

Sistema de detección y reconocimiento facial mediante técnicas híbridas en imágenes y secuencias de vídeo.

Facial detection and recognition system using hybrid techniques in images and video sequences

Renzo Apaza¹, <https://orcid.org/0000-0002-7089-2812>

Gina F.², <https://orcid.org/0009-0008-8718-8099>

Charaja, Vladimiro Ibañez³, <https://orcid.org/0000-0002-0277-4945>

Ernesto N. Tumi⁴, <https://orcid.org/0000-0003-0401-9027>

¹Universidad Nacional del Altiplano, Puno., delujoprimer@gmail.com, Puno, Perú

²Universidad Nacional del Altiplano, fiorechs20@hotmail.com, Puno, Perú

³Universidad Nacional del Altiplano, viq_ibanez@hotmail.com, Puno, Perú

⁴ Universidad Nacional del Altiplano, viq_ibanez@hotmail.com, Puno, Perú

Resumen - El reconocimiento automático de la identidad de un individuo aún no alcanza una tasa de éxito del 100%. Por ello, este trabajo buscó mejorar dicha tasa, haciendo hincapié en los métodos de extracción y clasificación de características. En este sentido, se propuso mejorar la tasa de reconocimiento facial mediante la representación de las imágenes, utilizando la transformada wavelet de Gabor aplicada a las imágenes en escala de grises obtenidas tras la normalización de las originales. Posteriormente, se aplicó el análisis de componentes principales (PCA) a la nueva representación para obtener y formar el vector de características de las imágenes faciales. A continuación, se aplicó un clasificador basado en máquinas de vectores de soporte (SVM). El método se probó en una base de datos de imágenes faciales compuesta por imágenes de los bancos de rostros FERET y ORL, así como imágenes obtenidas por los responsables de la investigación. Como resultado de la combinación de la transformada wavelet de Gabor y el análisis de componentes principales en el proceso de extracción de características y la clasificación de imágenes mediante máquinas de vectores de soporte, se logró una tasa de reconocimiento superior al 95%.

Palabras clave: Reconocimiento facial, ondícula de Gabor, análisis de componentes principales (PCA), máquina de soporte vectorial (SVM), vector de características.

Abstract - Recognizing the identity of an individual automatically is still a task that fails to achieve a success rate of 100%, so the present work sought to improve the recognition rate; giving emphasis to the methods of extraction and classification of characteristics. In this sense, it was proposed to improve the ratio for the recognition of faces through the representation of the images, using the Wavelets transform of Gabor on the grayscale images obtained after normalizing the original images, afterwards the new representation obtained was applies the Principal Component Analysis (PCA) technique to obtain and then form the feature vector of face images. Next, a classifier based on Vector Support Machines (SVM) is applied. The method was tested on a database of images of faces constituted between the banks of faces FERET, ORL and images obtained by those responsible for the investigation. As a result of the combination of Gabor Wavelet Transform and Principal Components Analysis techniques in the process of extracting features and image classification based on Vector Support Machines, a recognition rate higher than 95% is achieved.

Keywords: :Face Recognition, Gabor Wavelet, Principal ComponentAnalysis (PCA), Vector Support Machine (SVM), Feature Vector.

1. INTRODUCCIÓN

Dentro de la amplia gama de posibles actividades que pueden automatizarse, aquella relacionada con la capacidad para establecer la identidad de los individuos ha cobrado una gran importancia y como consecuencia directa, la biometría ha transformado en un área resaltante.

Las técnicas de identificación biométrica permiten el reconocimiento de la identidad de un individuo basado en las características conductuales o en los rasgos físicos propios intrínsecos de los individuos, garantizando de esta manera que un individuo no pueda usurpar la identidad de otro. El reconocimiento de rostros es considerado por los investigadores como una técnica importante de la

biometría, siendo una habilidad de identificar a los sujetos basados en sus características faciales.

La seguridad de las personas, bienes o información es una de las mayores preocupaciones de nuestra sociedad local, nacional e internacional hoy en día. En tal sentido son necesarios los medios de verificación de identidad para salvaguardar la seguridad e integridad personal, material e intelectual. Así mismo los actuales medios de verificación de identidad están relacionados a las pertenencias de los individuos como los documentos de identidad, tarjetas de crédito y pasaporte. Sin embargo, las pertenencias personales pueden ser sustraídas sin autorización, dañadas físicamente o ser sometidas a procedimientos para revelar las contraseñas a través de algoritmos de fuerza bruta que conlleva a la usurpación de la identidad.

En tal sentido, la presente investigación aborda la detección y el reconocimiento automático de rostros haciendo énfasis en el proceso de extracción de características mediante la combinación de las técnicas Wavelets de Gabor y Análisis de Componentes Principales para posteriormente realizar la identificación mediante un clasificador basado en Maquinas de soporte Vectorial.

2. REVISIÓN LITERARIA

El reconocimiento de rostros, se considera como una parte importante de la técnica de biometría, y significativa en la investigación científica [1]. Es la capacidad para establecer la identidad de un individuo en base a sus características faciales. El reconocimiento automático del rostro, ha sido ampliamente estudiado en las últimas décadas debido a su importante papel en una serie de dominios de aplicación, tales como control de acceso, vigilancia visual [2].

Desde los primeros trabajos realizados por Takeo Kanade [3] que desarrolló un sistema automático de rostros, [4][5] no fue sino hasta el trabajo de Turk y Pentland [6] sobre Eigenface que la investigación en reconocimiento de rostros cobró mayor relevancia.

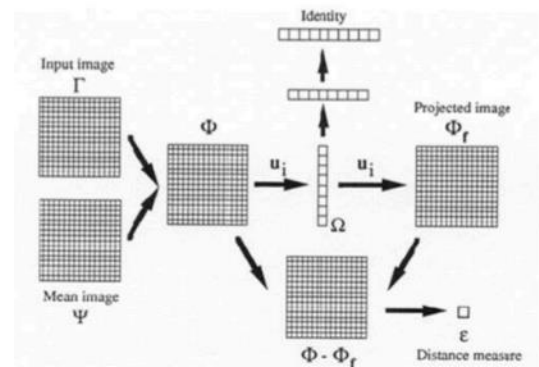


Fig 1. Colección de redes para implementar el cálculo de vector de patrones, proyección en el espacio, medida de distancia del espacio de rostros, e identificación.

Desde que se propuso el problema, [6] que presentaron un sistema de reconocimiento de rostros en tiempo casi real mediante la introducción de Eigenface en la extracción de características de imagen facial, otros trabajos que marcaron un hito importante fueron: el método de Fisherface [7] [8] que aplicaba el Análisis Discriminatorio Lineal LDA después de una etapa de PCA para lograr una mayor precisión, también se resalta el uso de filtros locales como los Jets de Gabor [9] [10] y el trabajo de [11] que propusieron una técnica de reconocimiento efectivo usando Análisis de Componentes Principales y Support Vector Machine.

El reconocimiento facial mediante filtros Gabor se introdujo en primer lugar por Martin Lades [9], y pronto demostró ser un medio muy eficaz en la extracción de rasgos faciales. Xiaoming [12] propuso un algoritmo de reconocimiento facial combinando una serie de características de vectores que consiste en la magnitud de Gabor, PCA y SVM para la clasificación.

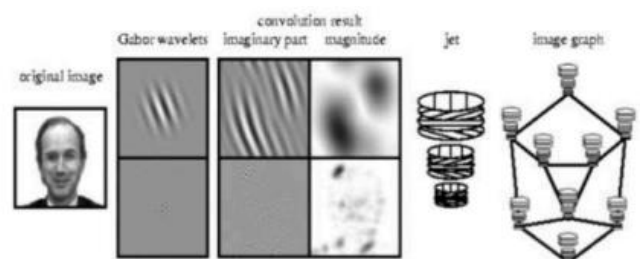


Fig. 2. Representación gráfica de un rostro basado en la transformada Wavelets de Gabor. una convolución con un conjunto de kernels wavelets que tienen la forma de olas planas restringidas por una función envolvente gaussiana.

En general, muchos métodos son propuestos para abordar la difícil tarea de reconocimiento de rostros. Un buen método de reconocimiento de rostros debe considerarse como la representación del rostro en los problemas de clasificación, y un

buen método de representación de rostros requiere de un mínimo de anotaciones manuales.

3. METODOLOGÍA

3.1 Método general de reconocimiento de rostros

Para la elaboración del sistema y posterior tratamiento de los datos obtenidos, se optó por los procesos de un sistema general de reconocimiento de rostros, que contempla cuatro etapas la adquisición de una imagen o video, la detección de rostros, localización y extracción de características y la clasificación (Fig. 3).

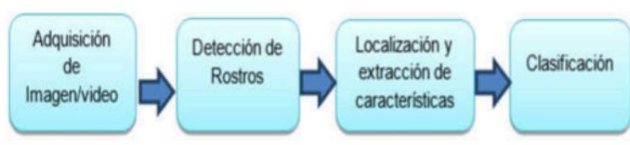


Fig.3. Diagrama de bloques de un esquema general para el reconocimiento de rostros.

3.2 Etapas del reconocimiento de rostros

El método propuesto para el reconocimiento de rostros contempla dos etapas:

- Entrenamiento y
- Reconocimiento.

Entrenamiento

El entrenamiento prepara al sistema con datos etiquetados que permiten obtener los vectores característicos que son almacenados para realizar posteriormente la recuperación de estos. En esta etapa se siguen los procesos de: adquisición de la imagen desde una base de datos, detección de rostro en la imagen, representación de la imagen mediante la Wavelet de Gabor y extracción de características mediante el análisis de componentes principales PCA.

Reconocimiento.

el reconocimiento consiste en la operación de adquirir una imagen nueva no incluida en la etapa de entrenamiento y posteriormente asignar a la imagen, una clase almacenada en la base de datos, los procesos seguidos en esta etapa son: adquisición de una imagen de prueba, detección de rostro en la imagen, representación de la imagen mediante la Wavelet de Gabor, extracción de características mediante el análisis de componentes principales (PCA) y clasificación de la imagen a través de máquina de soporte vectorial SVM.

La Fig. 4 resume los dos procesos propuestos para la tarea de reconocimiento de rostros.

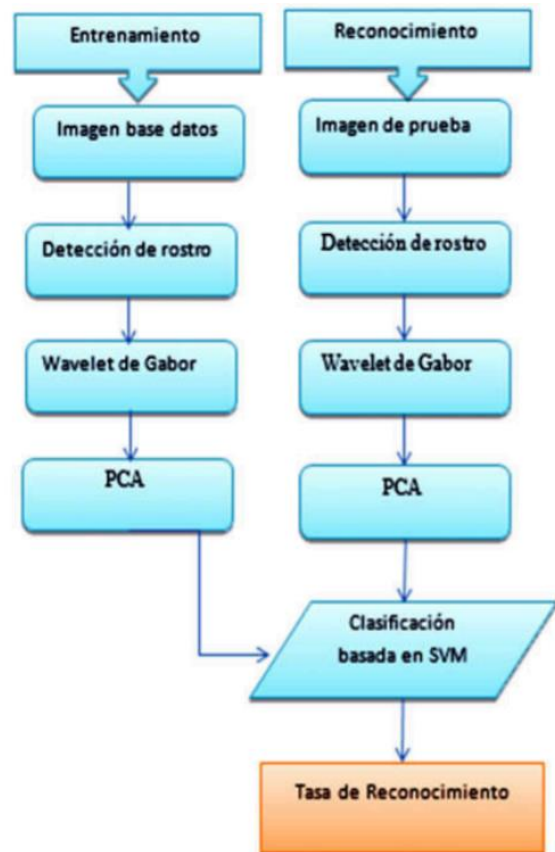


Fig. 4 Diagrama de procesos de reconocimiento de rostros propuesto.

3.3 Métodos para la detección de rostros

Para lograr el propósito de detectar rostros, se utilizó el clasificador en cascada que usa las características basadas en Haar que es un método efectivo para la detección de objetos propuesto por Paul Viola y Michael Jones en el paper, "Rapid Object Detection using a Boosted Cascade of Simple Features" en el año 2001. Es un método basado en aprendizaje de la máquina donde una función en cascada es entrenada con imágenes positivas e imágenes negativas, que es usada para detectar diferentes tipos de objetos.



3.4 Método para la extracción de características a.Wavelet de Gabor

Utiliza el filtro de Gabor que es definido como:

$$\Psi_{u,v}(z) = \frac{||k_{u,v}||^2}{\sigma^2} e^{(-||k_{u,v}||^2 ||z||^2 / 2\sigma^2)} \left[e^{ik_{u,v}z} - e^{-\frac{\sigma^2}{2}} \right]$$

Donde u y v definen las escalas y orientaciones de los filtros de Gabor, $z = (x, y)$ y $k_{u,v}$ está definida como:

$$k_{u,v} = k_v e^{i\varphi u}$$

$k_v = k_{max}/f^v$ y $\varphi u = \pi u/8$, k_{max} es la frecuencia máxima, y f es el factor de espacio entre los Kernels en el dominio de la frecuencia. En el presente

$$\sigma = 2\pi, k_{max} = \frac{\pi}{2}, f = \sqrt{2}, u \in \{0, 1, \dots, 7\} \text{ y } v \in \{1, 2, 3, 4\}.$$

La representación de la imagen de rostro mediante la Wavelet de Gabor es obtenida mediante la convolución entre la imagen y la familia de filtros de Gabor descritas en la ecuación:

$$F_{u,v}(Z) = I(Z) * \Psi_{u,v}(Z)$$

Donde $z = (x, y)$, $*$ denota el operador de convolución y $F_{u,v}$ es la respuesta del filtro de Gabor de la imagen con orientación u y escala v .

Se convolucionan la imagen con 40 kernels de Gabor para generar las características de Gabor. La imagen de entrada es una imagen de rostro que es geométricamente normalizado de 64 x 64 píxeles. Fig. Filtros de Gabor con diferentes orientaciones y escala

Así el tamaño del vector es (64 x 64 x 40 x 2).

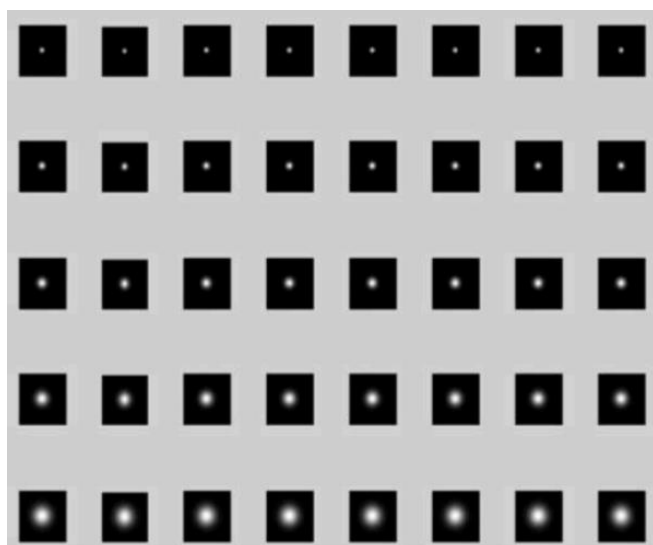


Fig. 6 Filtros de Gabor con diferentes orientaciones y escala

b. Convolución de la Imagen

La convolución es realizada entre todas las partes de la imagen y un operador (Kernel), el kernel para el presente trabajo son el banco de filtros de Gabor, se siguió el procedimiento para efectuar la operación:

Ubicar el ancla del kernel (centro de la imagen) en lo alto de un determinado pixel, con el resto del kernel sobrepuesto en los píxeles locales de la imagen.

- Multiplicar los coeficientes del kernel por el valor del pixel correspondiente y sumar el resultado.
- Localizar el resultado en la ubicación del ancla en la imagen de entrada.
- Repetir el proceso para todos los píxeles desplazando el kernel sobre la totalidad de la imagen.

La siguiente ecuación, expresa el procedimiento descrito anteriormente

$$H(x, y) = \sum_{i=0}^{M_i-1} \sum_{j=0}^{M_j-1} I(x+i-a_i, y+j-a_j) K(i, j)$$

c. Análisis de componentes Principales (PCA)

En esta fase, cargamos la base de datos. En general, aplicamos muchas transformaciones antes de cargarlas. En efecto, la señal contiene información útil para el reconocimiento y solo los parámetros relevantes son extraídos. El modelo es una representación compacta de la señal que facilita la fase de reconocimiento, pero también reduce la cantidad de datos a ser almacenados.

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,m} \end{pmatrix} \rightarrow \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{n,1} \\ \vdots \\ a_{1,m} \\ \vdots \\ a_{n,m} \end{pmatrix}$$

Entonces, el conjunto de entrenamiento de imágenes de rostros será $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M$. El rostro promedio del conjunto es definido por la ecuación:

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i$$

Cada rostro, se diferencia del promedio por el vector de la siguiente ecuación:

$$\Phi_i = \Gamma_i - \Psi, i = 1, \dots, M$$

A continuación, calculamos la matriz de covarianza C, según:

$$C = \sum_{i=1}^N \Phi_i \Phi_i^T = AA^T$$

La matriz C es N2 por N2, y determina los N2 vectores y valores propios, se vuelve una tarea intratable para el tamaño de una imagen típica. Necesitamos un método factible para encontrar los eigen vectores ecuación:

$$\begin{pmatrix} e_i = & Av_i \\ \lambda_i = & \mu_i \end{pmatrix}$$

De M eigen vectores (eigen faces) e_i , solo escogemos M1 que tiene el valor propio (eigen value) más alto. Cuanto más alto sea el valor propio, describe la característica más representativa del rostro de un particular eigenvector. Los eigen face con bajos eigen values pueden ser omitidos, porque ellos solo explican una pequeña parte de los rasgos característicos del rostro. Luego M1 eigen faces e_i son determinados, la fase de entrenamiento del algoritmo es finalizado.

El proceso de clasificación de un nuevo rostro F a una de las clases procede en dos pasos.

Primero, la nueva imagen es transformada en sus componentes eigen face, Los pesos resultantes forman el vector de pesos mostrados Ω_T según la siguiente ecuación

Los pesos forman un vector $\Omega_T = [w_1, w_2, \dots, w_M]$ que describen la

$$W_k = e_k^T (\Gamma_{new} - \Psi), k=1, \dots, M', \Omega_T = [w_1, w_2, \dots, w_M]$$

contribución de cada eigenface de la representación de la imagen de rostro de entrada, tratando los eigen faces como un conjunto de bases para las imágenes de rostro. El vector puede ser usado con un algoritmo de reconocimiento de patrón estándar para encontrar a qué número de clases predefinidas pertenece. El método utilizado para determinar qué clase de rostro provee la mejor descripción de una imagen de rostro de entrada es la máquina de soporte vectorial SVM.

3.5 Método de evaluación de eficiencia del sistema

Para determinar la eficiencia del sistema, se estima la tasa de falsa aceptación (rate of false acceptance- FAR):

$$FAR = \frac{\text{número de falsas aceptaciones}}{\text{número de impostores}}$$

la tasa de falso rechazo (rate of false rejection - FRR):

$$FRR = \frac{\text{número de falsos rechazos}}{\text{número de rostros evaluados}}$$

y la tasa de igual error (equal error rate -ERR) ecuación:

$$EER = \frac{\# \text{ de falsas aceptaciones} + \# \text{ falsos rechazos}}{\# \text{ total de entradas}}$$

4. ANÁLISIS DE RESULTADOS

Se implementaron los módulos en función de las arquitecturas diseñadas en lenguaje de programación C++ y con la librería Opencv posteriormente se evalúa la efectividad del sistema:

4.1 Adquisición de imágenes y secuencias de video

La arquitectura diseñada para el desarrollo del Módulo, se resume la interacción con el usuario, el componente administrador de archivos y el uso de la cámara web. Se resalta que para el uso de las imágenes adquiridas por la Webcam es necesaria la interacción con los drivers del Sistema Operativo.

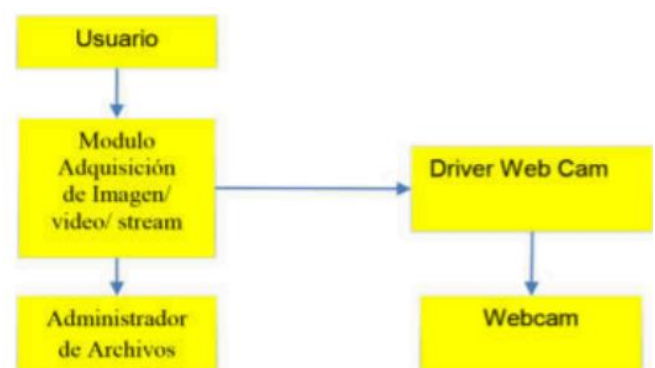


Fig. 7. Arquitectura del Módulo de adquisición de Imágenes y Secuencias de video.



Fig. 8. Adquisición de imagen en tiempo real realizada con el sistema.

4.2 Detección de rostros en imágenes y secuencias de video

En la Fig. 9, se muestra el modelo de la Arquitectura del módulo.

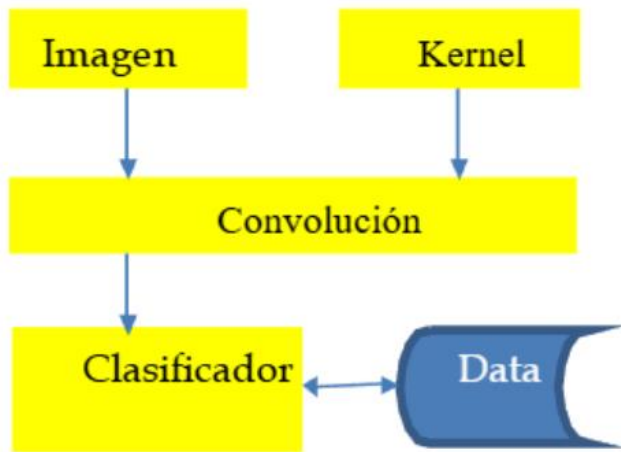


Fig. 9. Adquisición de imagen en tiempo real realizada con el sistema

La captura de la imagen a ser analizada es realizada en el espacio de colores RGB (red, green y blue), si la imagen tiene dimensiones muy altas, los procesos para la detección de rostros son lentos es por ello que las imágenes son reducidas a un espacio de 240 x 320 píxeles y la profundidad de la imagen es reducida a un solo canal de escala de Grises. La Fig. 10, muestra la reducción de dimensión y color de la imagen original capturada por el software.

A continuación, se realiza la clasificación de la imagen mediante el clasificador en cascada basado en las características de Haar iniciando un kernel de 20 x 20 píxeles, la Fig. 10, muestra la detección de rostro basado en las características de Haar.

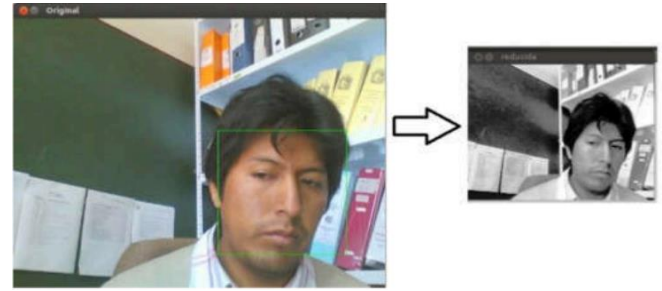


Fig. 10. Reducción de dimensiones de la imagen en tamaño y color.

4.3 Extracción de características basadas en la Wavelet de Gabor y Análisis de Componentes.

Principales El modelo de la Arquitectura del módulo de extracción de características es mostrado en la Fig. 11, se detallan los componentes de pre procesamiento, Wavelets de Gabor, convolución PCA y los vectores de características.

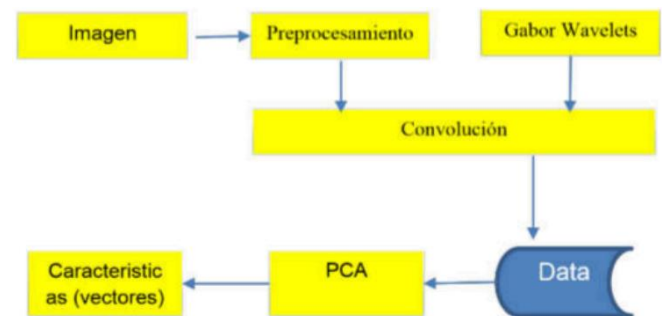


Fig. 11: Arquitectura del módulo de extracción de características.

Se aplicaron 40 filtros de Gabor a la imagen con ocho escalas diferentes y cinco direcciones diferentes, estas imágenes constituyen los vectores de características basadas en la Wavelet de Gabor.

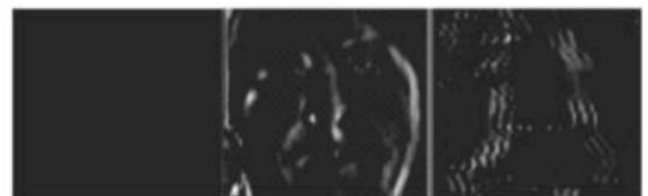


Fig. 12: Imágenes filtradas con la Wavelet de Gabor.

Finalmente aplicamos el Análisis de Componentes Principales a las imágenes concatenadas de cada rostro que se aplicó la Wavelet de Gabor. De esta manera reducimos la dimensionalidad de la imagen, y se constituye el vector de características de la imagen.

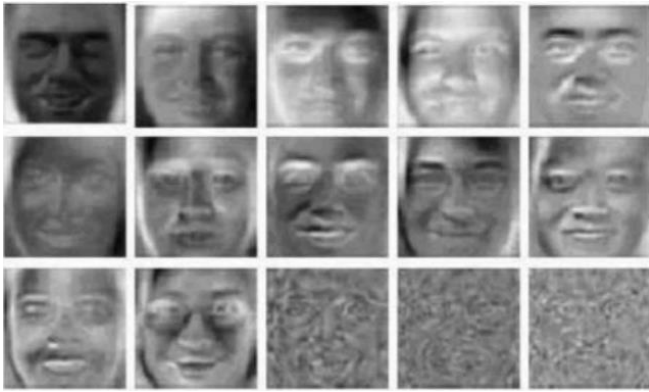


Fig. 13: Imágenes luego de aplicar PCA.

4.4 Clasificación de rostros basado en Máquinas de Soporte Vectorial

La Fig. 14, resume la arquitectura del módulo de clasificación de rostros

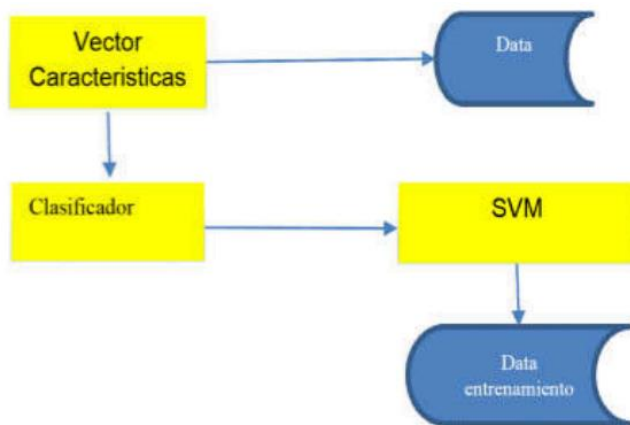


Fig. 14: Arquitectura del módulo de clasificación de rostros.

4.5 Evaluación del sistema

a. Tasa falsa de aceptación (FAR)

Datos:

- Número de falsas aceptaciones = 1
- Numero de impostores = 48

Cálculo de FAR

$$FAR = \frac{\text{número de falsas aceptaciones}}{\text{número de impostores}} = \frac{1}{48} = 0.02 = 2\%$$

La tasa falsa de aceptación (FAR) es igual al 0.02 que representa el 2%, que indica que para la prueba realizada existe un 2% de posibilidad de aceptar un usuario no registrado en la base de datos.

B Tasa falso de rechazo (FRR)

Datos:

- Número de falsos rechazos = 1

- Numero de rostros evaluados = 48

$$FRR = \frac{\text{número de falsos rechazos}}{\text{número de rostros evaluados}} = \frac{1}{48} = 0.02 = 2\%$$

La tasa falso de rechazo (FRR) es igual a 0.02 que representa el 2%, que indica que para la prueba realizada existe un 2% de posibilidad de rechazar un usuario registrado en la base de datos.

C Tasa de Error Igual

Datos:

- Número de falsas aceptaciones = 1
- Número de falsos rechazos = 1
- Número total de entradas = 96

La tasa de error igual es 0.02 que representa el 2% que

$$EER = \frac{\text{número de falsas aceptaciones} + \text{número falsos rechazos}}{\text{número total de entradas}}$$

$$EER = \frac{1 + 1}{96} = 0.02 = 2\%$$

indica que para la prueba realizada existe un 2% de cometer un error (tipo I o tipo II), también podemos afirmar que el sistema proporcionó resultados exactos en un 98%.

5. CONCLUSIONES

Se verificó que el desarrollo del sistema para la detección y reconocimiento de rostros basadas en técnicas híbridas logró alcanzar un porcentaje del 98% superando al porcentaje esperado que fue 95%, además, se obtuvo una tasa de falsos aceptados del 2% y una tasa de falsos rechazados del 2%.

La base de datos de imágenes utilizadas constituida por imágenes de tres colecciones diferentes (FERET, ORL, y la de los investigadores) no presentaron diferencias significativas para el desarrollo y prueba del sistema. Además, la implementación del módulo de detección de rostros basada en un clasificador a cascada fue útil para poder realizar la posterior tarea de reconocimiento de rostros. El módulo implementado para la detección de rostros fue satisfactoriamente.

La combinación de las técnicas Wavelet de Gabor y Análisis de Componentes Principales en las tareas de extracción de características, logró alcanzar una tasa de falsos aceptados del 2% y una tasa de falsos rechazados del 2%,

demostrando un alto grado de confiabilidad para la constitución de vectores característicos. De esta manera las Wavelets de Gabor demostraron ser adecuadas para la extracción de característica, debido a sus propiedades de orientación, localización espacial y óptima localización en el espacio y dominio de la frecuencia. El algoritmo de Análisis de Componentes Principales (PCA) es un método global que usa principalmente los niveles de gris de los píxeles de una imagen. La simplicidad para la implementación del algoritmo contrasta con una fuerte sensibilidad a cambios en la luz, posición y expresiones faciales. Es por eso que aumentamos el número de posturas para cada persona. Sin embargo, el ACP no requiere ningún conocimiento a priori sobre la imagen. El principio es construir un espacio de sub-vectores conservando sólo los mejores vectores propios, al tiempo que conserva una gran cantidad de información útil, hace que el PCA sea un algoritmo efectivo y de uso común en la reducción de la dimensionalidad, donde puede ser utilizado con otros algoritmos para mejorar los resultados de la aplicación.

La ejecución del clasificador basado en máquinas de Soporte Vectorial, realiza de manera efectiva el reconocimiento de rostros, alcanzando un porcentaje del 98% de efectividad, luego de la evaluación. Así, se confirma que el reconocimiento de individuos mantiene un problema complejo y que a pesar de ser un área activa de investigación aún existen muchas dificultades en su aplicación en condiciones reales, se ha logrado una tasa de reconocimiento efectivo de 98%, no logrando resolver el problema en un 100% de los casos.

REFERENCIAS

- [1] J. Zhou, Z. Ji, L. Shen, Z. Zhu and S. Chen, "PSO Based Memetic Algorithm for Face Recognition Gabor," in IEEE Conference Memetic Computing, 2011.
- [2] A. K. Jain, B. Klare and U. Park, "Face Recognition: Some Challenges in Forensics," in IEEE International Conference Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011.
- [3] K. Takeo, "Picture processing system by computer complex and recognition of human faces," Kyoto University, Kyoto, 1973.
- [4] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the Characterization," in IEEE Trans. Pattern Anal, Vol. 12, Nro 1, 1990.
- [5] L. Sirovich and M. Kirby, "Low dimensional procedure for the characterization of human faces," in J. Opt. Soc. Am, Vol 4 Nro 3, 1987.
- [6] M. Turk and A. Pentland, "Face recognition using eigenfaces," in IEEE Conference on computer Vision and Pattern Recognition, Hawaii, 1991.
- [7] P. Belhumeur, J. Hespanha and D. Kriegman, "Eigenfaces vs. Fisherfaces," in Trans. Pattern Anal. Mach, vol. 19, Nro 7, 1997.
- [8] K. Etemad and R. Chellapa, "Face recognition using discriminant eigenvectors," in Proceedings of the International Conference on Acoustic, Speech and Signal Processing, 1996.
- [9] M. Lades, J. Buhmann, C. Vonder Malsburg, R. Wurtz and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," in IEEE Trans. Comput, vol. 42, 1993.
- [10] L. Wiskott, J. Fellous, N. Kruger and C. Malsburg, "Face recognition by elastic bunch graph," in IEEE Trans. Pattern Anal Mach, vol 19 Nro 7, 1997.
- [11] C. Wang, L. Lan, Y. Zhang and M. Gu, "Face Recognition Based on Principle Component Analysis and Support Vector Machine," in IEEE Conference Intelligent Systems and Applications (ISA), 2011.
- [12] M. Lades, J. Vorbruggen and J. Buhmann, "Distortion invariant object recognition in the dynamic link architecture," in IEEE Transactions on Computers, 1993.