
43	0	0	0	1
44	0	0	0	1
45	0	0	0	1
46	0	0	0	1
47	0	0	0	1
48	0	0	0	1
49	0	0	0	1
50	0	0	0	1
51	0	0	0	1
52	0	0	0	1
53	0	0	0	1
54	0	0	0	1
55	0	0	0	1
56	0	0	0	1
57	0	0	0	1
58	0	0	0	1
59	0	0	1	0
60	0	0	0	1
61	1	0	0	0
62	1	0	0	0
63	1	0	0	0
64	1	0	0	0
65	1	0	0	0
66	1	0	0	0
67	1	0	0	0
68	1	0	0	0
69	1	0	0	0
70	1	0	0	0
71	1	0	0	0
72	1	0	0	0
73	1	0	0	0
74	1	0	0	0
75	1	0	0	0
76	1	0	0	0
77	1	0	0	0
78	1	0	0	0



79	1	0	0	0
80	0	1	0	0
81	1	0	0	0
82	1	0	0	0
83	1	0	0	0
84	1	0	0	0
85	1	0	0	0
86	1	0	0	0
87	1	0	0	0
88	1	0	0	0
89	1	0	0	0
90	1	0	0	0
91	1	0	0	0
92	1	0	0	0
93	1	0	0	0
94	1	0	0	0
95	1	0	0	0
96	0	0	0	1
97	0	0	0	1
98	0	0	0	1
99	0	0	0	1
100	0	0	0	1
101	0	0	0	1
102	0	0	0	1
103	0	0	0	1
104	0	0	1	0
105	0	0	0	1
106	0	0	0	1
107	0	0	0	1
108	0	0	0	1
109	0	0	0	1
110	0	0	0	1
111	0	0	0	1
112	0	0	0	1
113	0	0	0	1
114	0	0	0	1
115	0	0	0	1
116	0	0	0	1
117	0	0	0	1
118	0	0	0	1
119	0	0	0	1
120	0	0	0	1



121	0	0	0	1
122	0	0	0	1
123	0	0	0	1
124	0	0	0	1
125	0	0	0	1
126	0	0	0	1
127	0	0	0	1
128	0	0	0	1
129	0	0	0	1
130	0	0	0	1
131	0	0	1	0
132	0	0	0	1
133	0	0	0	1
134	0	0	0	1
135	0	0	0	1
136	1	0	0	0
137	1	0	0	0
138	1	0	0	0
139	1	0	0	0
140	1	0	0	0
141	1	0	0	0
142	1	0	0	0
143	1	0	0	0
144	0	1	0	0
145	1	0	0	0
146	1	0	0	0
147	1	0	0	0
148	1	0	0	0
149	0	1	0	0
150	1	0	0	0
151	1	0	0	0
152	1	0	0	0
153	1	0	0	0
154	1	0	0	0
155	1	0	0	0
156	1	0	0	0
157	0	1	0	0
158	1	0	0	0
159	1	0	0	0
160	1	0	0	0
161	1	0	0	0
162	1	0	0	0



163	1	0	0	0
164	1	0	0	0
165	1	0	0	0
166	0	1	0	0
167	1	0	0	0
168	1	0	0	0
169	1	0	0	0
170	1	0	0	0
171	1	0	0	0
172	1	0	0	0
173	1	0	0	0
174	1	0	0	0
175	1	0	0	0
176	1	0	0	0
177	1	0	0	0
178	1	0	0	0
179	0	1	0	0
180	1	0	0	0
181	1	0	0	0
182	1	0	0	0
183	1	0	0	0
184	1	0	0	0
185	1	0	0	0
186	1	0	0	0
187	1	0	0	0
188	1	0	0	0
189	1	0	0	0
190	1	0	0	0
191	1	0	0	0
192	1	0	0	0
193	1	0	0	0
194	1	0	0	0
195	1	0	0	0
196	1	0	0	0
197	1	0	0	0
198	1	0	0	0
199	1	0	0	0
200	0	1	0	0
201	1	0	0	0
202	1	0	0	0
203	1	0	0	0
204	1	0	0	0



205	0	0	0	1
206	0	0	0	1
207	0	0	0	1
208	0	0	0	1
209	0	0	0	1
210	0	0	0	1
211	0	0	0	1
212	0	0	0	1
213	0	0	0	1
214	0	0	0	1
215	0	0	0	1
216	0	0	0	1
217	0	0	0	1
218	0	0	0	1
219	0	0	0	1
220	0	0	0	1
221	0	0	0	1
222	0	0	0	1
223	0	0	0	1
224	0	0	0	1
225	0	0	0	1
226	0	0	0	1
227	0	0	0	1
228	0	0	0	1
229	0	0	0	1
230	0	0	0	1
231	0	0	0	1
232	0	0	0	1
233	0	0	0	1
234	1	0	0	0
235	1	0	0	0
236	1	0	0	0
237	1	0	0	0
238	1	0	0	0
239	1	0	0	0
240	1	0	0	0
241	1	0	0	0
242	1	0	0	0
243	1	0	0	0
244	0	0	0	1
245	0	0	0	1
246	0	0	1	0



247	0	0	0	1
248	0	0	0	1
249	0	0	0	1
250	0	0	0	1
251	0	0	0	1
252	0	0	0	1
253	0	0	0	1
254	0	0	0	1
255	0	0	0	1
256	0	0	0	1
257	0	0	0	1
258	0	0	0	1
259	0	0	0	1
260	0	0	0	1
261	0	0	0	1
262	0	0	0	1
263	0	0	0	1
264	0	0	0	1
265	0	0	0	1
266	0	0	0	1
267	0	0	0	1
268	0	0	0	1
269	0	0	0	1
270	0	0	0	1
271	0	0	0	1
272	0	0	0	1
273	0	0	0	1
274	0	0	0	1
275	0	0	0	1
276	0	0	0	1
277	0	0	0	1
278	0	0	0	1
279	0	0	0	1
280	0	0	0	1
281	0	0	0	1
282	0	0	0	1
283	0	0	1	0
284	0	0	0	1
285	0	0	0	1
286	0	0	0	1
287	0	0	0	1
288	0	0	0	1



289	0	0	0	1
290	0	0	0	1
291	0	0	0	1
292	0	0	0	1
293	0	0	0	1
294	0	0	0	1
295	0	0	0	1
296	0	0	0	1
297	0	0	0	1
298	0	0	0	1
299	0	0	0	1
300	0	0	0	1
301	0	0	0	1
302	0	0	0	1
303	0	0	0	1
304	0	0	1	0
305	0	0	0	1
306	0	0	0	1
307	0	0	0	1
308	0	0	0	1
309	0	0	0	1
310	1	0	0	0
311	1	0	0	0
312	1	0	0	0
313	1	0	0	0
314	1	0	0	0
315	1	0	0	0
316	1	0	0	0
317	1	0	0	0
318	1	0	0	0
319	1	0	0	0
320	1	0	0	0
321	1	0	0	0
322	1	0	0	0
323	1	0	0	0
324	0	1	0	0
325	1	0	0	0
326	1	0	0	0
327	0	1	0	0
328	1	0	0	0
329	1	0	0	0
330	1	0	0	0



331	1	0	0	0
332	1	0	0	0
333	1	0	0	0
334	1	0	0	0
335	1	0	0	0
336	1	0	0	0
337	1	0	0	0
338	1	0	0	0
339	1	0	0	0
340	1	0	0	0
341	1	0	0	0
342	1	0	0	0
343	1	0	0	0
344	1	0	0	0
345	1	0	0	0
346	1	0	0	0
347	1	0	0	0
348	1	0	0	0
TOTAL	163	11	8	166



ANEXO 02: Técnicas de programación empleadas

2.1. Detectar Rostro

```
public static final String XML_FILE =
"recursos/haarcascade_frontalface_default.xml";

public CvRect detectFace(IplImage src) {
    CvHaarClassifierCascade cascade = new
    CvHaarClassifierCascade(cvLoad(XML_FILE));
    CvMemStorage storage = CvMemStorage.create();
    CvSeq sign = cvHaarDetectObjects( imagen, cascade, storage, 1.5,
    3, CV_HAAR_DO_CANNY_PRUNING);
    cvClearMemStorage(storage);
    int total_Faces = sign.total();
    if(total_Faces > 0){
        CvRect r = mayorR(total_Faces, sign);
        return r;
    } else {
        CvRect r1 = new CvRect(-1,-1,-1,-1);
        return r1;
    }
}
```

2.2. Recortar imagen

```
public IplImage recort( IplImage src ) {
    IplImage cropped;
    IplImage resizedImage = null;
    FaceDetection fd = new FaceDetection();
    CvRect r = fd.detectFace(src);
    if(r.x() > 0){
        cvSetImageROI(src, r);
        cropped = cvCreateImage(cvGetSize(src), src.depth(),
src.nChannels());
        cvCopy(src, cropped);
        resizedImage = IplImage.create(280, 280, cropped.depth(),
cropped.nChannels());
        cvResize(cropped, resizedImage);
        return resizedImage;
    } else {
        System.out.println("error al recortar");
        return src;
    }
}
```



2.3. Reconocimiento con PCA

```
File root = new File(this.trainingDir);
FilenameFilter jpggFilter = new FilenameFilter() {
    public boolean accept(File dir, String name) {
        return name.toLowerCase().endsWith(".jpg");
    }
};
File[] imageFiles = root.listFiles(jpggFilter);
MatVector images = new MatVector(imageFiles.length);
int[] labels = new int[imageFiles.length];
String[] labels_names = new String[imageFiles.length];

for (File image : imageFiles) {
    img = cvLoadImage(image.getAbsolutePath());
    label = Integer.parseInt(image.getName().split("\\-")[0]);
    label_name = image.getName();
    labels_names[counter] = getName(label_name);
    grayImg = IplImage.create(img.width(), img.height(), IPL_DEPTH_8U,
1);
    cvCvtColor(img, grayImg, CV_BGR2GRAY);
    images.put(counter, grayImg);
    labels[counter] = label;
    counter++;
}

IplImage greyTestImage = IplImage.create(testImage.width(),
testImage.height(), IPL_DEPTH_8U, 1);
FaceRecognizer faceRecognizer = createEigenFaceRecognizer();
faceRecognizer.train(images, labels);
cvCvtColor(testImage, greyTestImage, CV_BGR2GRAY);
int predictedLabel = faceRecognizer.predict(greyTestImage);
```

2.4. Filtro Rojo

```
public BufferedImage SumarImagen(BufferedImage imagenRetorno ,
BufferedImage imagenRetorno1 ) {
public BufferedImage FiltroImagenRojo(BufferedImage image){
    BufferedImage imagenRetorno = clonarBufferedImage(image);
    for( int i = 0; i < imagenRetorno.getWidth(); i++ ){
        for( int j = 0; j < imagenRetorno.getHeight(); j++ ){
            int srcPixel = imagenRetorno.getRGB(i, j);
            Color c = new Color(srcPixel);
            int valR = c.getRed();
            imagenRetorno.setRGB(i, j, this.colorRGBaSRGB(new
Color(valR,0,0, c.getAlpha())));
        }
    }
    return imagenRetorno;
}
```



2.5. Filtro Azul – Verde

```
public BufferedImage FiltroImagenAzul(BufferedImage imagen){
    BufferedImage imagenRetorno=this.clonarBufferedImage(imagen);
    for( int i = 0; i < imagenRetorno.getWidth(); i++ ){
        for( int j = 0; j < imagenRetorno.getHeight(); j++ ){
            int srcPixel = imagenRetorno.getRGB(i, j);
            Color c = new Color(srcPixel);
            int valA = c.getBlue();
            int valV = c.getGreen();
            imagenRetorno.setRGB(i, j,this.colorRGBaSRGB(new
Color(0,valV,valA,c.getAlpha())));
        }
    }
    return imagenRetorno;
}
```

2.6. Suma de imágenes (imagen estereoscópica)

```
for( int i = 0; i < imagenRetorno.getWidth(); i++ ){
    for( int j = 0; j < imagenRetorno.getHeight(); j++ ){
        int srcPixel = imagenRetorno.getRGB(i, j);
        int srcPixel1 = imagenRetorno1.getRGB(i, j);
        Color c = new Color(srcPixel);
        Color c1 = new Color(srcPixel1);
        int valR = c.getRed();
        int valA = c1.getBlue();
        int valV = c1.getGreen();
        imagenRetorno.setRGB(i, j,this.colorRGBaSRGB(new
Color(valR, valV, valA)));
    }
}
return imagenRetorno;
}
```

