

Análisis de un sistema multi-biométrico de reconocimiento facial usando wavelet $up(x)$

Analysis of a multi-biometric face recognition system using wavelet $up(x)$

E. García-Rios ^a, E. Escamilla-Hernández ^b, H. M. Pérez-Meana ^b, L. E. Ramos-Velasco ^c, M. Pérez-Bautista ^a,
O. V. Kravchenko ^d

^a Área Académica de Sistemas Computacionales, Instituto Tecnológico Superior del Occidente del Estado de Hidalgo, 42700 Mixquiahuala, Hgo., México.

^b Sección de Estudios de Posgrado e Ingeniería, Escuela Superior de Ingeniería Mecánica y Eléctrica Unidad Culhuacán, 04440, CDMX., México.

^c Maestría en Ingeniería Aeroespacial, Universidad Politécnica Metropolitana de Hidalgo, 43860 Tolcayuca, Hgo., México.

^d Instituto de investigación de Ingeniería Mecánica de RAS, Universidad Técnica Estatal de Bauman de Moscú, 2-Ya Baumanskaya Utilisa, 105005, Moscú, Rusia.

Resumen

En este artículo se presenta el análisis de un sistema multi-biométrico de reconocimiento facial con uso de imágenes estéreo. En donde, se propone como algoritmo de extracción de características el uso de la Wavelet $up(x)$, la cual está basada en funciones atómicas y como clasificador la Máquina de Soporte Vectorial (SVM). El sistema multi-biométrico de rostros está constituido por el uso de distintos niveles de fusión para utilizar las características faciales estéreo. Al evaluar los resultados del sistema con la wavelet $up(x)$ en comparación con otros algoritmos de extracción de características (Filtros de Gabor, DWT de Haar y la DWT Daubechies de tercer orden) se obtiene que este proporciona un mayor porcentaje de reconocimiento; debido a que los coeficientes de esta función wavelet tienen una mejor respuesta en tiempo - frecuencia, otorgando así una mejor extracción de características dando un mayor porcentaje de reconocimiento dentro del sistema.

Palabras Clave: Sistema multi-biométrico, Función Atómica, wavelet $up(x)$, niveles de fusión, reconocimiento facial, imágenes estéreo.

Abstract

This paper presents an analysis of a multi-biometric face recognition system using stereo images. It's proposed the use of Wavelet $up(x)$ which is based on atomic functions as feature extraction algorithm, and the Support Vector Machine (SVM) as classifier. The multi-biometric face system is constituted by different fusion levels employing stereo facial features. The evaluation the results obtained using the proposed recognition system with the wavelet $up(x)$ in comparison with other feature extraction algorithms (Gabor Filters, Haar DWT and the third order Daubechies DWT) show that it provides a higher percentage of recognition; because the coefficients of this wavelet function have a better performance in time – frequency response, thus giving a better extraction of features, resulting in a higher recognition percentage of the proposed system.

Keywords: Multi-biometric system, Atomic Function, wavelet $up(x)$, fusion levels, facial recognition, stereo vision.

1. Introducción

Los sistemas de reconocimiento multi-bimétrico (Anil K, 2008) son aquellos que permiten hacer la combinación de

información en una o más modalidades de información biométrica, sensores o algoritmos para el reconocimiento biométrico; donde en los sistemas de reconocimiento facial se han propuesto ser sistemas de este tipo como estrategia en la

*Autor para la correspondencia: egarcia@itsoeh.edu.mx

Correo electrónico: egarcia@itsoeh.edu.mx (Elizabeth García-Rios), eescamillah@ipn.mx (Enrique Escamilla-Hernández), hmperez@ipn.mx (Héctor Manuel Pérez-Meana), lramos@upmh.edu.mx (Luis Enrique Ramos-Velasco), mperez@itsoeh.edu.mx (Mario Pérez-Bautista), olekravchenko@gmail.com (Oleg Victorovich - Kravchenko)

búsqueda de mejorar su rendimiento funcional. Además, el interés de mejorar dichos sistemas biométricos es porque son un método no intrusivo los cuales son utilizados cada vez más en aplicaciones comerciales debido a la rápida evolución de la tecnología electrónica. Como resultado, se han propuesto varios algoritmos eficientes durante las últimas dos décadas y muchos de los cuales, incluso están actualmente disponibles en el mercado. Sin embargo, casi todos aún requieren algunas condiciones especiales para su funcionamiento óptimo, tales como: la pose, orientación facial, iluminación de fondo controlado y expresiones faciales neutras o bien sin expresión. Por lo que, en el presente trabajo, se propone analizar un sistema multi-biométrico de reconocimiento facial usando imágenes estéreo de rostros empleando distintos niveles de fusión (García-Ríos E., 2013; García E., 2014), reemplazando los métodos de extracción de características antes implementados en los trabajos previos, por la wavelet $up(x)$ con Funciones Atómicas (Kravchenko VF, 2009; Kravchenko V. K., 2015). Donde, el principal objetivo es obtener un análisis comparativo del sistema multi-biométrico facial, entre los métodos de extracción de características antes implementados en (García-Ríos, 2013; García, 2014) y la función wavelet $up(x)$, realizando un cambio en el uso de los coeficientes de la wavelet con las Funciones Atómicas (AF 's). Esto en la búsqueda de obtener un mejor rendimiento en el reconocimiento e identificación facial; debido a que el uso de estas wavelets, han presentado mejores resultados en diferentes aplicaciones (Hernández-Matamoros, 2020; Perez-Daniel, 2012). Prestando especial atención a la extracción de características del rostro, ya que juega un papel importante para la caracterización y reconocimiento facial.

2. Sistema multi-biométrico de reconocimiento facial

En el presente trabajo se hace un análisis del comportamiento de un sistema multi-biométrico, en donde el sistema esta principalmente constituido con el uso de distintos niveles de fusión y el uso de imágenes estéreo. A continuación, se presenta el diagrama del sistema propuesto, ver Figura 1.

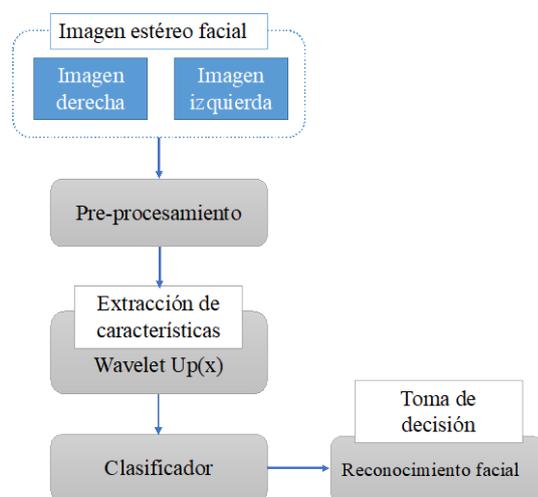


Figura 1: Diagrama general del sistema multi-biométrico facial.

Básicamente la integración del sistema propuesto está constituida por los siguientes módulos: entrada de datos, pre-procesamiento, extracción de características, clasificación y la

toma de decisión para poder realizar reconocimiento de los rostros.

2.1. Entrada de datos.

En este primer módulo del sistema se ingresan los datos que servirán para el análisis y aprendizaje del sistema. En donde, se obtienen una base de datos con imágenes estéreo, la cual fue creada con la cámara FUJI FILM Fine Pix Real 3D W1, adquiriendo fotografías de distintos rostros. Cabe destacar que dichas imágenes fueron tomadas bajo un entorno controlado, es decir, que el fondo que contiene cada una de las imágenes capturadas, fue bajo condiciones controladas de: iluminación, pose y distancia.

2.2. Pre-procesamiento.

Después de la entrada de los datos, el siguiente módulo desempeña el primer procesamiento hacia las imágenes de entrada, teniendo como objetivo la eliminación de ruido que se llega a presentar ante la imagen capturada, donde en este caso esencialmente se realiza una reducción de dimensión de la imagen tratando de excluir la mayor parte del fondo, de manera que se conserve solamente la región de la imagen que contiene el rostro para después poder realizar un mejor análisis de esta en los siguientes procesos. Para la reducción de la dimensión de la imagen de entrada, esta se efectúa mediante una operación de filtrado pasa bajas, programado en lenguaje de Matlab (García E., 2014).

2.3. Extracción de características con la Wavelet $up(x)$.

Una vez que la etapa de pre-procesamiento elimina la mayor parte del fondo y la imagen bajo análisis corresponde solamente al rostro, se procede a la obtención de los vectores característicos con el uso de la wavelet $up(x)$, como método de extracción de características; continuando la programación de este, en el lenguaje de programación de Matlab.

La wavelet que se implementarán en este sistema, esta basada en la teoría de Funciones Atómicas (AF 's), donde las Funciones Atómicas es la parte fundamental para comprender la wavelet $up(x)$. La teoría de estas funciones, las cuales son soluciones infinitamente diferenciables compatibles de soporte compacto definidas en \mathbb{R}^2 para ecuaciones con argumento desplazado.

$$Lf(x) = \lambda \sum_{k=1}^M c(k) f(ax - b), \quad |a| > 1 \quad (1)$$

Donde L denota el operador diferencial lineal con coeficientes constantes; la función madre de las AF 's, llamada $up(x)$, la cual es utilizada en este trabajo, se representa por su Transformada de Fourier como una convolución infinita de impulsos rectangulares.

$$up(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega x} \prod_{k=1}^{\infty} \frac{\sin \omega \cdot 2^{-k}}{\omega \cdot 2^{-k}} d\omega \quad (2)$$

Algunas Propiedades de esta ventana son:

$$up(x) = up(-x); \int_{-1}^1 up(x)dx = 1; up(x) = 1 - up(x);$$

$$up'(x) = 2 * up(x)(2x + 1) - 2 * up(2x - 1) \tag{3}$$

Entonces la wavelet $up(x)$ es construida con base a esta función $up(x)$, que se muestra en Figura 2(a); donde en la Figura 2(b), se observa el espectro de $up(x)$ es una función par con un decrecimiento acelerado conforme nos alejamos de frecuencia central, que es muy apropiado para la construcción de la función wavelet. La construcción de la función wavelet inicia con la obtención la función de escala $\phi(x)$, que es expresada como una suma de traslaciones y dilataciones.

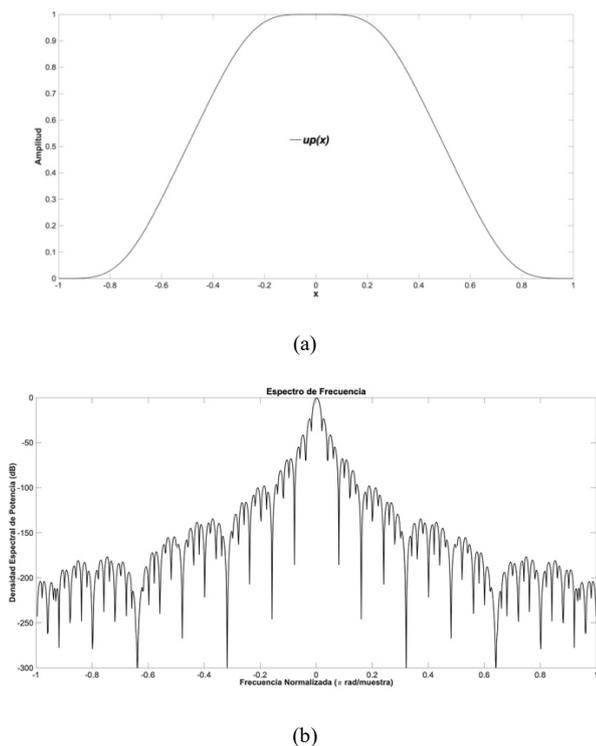


Figura 2: (a) Ventana AF $up(x)$ y (b) espectro de ventana AF $up(x)$

Para la creación de esta wavelet ortonormal basada en la $up(x)$, es necesario que cumpla la teoría del análisis de multiresolución (MRA); debemos de encontrar una función de escala $\phi(\omega)$ que forme una sucesión de subespacios cerrados $V_j \in L^2(R)$ con $(j \in \mathbb{Z})$, que cumplen las siguientes propiedades:

Propiedad 1

$$V_j \subset V_{j+1}, \forall j \in \mathbb{Z}$$

Propiedad 2

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{Z})$$

Propiedad 3

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$$

Propiedad 4

Existe $\phi(\omega) \in V_0$ tal que $\{\phi(\cdot - k): k \in \mathbb{Z}\}$ forman una base de Riesz para V_0 .

Propiedad 5

Para garantizar la base del sistema de Riesz la función

$$\sum_{n \in \mathbb{Z}} |\hat{\phi}(\omega + 2\pi n)|^2 = 1 \tag{4}$$

Con base a sumas de la función $up(x)$ desplazada horizontalmente y con un factor de dilatación $\frac{3}{2\pi}$, se obtiene $up^{sumn}(t)$

$$up^{sumn}(t) = up\left(\frac{3}{2\pi}t + 1\right) + up\left(\frac{3}{2\pi}t\right) + up\left(\frac{3}{2\pi}t - 1\right)$$

$$t \in \left[-\frac{4\pi}{3}, \frac{4\pi}{3}\right] \tag{5}$$

Se define el espectro de la función de escala $\hat{\phi}(t) = up^{sumn}(t)$, realizando un cambio de variable en (5) de t por ω , obteniendo la raíz cuadrada que cumple con (4)

$$\hat{\phi}(\omega) = \sqrt{up^{sumn}(\omega)}$$

$$\hat{\phi}(\omega) = \sqrt{up\left(\frac{3}{2\pi}\omega + 1\right) + up\left(\frac{3}{2\pi}\omega\right) + up\left(\frac{3}{2\pi}\omega - 1\right)}$$

$$\omega \in \left[-\frac{4\pi}{3}, \frac{4\pi}{3}\right] \tag{6}$$

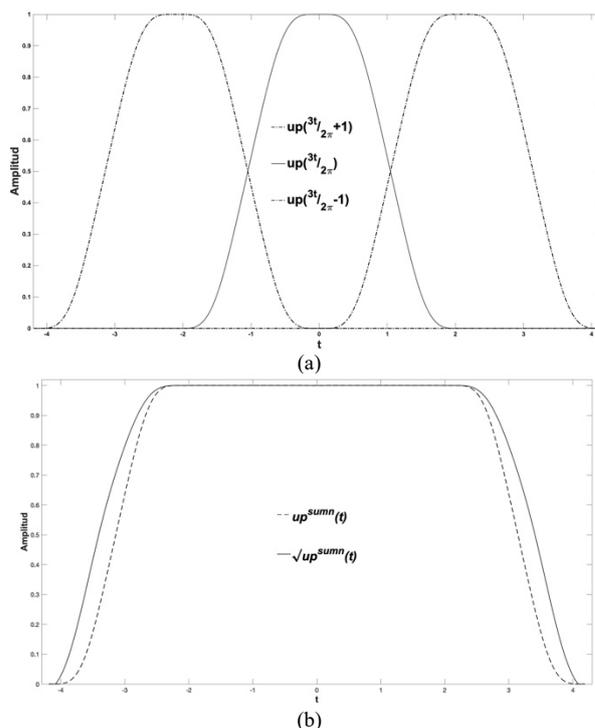


Figura 3: (a) Desplazamiento y dilataciones de $up(x)$
(b) $up^{sumn}(t)$ & $\sqrt{up^{sumn}(t)}$

Calculando la transformada inversa de Fourier de $\hat{\varphi}(\omega)$ se obtiene la Función de Escala

$$\varphi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\varphi}(\omega) e^{j\omega x} d\omega; \quad x \in [-8, 8] \quad (7)$$

La obtención de wavelet está representada al calcular la transformada inversa de Fourier de:

$$\Psi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega/2} \hat{\varphi}\left(\frac{\omega}{2}\right) (\hat{\varphi} + \hat{\varphi}(\omega + 2\pi)) e^{j\omega x} d\omega; \quad x \in [-8, 8] \quad (8)$$

$$\varphi(x) = \sum_{k=1}^{N-1} c_k \phi(2x - k) \leftrightarrow \Psi(\omega) = \frac{1}{2} \sum_k c_k e^{-j\omega k/2} \psi\left(\frac{\omega}{2}\right) \quad (9)$$

$$\psi(\omega) = e^{j\omega/2} \hat{\varphi}\left(\frac{\omega}{2}\right) (\hat{\varphi}(\omega - 2\pi) + \hat{\varphi}(\omega + 2\pi)) \quad (10)$$

Los coeficientes de descomposición del filtro h_n son obtenidos desarrollando $H_o(\omega)$ en una serie de Fourier

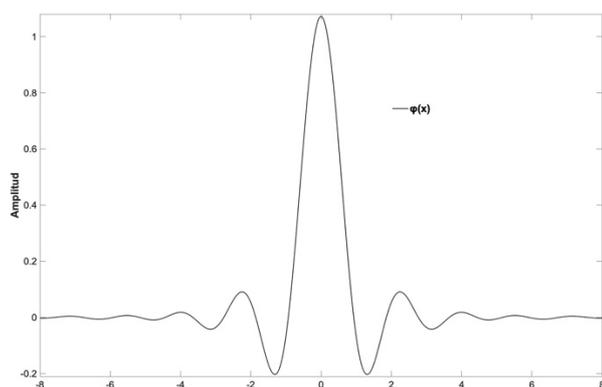
$$h_n = \frac{\sqrt{2}}{2\pi} \int_{-\pi}^{\pi} H_o(\omega) e^{jn\omega} d\omega \quad (11)$$

$$H_o(\omega) = \begin{cases} \hat{\varphi}(2\omega) & \omega \in \left[-\frac{2\pi}{3}, \frac{2\pi}{3}\right] \\ 0 & \text{otro caso} \end{cases} \quad (13)$$

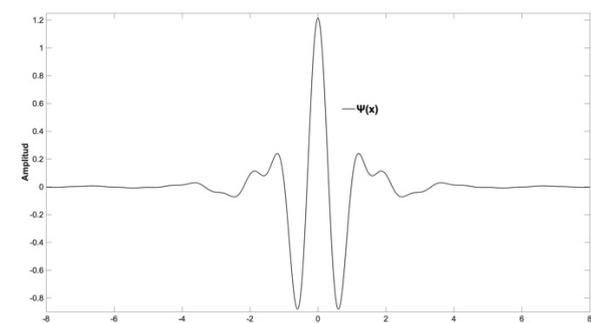
Para lograr trabajar con la wavelet $up(x)$ se utilizaron (13) y (11), para obtener 20 coeficientes que son mostrados en Tabla 1, en lugar de los presentados en (Kravchenko V. F., 2009).

Tabla 1: Coeficientes utilizados de la wavelet $up(x)$

No.	Coficiente	No.	Coficiente
1	-0.002828221	11	0.534450174
2	0.003514429	12	0.309840492
3	0.008659228	13	-0.033372304
4	-0.012513324	14	-0.085152986
5	-0.017078256	15	0.026227312
6	0.034064434	16	0.034064434
7	0.026227312	17	-0.017078256
8	-0.085152986	18	-0.012513324
9	-0.033372304	19	0.008659228
10	0.309840492	20	0.003514429



(a)



(b)

Figura 3: a) Función de escala $\varphi(x)$ y b) Wavelet $\Psi(x)$

Para lograr los términos de MRA hay que cumplir con la propiedad de

$$\hat{\varphi}(\omega) = H_0\left(\frac{\omega}{2}\right) \hat{\varphi}\left(\frac{\omega}{2}\right) \Leftrightarrow \hat{\varphi}(\omega) = \prod_{k=1}^{\infty} H_0\left(\frac{\omega}{2^k}\right) \quad (12)$$

La cual se logra con:

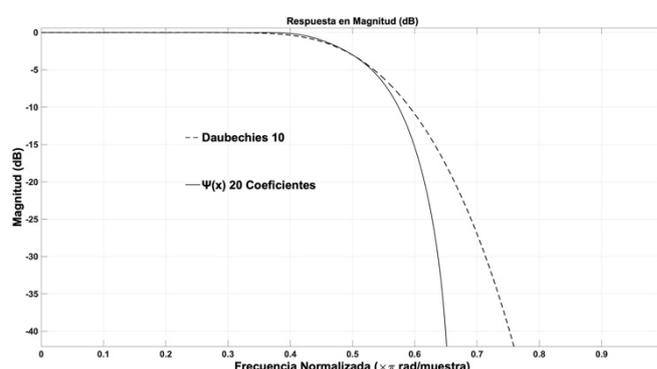


Figura 4: Espectro de frecuencia de Db10 vs Wavelet $up(x)$.

Comparando el espectro de frecuencia de la wavelet $up(x)$ con una de las wavelets clásicas como Daubechies de decimo orden (Db10) es posible visualizar que la wavelet $up(x)$ tiene una mejor respuesta en frecuencia, ver la Figura 4.

Obteniendo los coeficientes de la wavelet $up(x)$, tabla 1, se procede con la obtención de los vectores característicos faciales aplicando dichos coeficientes.

2.4. Clasificación.

Después de obtener los vectores característicos en el módulo anterior, se lleva a cabo la clasificación de los datos para después hacer la identificación/verificación de la identidad de la persona bajo análisis, empleando una Máquina de Soporte Vectorial (SVM) (Yu X., 2011). Este clasificador aprende la superficie de decisión de dos clases diferentes de puntos de entrada. Mapeando los datos mediante un Kernel polinomial del tipo $(gamma * u * v + coef0)^{degreec}$, donde $coef0=1$, $gamma=1$, $cost=100$. Donde se busca la máxima separación entre clases. Esta función de frontera, cuando se lleva de nuevo

al espacio de entrada, puede separar los datos en todas las clases diferentes, formando cada una de ellas un clúster. Para este caso se implementan los mismos parámetros del programa LIBSVM library en Matlab, utilizando los valores de “default” sugeridos por los autores (Chang, 2011).

2.5. Toma de decisión.

Una vez que se extraen las características faciales de las imágenes estéreo que pertenecen a las personas bajo análisis, se combina o bien se fusiona la información implementando tres niveles de fusión diferentes. Los niveles de fusión utilizados, son: fusión a nivel sensor, fusión a nivel características y fusión a nivel decisión. Para cada uno de estos, se obtiene una forma de distinta de tratar y analizar la información, por lo que la toma de decisión depende de que cada uno de estos casos. A continuación, se describe cada uno de los niveles de fusión implementados.

2.5.1 Fusión a nivel sensor.

Para llevar a cabo el análisis del sistema con la fusión a nivel sensor, este realiza la combinación del par de imágenes faciales que componen la imagen estéreo, donde principalmente se recorta por mitades el rostro para después hacer la combinación de estas con solo las mitades externas del rostro, obteniendo como resultado una sola imagen, ver Figura 5. Después de hacer la combinación de las imágenes en la entrada de los datos, se procede a la caracterización de estas con la wavelet $up(x)$ y por último se analizan con el uso de la SVM, para de esta manera determinar que muestras son reconocidas o no (García E., 2014).

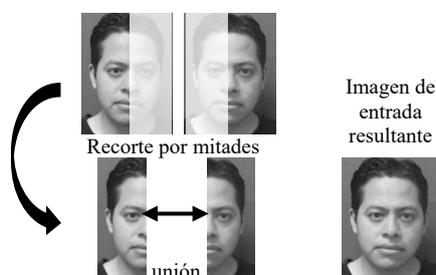


Figura 5: Fusión del par de imágenes estéreo.

2.5.2 Fusión a nivel características.

El siguiente modo de analizar la información es mediante la fusión a nivel características. En la cual se realizan dos procedimientos. Una de estas formas es después de la entrada de los datos y la otra forma es después de someterlas al pre-procesamiento, donde, viene la parte de la extracción de características, implementando la wavelet $up(x)$ para cada par de imágenes que conforman la imagen facial estéreo; es entonces, que se procede con las formas de combinación de información, ver Figura 6. Una de estas formas, se trata del cálculo de la media aritmética de los vectores característicos que conforman la imagen facial estéreo, después de esto se somete a la clasificación con la SVM y por último se toma la decisión. La segunda forma de combinación de información, es que, al contar con los vectores característicos del par de imágenes estéreo, estos se concatenan de tal manera que se obtiene como resultado un vector con el doble de longitud; el

cual este es usado para el entrenamiento del sistema y por último realizar el análisis en la toma de decisión de reconocimiento e identificación de rostros (García E., 2014).

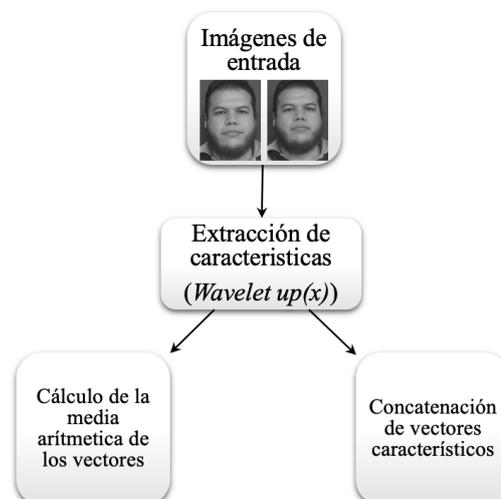


Figura 6: Fusión a nivel características.

2.5.3 Fusión a nivel decisión.

Para el análisis de esta última forma de fusión, se realizan todos los procesos de forma independiente para cada par de imágenes que componen la imagen estéreo, donde al final en el módulo de toma de decisión se realiza la combinación de información. Es decir, desde la entrada de los datos se hace el procedimiento de forma independiente para las dos componentes de la imagen estéreo, realizando el pre-procesamiento, la extracción de características y clasificación de los datos, donde al llegar a la toma de decisión, el sistema determina si la identidad es positiva solamente si ambas muestras determinan que la persona bajo análisis es quien dice ser (García E., 2014).

3. Evaluación de resultados

Para la evaluación de resultados del sistema, fue utilizada una base de datos con imágenes de 60 personas distintas con 15 pares de imágenes estéreo por cada persona, de las cuales 10 imágenes se usaron para el entrenamiento del sistema usando la SVM y las otras 5 imágenes restantes para la realización de pruebas de identificación y verificación de los rostros (García E., 2014). Así, el número total de imágenes contenidas en esta base de datos es de 1800 imágenes, capturadas en un ambiente con iluminación y fondo controlado, ver Figura 7.



Figura 7: Ejemplos de imágenes estéreo contenidas en la base de datos.

A continuación, se presentan los resultados obtenidos, después de analizar el comportamiento de cada uno de los tipos de niveles de fusión propuestos, los cuales ayudan a desempeñar un sistema multi-biométrico. Para todos los casos de niveles de fusión, los vectores característicos fueron obtenidos mediante la función wavelet $up(x)$.

Tabla 2: Porcentajes de identificación.

Métodos de extracción de características	Fusión a nivel sensor	Fusión a nivel características		Fusión a nivel decisión
		Media aritmética	Concatenación	
Wavelet $up(x)$	100%	100%	100%	100%
Filtros de Gabor	99.50%	99.55%	99.82%	99.50%
Daubechies de orden 3	Nivel 4	99.77%	99.66%	99.66%
	Nivel 5	99.88%	99.88%	99.55%
	Nivel 6	99.77%	99.55%	99.88%

En comparación con (García E., 2014), los resultados con la función wavelet $up(x)$ muestran que las cuatro formas de combinar la información en los distintos niveles de fusión, se logra alcanzar el total por ciento de identificación, consiguiendo así, la optimización del sistema de reconocimiento, ver Tabla 2.

4. Conclusiones

El propósito de este trabajo es analizar un sistema multi-biométrico facial usando imágenes estéreo, enfocándose en el módulo de extracción de características, incorporando el uso de la wavelet $up(x)$ con funciones atómicas para la caracterización de los rostros. Dando resultados visibles a la mejora del sistema con respecto a otros métodos clásicos de extracción, como: filtros de Gabor y Daubechies. Esta función proporciona resultados destacados; debido a que los coeficientes de esta wavelet tienen una mejor respuesta en tiempo-frecuencia, otorgando así una mejor extracción de

características dando un mayor porcentaje de reconocimiento dentro del sistema.

5. Agradecimientos

Agradecemos a las instituciones que pertenecen a cada uno de los autores de este trabajo, por el apoyo otorgado para la realización de este.

Referencias

- Chang, C. y. (2011). *LIBSVM: A Library for Support Vector Machines*. ACM T Intel Syst Tec.
- Hernandez-Matamoros A, H. F.-H.-M.-M. (2020). Recognition of ECG signals using wavelet based on atomic functions. *Biocybernetics and Biomedical Engineering*, 803-814.
- Anil K, J. P. (2008). *Handbook of Biometrics*. Boston, USA: Springer Science+Business Media.
- García E., E. E.-M. (2014). Sistema de Reconocimiento de Rostros usando Visión Estéreo. *Información Tecnológica SciELO*, 25(6), 117 - 130.
- García-Rios E., E.-H. E.-T.-S.-M.-M. (2013). Multi-biometric Face Recognition System using Levels of Fusion. *NAUN International Journal of Computers*, 7(3), 99-108.
- Perez-Daniel K, E. E.-M.-M. (2012). A novel image interest points detector based on Harris method using Atomic functions. *Telecommunications and Radio Engineering*, 71(17), 1575-1588.
- Kravchenko VF, H. M. (2009). *Adaptive digital processing of multidimensional signals with applications*. (V. F. Kravchenko, Ed.) Russia: M.:FIZMATLIT.
- Kravchenko, V. F. (2009). New system of Kravchenko orthogonal wavelets $\{up(t)\} \{up(t)\sim\}$. *Dokl. Math*, 80, 775-780.
- Kravchenko, V. K. (2015). Atomic and WA-systems and R-functions applied in modern radio physics problems: Part II. *J. Commun Technol. Electron*, 60, 103 - 141.
- Yu X., W. X. (2011). Support Vector Machine Regression for Reactivity Parameters of Vinyl monomers. *Journal of the Chilean Chemical Society*, 56(3), 746-751.