



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

**CARRERA DE INGENIERÍA EN ELECTRÓNICA Y
TELECOMUNICACIONES**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL
TÍTULO DE INGENIERO EN ELECTRÓNICA Y
TELECOMUNICACIONES**

**TEMA: DISEÑO, IMPLEMENTACIÓN Y ANÁLISIS DE UN
SISTEMA BIOMÉTRICO VASCULAR.**

AUTOR: IZURIETA SALAZAR, SANTIAGO DAVID

DIRECTOR: ING. CARRERA ERAZO, ENRIQUE VINICIO PhD

SANGOLQUÍ

2017

CERTIFICACIÓN



DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA
CARRERA DE ELECTRÓNICA Y TELECOMUNICACIONES

CERTIFICACIÓN

Certifico que el trabajo de titulación, **“DISEÑO, IMPLEMENTACIÓN Y ANÁLISIS DE UN SISTEMA BIOMÉTRICO VASCULAR”** realizado por el señor **SANTIAGO DAVID IZURIETA SALAZAR**, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar al señor **SANTIAGO DAVID IZURIETA SALAZAR** para que lo sustente públicamente.

Sangolquí, 03 de Marzo de 2017



Ing. Enrique V. Carrera PhD

DIRECTOR

AUTORÍA DE RESPONSABILIDAD



DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA
CARRERA DE ELECTRÓNICA Y TELECOMUNICACIONES

AUTORÍA DE RESPONSABILIDAD

Yo, **SANTIAGO DAVID IZURIETA SALAZAR**, con cédula de identidad N° 1718812066 declaro que este trabajo de titulación “**DISEÑO, IMPLEMENTACIÓN Y ANÁLISIS DE UN SISTEMA BIOMÉTRICO VASCULAR**” ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, 03 de Marzo de 2017

SANTIAGO DAVID IZURIETA SALAZAR

C.C.: 1718812066

AUTORIZACIÓN



DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA
CARRERA DE ELECTRÓNICA Y TELECOMUNICACIONES

AUTORIZACIÓN

Yo, **SANTIAGO DAVID IZURIETA SALAZAR**, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación “**DISEÑO, IMPLEMENTACIÓN Y ANÁLISIS DE UN SISTEMA BIOMÉTRICO VASCULAR**” cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Sangolquí, 03 de Marzo de 2017

SANTIAGO DAVID IZURIETA SALAZAR

C.C.: 1718812066

DEDICATORIA

A Dios por darme la vida y las fuerzas necesarias para alcanzar esta meta.

A mis padres Ruperto Izurieta y Elvia Salazar, por dejarme la mejor de las herencias que un ser humano puede recibir, el estudio.

A mi hermano Diego Izurieta, por ser mi mejor amigo, hermano y ejemplo en esta vida.

A mi tía y mi prima Margoth y Mónica Izurieta, por su incondicional y tan invaluable apoyo.

A mis amigos de siempre, por sus ánimos y abrazos en todo momento.

Dios les bendiga.

Santiago David Izurieta Salazar

AGRADECIMIENTO

Agradezco a Dios, por dejarme saborear su amor y fidelidad en absolutamente todos los momentos de mi vida.

Agradezco a mis padres Ruperto Izurieta y Elvia Salazar, por su tan valioso ejemplo, por sus consejos con y sin palabras pero sobre todo por sus infinitas oraciones pidiendo por mí. Gracias papi y mami, sin ustedes no lo hubiese logrado.

Agradezco a mi hermano Diego Izurieta, por todos los consejos y palabras de ánimo que siempre han llegado. ¿Seguimos soñando?, siempre SI. Eres extraordinario Body.

Agradezco a mi tía y a mi prima Margoth y Mónica Izurieta, porque han sido un pilar fundamental en mi vida. Gracias por mostrarme su amor y respaldo en todo momento.

Agradezco a mis amigos y amigas. Gracias por estar siempre en las buenas y en las malas. Sin sus ánimos y abrazos nada de esto hubiese sido realidad.

Agradezco a mi Director, PhD. Vinicio Carrera, por aconsejarme, motivarme y guiarme con sus conocimientos para la finalización de este trabajo de titulación.

Santiago David Izurieta Salazar

ÍNDICE

CERTIFICACIÓN.....	ii
AUTORÍA DE RESPONSABILIDAD.....	iii
AUTORIZACIÓN.....	iv
DEDICATORIA.....	v
AGRADECIMIENTO.....	vi
ÍNDICE.....	vii
ÍNDICE DE TABLAS.....	ix
ÍNDICE DE FIGURAS.....	x
RESUMEN.....	xi
ABSTRACT.....	xii
CAPÍTULO I.....	1
ASPECTOS GENERALES.....	1
1.1. Antecedentes.....	1
1.2. Justificación e Importancia.....	2
1.3. Alcance del Proyecto.....	4
1.4. Objetivos.....	6
1.4.1. General.....	6
1.4.2. Específicos.....	6
CAPÍTULO II.....	7
MARCO TEÓRICO.....	7
2.1. Biometría: Concepto y Origen.....	7
2.2. Sistema Biométrico Vascular.....	7
2.2.1. Reconocimiento de patrones vasculares.....	7
2.2.2. Modos de funcionamiento de los sistemas biométricos.....	10
2.3. Transformada Wavelet Discreta en dos dimensiones (2D-WTD).....	12

2.4. Extracción de características de textura.....	14
2.4.1. Matriz de co-ocurrencia.....	14
2.5. Procedimientos de clasificación.....	20
2.5.1. Análisis discriminante lineal	20
2.5.2. Análisis discriminante cuadrático.....	21
2.5.3. Método del vecino más cercano	23
CAPÍTULO III	25
DESCRIPCIÓN DEL SISTEMA DE BIOMETRÍA VASCULAR.....	25
3.1. Módulo de adquisición de imágenes.....	26
3.2. Módulo de pre procesamiento y extracción de características	26
3.3. Módulo de clasificación.....	35
CAPÍTULO IV	38
PRESENTACIÓN Y ANÁLISIS DE RESULTADOS.....	38
4.1. Presentación de los resultados obtenidos.....	38
4.1.1. Resultados utilizando el clasificador discriminante lineal	39
4.1.2. Resultados utilizando el clasificador discriminante cuadrático.....	40
4.1.3. Resultados utilizando el clasificador <i>k-nearest neighbor</i> (k-NN).....	41
4.2. Análisis de los resultados obtenidos	43
CAPÍTULO V	46
CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS.....	46
5.1. Conclusiones	46
5.2. Recomendaciones	47
5.3. Trabajos futuros	48
REFERENCIAS	49

ÍNDICE DE TABLAS

Tabla 1. Acontecimientos relevantes en la historia de la Biometría.....	9
Tabla 2. Combinaciones de los 4 niveles de gris de la figura 7.....	15
Tabla 3. Matriz de co-ocurrencia de la relación espacial (1,0) para la figura 7.....	16
Tabla 4. Matriz de co-ocurrencia simétrica de la relación espacial (1,0) para la figura 7.....	16
Tabla 5. Matriz de co-ocurrencia normalizada de la relación espacial (1,0) para la figura 7.....	17
Tabla 6. Detalle de características para cada nivel evaluado.....	34
Tabla 7. Resultados del clasificador discriminante lineal y validación A.....	39
Tabla 8. Resultados del clasificador discriminante lineal y validación B.....	40
Tabla 9. Resultados del clasificador discriminante cuadrático y validación A.....	40
Tabla 10. Resultados del clasificador discriminante cuadrático y validación B.....	41
Tabla 11. Evaluación de desempeño del clasificador k-NN en función de k, con el método de validación A.....	42
Tabla 12. Resultados del clasificador k-NN y validación A.....	43
Tabla 13. Resultados del clasificador k-NN y validación B. Imagen 1.....	43

ÍNDICE DE FIGURAS

Figura 1. Diagrama de bloques general de sistemas biométricos.	4
Figura 2. Nivel de absorción espectral de la hemoglobina desoxigenada.....	8
Figura 3. Diagrama de bloques del proceso de identificación biométrica.....	11
Figura 4. Diagrama de bloques del proceso de verificación biométrica.....	12
Figura 5. Descomposición de un nivel de la 2D-DWT.....	13
Figura 6. Primer nivel de descomposición de una imagen con la 2D – DWT con sus cuatro componentes (LL, LH, HL, HH).....	13
Figura 7. Imagen de ejemplo de dimensiones 4x4 pixeles con 4 valores de niveles de grises.....	15
Figura 8. Estructura funcional de un clasificador basado en funciones discriminantes.....	23
Figura 9. Ejemplo de clasificador k-NN para dos clases.....	24
Figura 10. Diagrama de flujo del sistema biométrico vascular.....	25
Figura 11. Sistema de adquisición de imágenes para un.....	26
Figura 12. (a) Muestra vascular del dedo índice de la mano izquierda.....	27
Figura 13. Wavelet spline biortogonal. Función para descomposición (lado izquierdo) y función para reconstrucción (lado derecho).....	29
Figura 14. Wavelet haar. Función para descomposición y reconstrucción.....	29
Figura 15. Diagrama de flujo de la función pre-procesamiento.....	30
Figura 16. Diagrama de flujo de la función extracción de características.....	31
Figura 17. Diagrama de flujo de la función Clasificación.....	37
Figura 18. Matriz de confusión realizando la validación A.....	38
Figura 19. Matriz de confusión realizando la validación B.....	39
Figura 20. Desempeño del clasificador k-NN en función del valor de k	41

RESUMEN

Biometría o autenticación biométrica se refiere a técnicas automatizadas de reconocimiento de un individuo a través de sus características de comportamiento o fisiológicas. Actualmente el campo del análisis de rasgos biométricos ha alcanzado bastante importancia. La biometría vascular es un tipo de biometría que enfrenta la creciente demanda de soluciones biométricas para fortalecer la seguridad informática en sus diferentes aplicaciones. La biometría vascular es un tipo de biometría con grandes ventajas. La ventaja más importante es que el rasgo biométrico que procesa se encuentra por dentro de la piel teniendo un único patrón para cada individuo. Esto hace que el sistema sea bastante robusto ante la falsificación. En el presente trabajo se ha diseñado la etapa de software de un sistema biométrico vascular. El objetivo es analizar los parámetros que influyen en cada etapa y en el resultado final. Las muestras de cada individuo se las ha obtenido de una base de datos multimodal de libre acceso llamada SDUMLA-HMT. Con el fin de caracterizar a cada patrón analizado se ha utilizado características texturales para cada muestra. Se ha procesado la muestra a través de una *wavelet spline* y a través de una *wavelet haar* comparando entre si los resultados obtenidos. Finalmente para la etapa de clasificación se lo ha hecho implementando los clasificadores discriminante cuadrático, discriminante lineal y un clasificador k-NN con el objetivo de analizar el desempeño de cada uno.

PALABRAS CLAVE:

- ❖ **BIOMETRÍA VASCULAR**
- ❖ **WAVELET**
- ❖ **TEXTURA**
- ❖ **APRENDIZAJE DE MÁQUINA**

ABSTRACT

Biometrics or biometric authentication refers to automated techniques for recognizing an individual through their behavioral or physiological characteristics. Currently, the analysis of biometric features has become a quite important area. The vascular biometry is a type of biometry that faces the growing demand for biometric solutions to strengthen computer security in its different applications. Vascular biometry is a type of biometry with great advantages. The most important advantage is that the biometric trait is located inside the skin having a single pattern for each individual. This makes it a fairly robust system against counterfeiting. In this project the software stage of a vascular biometric system has been designed. The goal is to analyze the parameters that influence each stage and the final result. The samples of each individual were obtained from a freely accessible multimodal database called SDUMLA-HMT. In order to characterize each pattern, the textural characteristics have been used for each sample. The sample was processed through a spline wavelet and through a haar wavelet comparing the obtained results. Finally, the classification stage was done by implementing a quadratic classifier, a linear classifier and a k- NN classifier with the objective of analyzing the performance of each one.

KEYWORDS:

- ❖ **VASCULAR BIOMETRY**
- ❖ **WAVELET**
- ❖ **TEXTURE**
- ❖ **MACHINE LEARNING**

CAPÍTULO I

ASPECTOS GENERALES

1.1. Antecedentes

A lo largo de la historia la identificación personal se ha basado en la utilización de varios mecanismos como tarjetas inteligentes y números de identificación personal (PIN). Estos métodos ofrecen una seguridad muy limitada y son poco confiables debido a la facilidad de ser copiados o suplantados. La identificación biométrica permite minimizar estas desventajas utilizando ciertas características biológicas, que al ser únicas para cada ser humano, convierten la identificación del usuario en parte integral de la seguridad del sistema (Motato Toro & Loaiza Correa, 2009).

A lo largo del siglo XX la mayor parte de los países del mundo han utilizado la biometría como una técnica de identificación. Los sistemas biométricos de referencia inmediata basan su proceso de identificación en el reconocimiento de huellas dactilares, el iris del ojo y la firma, pero ante la necesidad creciente de mejorar los controles de acceso y seguridad, estas soluciones están siendo desplazadas por una nueva gama de sistemas de mayor desempeño que basan su proceso de autenticación en el reconocimiento de nuevos rasgos fisiológicos (Im, Park, & Kim, 2001).

Actualmente, la tecnología biométrica presenta algunos métodos de identificación de personas posicionándose en el mercado con varios productos, y orientando las principales investigaciones y aplicaciones con los patrones de: iris, huella dactilar, geometría de la mano, rostros y venas (Kono, Ueki, & Umemura, 2002). De todos estos, el patrón venoso es un sistema que ha llamado la atención al poseer grandes ventajas biométricas (Miura, Miyatake, & Nagasaka, 2004).

La biometría vascular al igual que el reconocimiento de retina, usa luz infrarroja (IR) a corta distancia para detectar los patrones de la red vascular (ISO/IEC, 2011). La distribución del sistema venoso posee muchas características que las hacen únicas y dado que se requiere que la sangre esté fluyendo para registrar la imagen, la suplantación de identidad es casi imposible (Hu, Wang, & Yang, 2014) (Badawi, 2007). El uso de luz infrarroja (IR) se basa en el principio de absorción de los principales compuestos de esta radiación en la sangre los cuales son la oxihemoglobina (HnO_2), des-oxihemoglobina (Hb) y agua (H_2O) (Hidalgo, Marcotti, & Mathé, 2012). La Norma ISO/IEC 19794-9 estandariza la extracción de patrones vasculares de la palma de la mano, reverso de la mano y dedo (ISO/IEC, 2011).

El reconocimiento de patrones vasculares es una de las técnicas biométricas más novedosas. Numerosos estudios han analizado su unicidad, demostrando que la organización vascular es única en cada individuo, no solo en la retina (Bolle & Pankati, 1998) o la mano (Fujitsu Laboratories Ltd., 2006) (Schaffer, 2005), sino en todas las venas y capilares que forman parte del cuerpo humano (Willmore, 1994). El reconocimiento mediante patrones vasculares presenta un futuro prometedor y son muchas las líneas de investigación abiertas en esta área.

1.2. Justificación e Importancia

Los sistemas de autenticación basados en biometría vascular, se enfocan principalmente en los patrones de la red vascular de la palma de la mano, de la cara dorsal de la mano o del dedo, como dato de identificación personal (Choi, Chung, Im, Kim, & Park, 2000). Este rasgo fisiológico es útil como indicador biométrico ya que satisface las propiedades de (Fujitsu Laboratories Ltd., 2006) (Leedham & Wang, 2006):

- ✓ Universalidad: todas las personas poseen el rasgo.
- ✓ Unicidad: estos patrones son únicos en cada individuo, incluso diferentes entre gemelos además de distintos en la mano derecha e izquierda.

- ✓ Permanencia: no cambian con el crecimiento simplemente se amplían manteniendo el patrón.
- ✓ Cuantificación: se puede describir, representar y clasificar a cada patrón usando técnicas de adquisición y procesamiento de imágenes infrarrojas.

La biometría vascular, en contraste con los demás tipos de biometría, presenta la necesidad de que exista fluidez de sangre en la red subcutánea vascular del individuo. En otras palabras, el sistema de biometría vascular necesita que la persona esté con vida para registrar la imagen infrarroja o patrón. El uso de imágenes infrarrojas permite ver un contraste entre los vasos sanguíneos y el resto de la piel, además de eliminar algunos rasgos no deseados de la superficie cutánea y del entorno (Shimizu, 1992). Otra ventaja que aportan los patrones vasculares principalmente deriva del hecho de que son una característica interna del cuerpo humano, esto hace que no estén expuestos a condiciones ambientales, haciéndoles más invariantes en el tiempo que otras características como las huellas dactilares, que por trabajos manuales o envejecimiento pueden verse alteradas con distinto grado de intensidad (Silveyra, 2006). El patrón vascular en base a ser una propiedad interna, presenta un nivel de complejidad superior ante la alteración o falsificación, dado que los vasos sanguíneos pertenecen al tejido subcutáneo y son en su mayoría invisibles al ojo humano (Fujitsu Laboratories Ltd., 2003).

Por otro lado, haciendo referencia a la etapa de adquisición de imágenes, el diseño de estos dispositivos permite que la captura se realice sin contacto en la mayor parte de casos, por lo que se la cataloga como biometría higienizada (Fujitsu Laboratories Ltd., 2003). Los usuarios no tocan la superficie que contiene el sensor, de forma que mejora la higiene y con ello la aceptación del usuario. Adicionalmente, la luz ambiental no supone un inconveniente en el caso de efectuar la toma de la muestra de las venas, lo que si ocurre al realizar la captura de otros rasgos como el iris o la cara (Alberto, Mateos, & Pizarro, 2005).

Los sistemas basados en esta técnica se posicionan como un referente en el avance tecnológico de soluciones de control de accesos y seguridad, convirtiéndolos en sistemas robustos y prácticamente inviolables (ISO/IEC, 2011). Cabe recalcar que

en el territorio ecuatoriano no se ha realizado una investigación sobre biometría vascular, por lo que este proyecto, sentará un precedente en la investigación a nivel nacional sobre esta nueva tecnología biométrica.

1.3. Alcance del Proyecto

En general los sistemas biométricos están formados por cuatro bloques o etapas de funcionamiento, que son: Etapa de adquisición de rasgos, etapa de pre procesamiento, etapa de extracción de características y finalmente la etapa de clasificación. En la figura 1 se puede apreciar la estructura de estas etapas.



Figura 1. Diagrama de bloques general de sistemas biométricos.

Haciendo referencia a los sistemas de biometría vascular, la etapa de adquisición de rasgos se la realiza a través de un sensor infrarrojo acompañado de una adecuada iluminación. Para que todo este proceso sea realizado con éxito y la imagen sea obtenida correctamente, depende de factores propios de los dispositivos además de factores externos. Por esta razón y con el fin de estandarizar las muestras, se trabajará con una base de datos pública llamada SDUMLA-HMT (Liu, Yin, & Sun, 2011), diseñada en la Universidad de Shandong por el grupo llamado *Group of Machine Learning and Applications*. Se hará uso de la sección de muestras vasculares la cual está conformada por 6 imágenes de cada uno de los dedos índice; medio y anular de la mano derecha e izquierda de 106 individuos, dando un total de 3,816 imágenes cada una en formato “*bmp*”. Se escogerá 636 imágenes que corresponden a las 6 muestras del dedo índice de la mano izquierda de los 106 individuos.

En la etapa de pre procesamiento se utilizará algoritmos de filtrado y operaciones morfológicas para lograr obtener y mejorar la región de interés (ROI) que contiene el patrón venoso.

Dentro del pre procesamiento se implementarán procesos comúnmente utilizados por los autores en (Motato Toro & Loaiza Correa, 2009) (Im, Park, & Kim, 2001) (Hu, Wang, & Yang, 2014) (Barrera, Beltrán, & Camacho, 2016). Se variarán los parámetros de los algoritmos con el fin de hacer una comparación de las mejores características dentro del sistema de biometría vascular.

La etapa de extracción de características enfoca sus esfuerzos para extraer el patrón venoso. Para ello, se usará descriptores de textura del patrón venoso. Con esto se logrará pasar la información observada en las imágenes a valores numéricos o simbólicos. Este formato permitirá que la comparación posterior para la identificación o verificación sea más sencilla y exacta. Esta información característica de cada patrón venoso será guardada en un vector característico. Para esto, se utilizará dos tipos de wavelets. Una de ellas será una *wavelet spline* y la otra una *wavelet haar*. Esto se realizará con el fin de comparar sus resultados.

Finalmente, en la etapa de clasificación de patrones vasculares se implementarán los algoritmos clasificadores: Discriminante lineal, cuadrático y k-vecinos más cercanos (kNN); basados en aprendizaje de máquina. Se busca hacer una comparación de los mismos para definir de mejor manera al sistema de biometría vascular.

Para esto, en la fase de aprendizaje se utilizará dos métodos de validación. En el primero se separará 5 muestras para entrenamiento y 1 para validación. En el segundo método de separará 4 muestras para entrenamiento y 2 para validación para evaluar el desempeño del sistema. Una vez aislada las propiedades características de cada clase, ésta se representará por un patrón de clase que consiste en la media de todos los patrones que pertenecen a la misma clase (Tocoma, 2015).

Todo el procesamiento de las muestras vasculares y los algoritmos antes mencionados, se implementarán en la herramienta de software Matlab.

1.4. Objetivos

1.4.1. General

- Analizar los diferentes parámetros que afectan el desempeño de un sistema de biometría vascular basado en imágenes del dedo índice de la mano izquierda.

1.4.2. Específicos

- Realizar el estudio del estado del arte de los sistemas biométricos vasculares.
- Diseñar un sistema biométrico vascular para una base de datos estandarizada.
- Implementar las etapas de pre procesamiento, extracción de características y clasificación del sistema biométrico vascular.
- Evaluar los parámetros que caracterizan el desempeño de un sistema biométrico para identificación de patrones vasculares.

CAPÍTULO II

MARCO TEÓRICO

2.1. Biometría: Concepto y Origen

Biometría, desde el punto de vista etimológico proviene del griego “βιο” (bio) que significa vida y “μετρία” (metría) que significa medida (E-Cultura Group, 2015).

La Biometría es una ciencia, rama de la Biología dedicada al estudio estadístico o mensurativo de las características fisiológicas o conductuales cuantitativas de los seres vivos. Es el conjunto de métodos automáticos los mismos que analizan atributos humanos puntuales para identificación y verificación de personas (Aglío Caballero, 2016). Varios han sido los acontecimientos que han marcado el origen y el desarrollo de la biometría. En la tabla 1 se detallan los más relevantes.

En el presente siglo XXI es la época en que el campo de la biometría está experimentando un crecimiento y desarrollo continuo. El área de la biometría es muy extenso contando con numerosas aplicaciones como: Biometría Comercial y Residencial, Biometría del Cuidado Sanitario, Biometría para la Seguridad, Fortificación Judicial y Legal, Biometría Financiera y Biometría Móvil; por mencionar algunas (Alberto, Mateos, & Pizarro, 2005) (Jain, Maio, Maltoni, & Wayman, 2006).

2.2. Sistema Biométrico Vascular

2.2.1. Reconocimiento de patrones vasculares

El reconocimiento de patrones vasculares (*Vascular Pattern Recognition* o *Vein Pattern Authentication*) es una de las técnicas más novedosas. Este tipo de sistemas biométricos se fundamenta en el reconocimiento de patrones implícitos en la red vascular de la mano o de los dedos. Este rasgo biométrico característico no puede ser fácilmente obtenido para ser falsificado ya que la disposición de la red vascular es

subcutánea (Carrasco García, Fuentes García, García Benítez, & López Hernández, 2016).

El reconocimiento del patrón vascular del dedo es un método de autenticación biométrica basado en imágenes infrarrojas de la disposición de las venas del dedo. El infrarrojo está dividido en tres regiones: infrarrojo cercano, mediano y lejano (Centro de Procesamiento y Análisis Infrarrojo (IPAC), 2001). El sistema de biometría vascular funciona en la región del infrarrojo cercano. En esta región existe un rango espectral que oscila entre los 700nm hasta los 900nm aproximadamente en que la luz incidente con longitud de onda dentro de ese rango puede penetrar profundamente en los tejidos (Motato Toro & Loaiza Correa, 2009). En la figura 2 se puede observar que el nivel de absorción espectral de la hemoglobina des-oxigenada presenta un máximo valor en la región cerca del espectro infrarrojo cercano (Fujitsu Laboratories Ltd., 2003).

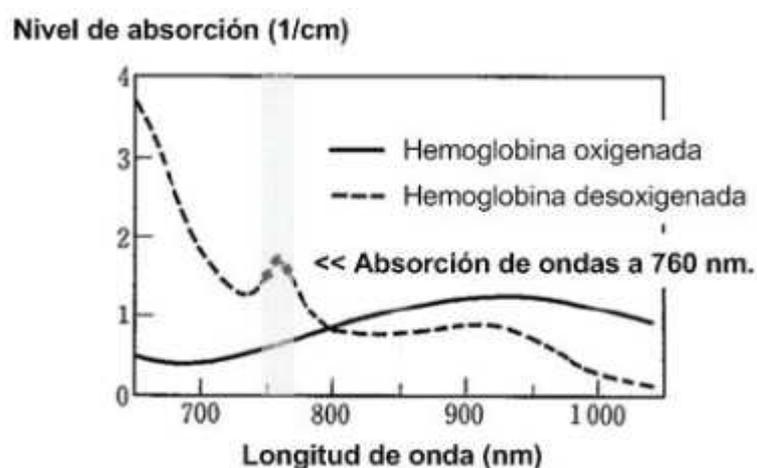


Figura 2. Nivel de absorción espectral de la hemoglobina desoxigenada
Fuente: (Motato Toro & Loaiza Correa, 2009)

El reconocimiento de las venas del dedo es una de las muchas formas de biometría utilizadas para identificar individuos y verificar su identidad. Tiene varias características importantes que lo diferencian de otras formas de biometría como un medio altamente seguro y conveniente de autenticación personal. Estas características son (Bansal & Kaur, 2015):

Tabla 1.
Acontecimientos relevantes en la historia de la Biometría.

ÉPOCA	ACONTECIMIENTO
Siglo VIII	Uso de huellas dactilares en China para caracterizar a un individuo tanto en documentos como en esculturas de arcilla.
Med. del Siglo XIV	China ya tenía aplicaciones biométricas. Sin embargo, no se pone en práctica hasta finales del siglo XIX.
Siglo XVII	Marcelo Malpigio realizó el primer estudio sistemático de huellas digitales.
Inicios del Siglo XX	La biometría gana terreno en el área de investigación gracias a la criminalística. Se realizaron estudios que relacionaban características físicas con tendencias criminales. Además, se ejecutaron estudios sobre el uso de rasgos biométricos para la detección de crímenes.
A lo largo del siglo XX	<p>1969: El FBI contrató al actualmente conocido como Instituto nacional de Estándares y Tecnología (NIST), con el fin de estudiar la automatización de la búsqueda, clasificación y concordancia de las huellas dactilares.</p> <p>1975: El FBI invirtió para el desarrollo de escáneres de huella dactilar. Esto condujo al desarrollo de un lector prototipo.</p> <p>1992: Se conforma el Consorcio Biométrico dentro del EEUU.</p> <p>1993: Da inicio el programa <i>Face Recognition Technology</i> (FERET).</p> <p>1994: Se patenta el primer algoritmo de reconocimiento de iris. En este algoritmo se fundamentan la mayoría de productos comerciales actuales.</p> <p>1996: En los Juegos Olímpicos se implanta la biometría de la mano para supervisar el acceso físico a la Villa Olímpica.</p> <p>2000: Se publica la primera investigación del uso de patrones vasculares en biometría.</p>
Siglo XXI	<p>2002: Creación del comité de estándares ISO/IEC en biometría.</p> <p>2010: Uso del reconocimiento de voz para identificación terrorista por parte de los servicios de inteligencia en Europa.</p> <p>2013: Apple incluye escáneres de huellas dactilares en dispositivos móviles.</p>

- ✓ Resistente a la manipulación criminal: Debido a que las venas están ubicadas de forma subcutánea, hay poco riesgo de falsificación o robo.
- ✓ Único y constante: Los patrones de la vena del dedo son diferentes, incluso entre gemelos idénticos y permanecen constantes a través de los años adultos.
- ✓ Sin contacto: El uso de la luz infrarroja cercana permite generar imágenes no invasivas y sin contacto que garantizan comodidad y limpieza para la experiencia del usuario. Por esta razón, a la biometría vascular se la califica como biometría higienizada.
- ✓ Facilidad de extracción de características: Los patrones de la vena del dedo son relativamente estables y claramente captables, permitiendo el uso de las cámaras de baja resolución para capturar imágenes de la vena en tamaño pequeño.
- ✓ Velocidad de autenticación rápida: La autenticación de los patrones vasculares en modo de verificación, es decir una comparación 1:1, tarda menos de un segundo.
- ✓ El usuario debe estar con vida: La cámara infrarroja puede capturar las imágenes de la vena del dedo sólo si la hemoglobina desoxigenada está presente en la red vascular del cuerpo y la imagen sólo puede ser tomada si el individuo está con vida.

Entre todas las técnicas biométricas, la tecnología de reconocimiento vascular en el dedo reduce considerablemente las preocupaciones relacionadas con la privacidad de la identidad e información del usuario y la seguridad de la misma. Actualmente, el interés sobre el uso de patrones vasculares del dedo en los sistemas biométricos ha crecido exponencialmente, aumentando el número de investigaciones y llegando a comercializarse sistemas basados en las venas de los dedos y de la mano (Fujitsu Laboratories Ltd., 2006).

2.2.2. Modos de funcionamiento de los sistemas biométricos

Desde la perspectiva del funcionamiento de los sistemas biométricos de reconocimiento de personas es necesario clasificarlos en dos modos:

- ✓ Sistemas de reconocimiento en modo de identificación

- ✓ Sistemas de reconocimiento en modo de verificación.

2.2.2.1. Sistemas de reconocimiento en modo de identificación

Estos sistemas tienen como objetivo, clasificar un rasgo biométrico de identidad desconocida como perteneciente a uno de entre un conjunto de N posibles rasgos biométricos que caracterizan a un determinado individuo. En el proceso de identificación, el usuario no le informa al sistema biométrico cuál es su identidad. El sistema realiza la captura del rasgo biométrico del usuario y lo procesa para crear el patrón. Posteriormente, el sistema biométrico compara el patrón con un conjunto de patrones de referencia para determinar la identidad de la persona. En la figura 3 se puede apreciar el diagrama de bloques del funcionamiento de los sistemas de reconocimiento en modo identificación.

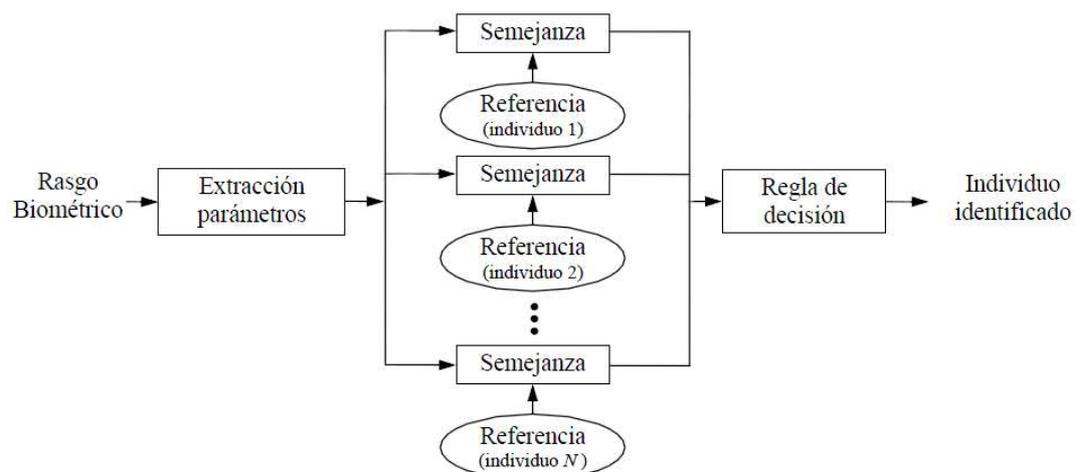


Figura 3. Diagrama de bloques del proceso de identificación biométrica
Fuente: (Ortega, 2015)

2.2.2.2. Sistemas de reconocimiento en modo de verificación

Estos sistemas tienen como objetivo, catalogar al usuario a prueba como usuario auténtico o como impostor. Para esto los sistemas de verificación toman dos entradas. La primera es la toma del rasgo biométrico y la segunda es una solicitud de identidad. Por lo tanto, las únicas dos salidas del sistema serán la decisión de aceptación o rechazo del usuario. En el proceso de verificación, la persona le informa al sistema cuál es su identidad. El sistema verifica la identidad del usuario

comparando el rasgo biométrico capturado con su propio rasgo biométrico previamente almacenado. El sistema biométrico dará como resultado una validación positiva o una validación negativa. En la figura 4 se observa el diagrama de bloques del funcionamiento de los sistemas de reconocimiento en modo de verificación.

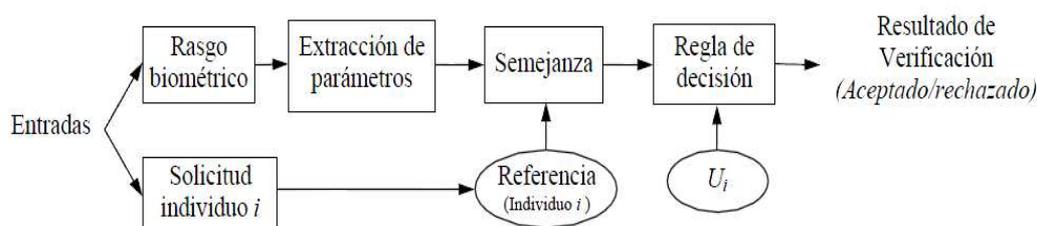


Figura 4. Diagrama de bloques del proceso de verificación biométrica
Fuente: (Ortega, 2015)

Según la aplicación del sistema biométrico a diseñar, se define el modo de funcionamiento en el que debe trabajar el sistema. Para este proyecto de investigación se desea identificar la identidad de los usuarios muestreados. Por lo tanto el funcionamiento del sistema es en modo de identificación, es decir, se realiza la comparación entre el patrón vascular de un usuario y un conjunto de patrones de varios usuarios previamente registrados en la base de datos con el fin de identificar su identidad.

2.3. Transformada Wavelet Discreta en dos dimensiones (2D-WTD)

La transformada wavelet discreta en dos dimensiones analiza por separado las filas y las columnas. Esta transformada considera a cada fila y a cada columna como una señal uni-dimensional. Una vez separadas las filas y las columnas, se aplica los filtros pasa-altas (H) y pasa-bajas (L) a cada fila y a cada columna (Gonzalez, 2015). En la figura 5 se muestra el desarrollo de este proceso. Para cada nivel de descomposición, se generan cuatro nuevas sub-matrices. Estas sub-matrices son (Albertos Cabañas, 2011):

- ✓ Matriz de aproximación LL.
- ✓ Matriz de detalles horizontales de la señal original LH.

- ✓ Matriz de detalles verticales de la señal original HL.
- ✓ Matriz de detalles diagonales de la señal original HH.

La DWT es un proceso reversible y para realizar la reconstrucción simplemente se re combina las cuatro sub-imágenes (LL, LH, HL y HH) utilizando un proceso de undecimado. En la figura 5, el símbolo $\downarrow 2$ representa el proceso de decimación. En la figura 6 se muestra un ejemplo de cómo un análisis en el primer nivel de descomposición genera las cuatro sub-imágenes.

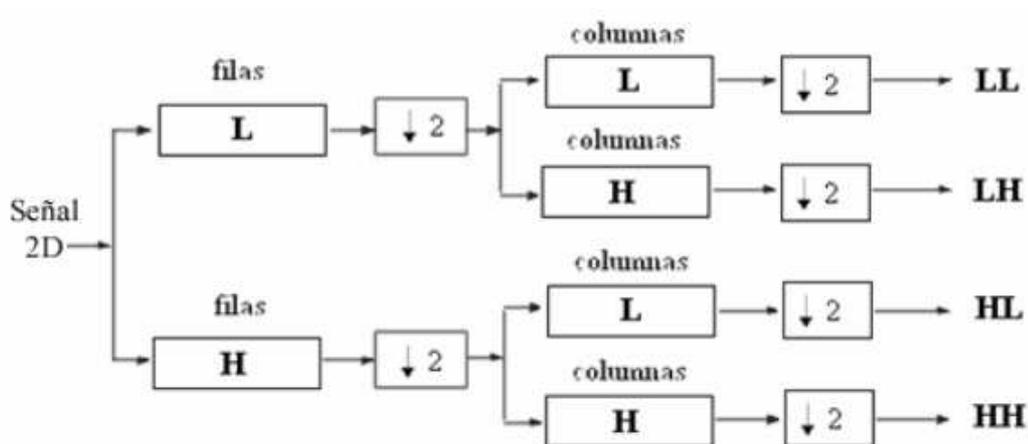


Figura 5. Descomposición de un nivel de la 2D-DWT
Fuente: (Gonzalez, 2015)

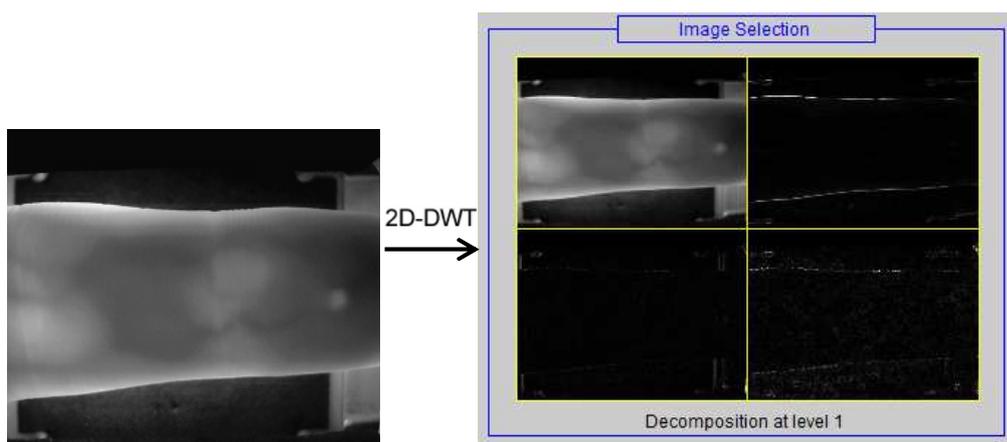


Figura 6. Primer nivel de descomposición de una imagen con la 2D – DWT con sus cuatro componentes (LL, LH, HL, HH)

La DWT no pierde la información de tiempo-frecuencia. La resolución del análisis depende del nivel de descomposición en que uno se encuentre, por lo que se

puede realizar un análisis en diferentes resoluciones (Análisis Multi-Resolución o MRA) (Mallat, 1989) (Gonzalez, 2015).

El uso de la transformada wavelet es una poderosa herramienta que tiene varios campos de aplicación además del análisis local de señales no estacionarias, el análisis de señales electrocardiográficas, de sonido, sísmicas, de radar, además de ser utilizada para la compresión y procesamiento de imágenes y reconocimiento de patrones (UNICEN).

2.4. Extracción de características de textura

Dentro de la identificación de objetos o regiones de interés en una imagen, la textura es una característica relevante. La textura de una imagen es una cuantificación de la variación espacial de valores de tono que es imposible definirlo precisamente por su carácter sensorial. La matriz de co-ocurrencia de niveles de grises o GLCM (*grey level co-occurrence matrix*), basadas en estadísticos de segundo orden es el método comúnmente utilizado para calcular matemáticamente la textura (Presutti, 2004).

2.4.1. Matriz de co-ocurrencia

Las medidas texturales de una imagen están contenidas en la relación espacial que los tonos de gris tienen entre ellos. Esas relaciones están especificadas en la matriz de co-ocurrencia espacial que son procesadas en una dirección específica o para todas: norte, sur, este, oeste y las 4 diagonales (Jiménez Guerrero, 2016). En lugar de contar estas direcciones por separado, hay formas más sencillas de realizarlo utilizando la matriz simétrica. Por lo tanto, para el cálculo se escoge las 4 direcciones 0° , 45° , 90° y 135° ; y se obliga a la matriz a ser simétrica para calcular las direcciones restantes (Dinstein, Haralick, & Shanmugan, 1973). Supongamos una imagen de ejemplo que se muestra en la figura 7. Esta imagen consta de 4 píxeles de lado y 4 niveles de grises: 0, 1, 2 y 3 (Presutti, 2004).

La matriz de referencia considera la relación espacial entre el píxel de referencia y el píxel vecino. Por ejemplo, si se elige el píxel vecino que está situado un píxel a

la derecha de cada pixel de referencia (0°), esto se expresa como (1,0): 1 pixel en la dirección x, 0 pixel en la dirección y. Para representar se lo hace de la siguiente manera:

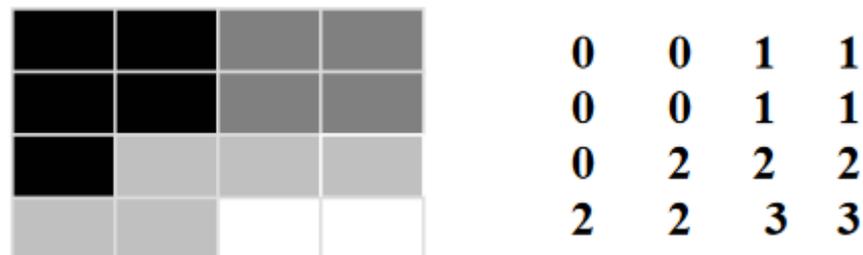


Figura 7. Imagen de ejemplo de dimensiones 4x4 pixeles con 4 valores de niveles de grises.

Fuente: (Presutti, 2004)

- ✓ (1,0) orientación horizontal, 0° .
- ✓ (0,-1) orientación vertical, 90° .
- ✓ (1,1) orientación diagonal hacia arriba, 45° .
- ✓ (1,-1) orientación diagonal hacia abajo, 135° .

Las posibles combinaciones se muestran en la siguiente tabla:

Tabla 2.

Combinaciones de los 4 niveles de gris de la figura 7.

Pixel vecino (columnas)	0	1	2	3
Pixel referencia (filas)				
0	(0,0)	(0,1)	(0,2)	(0,3)
1	(1,0)	(1,1)	(1,2)	(1,3)
2	(2,0)	(2,1)	(2,2)	(2,3)
3	(3,0)	(3,1)	(3,2)	(3,3)

Cada pixel en la ventana se va convirtiendo sucesivamente en el pixel de referencia. Por lo tanto, existen diferentes matrices de co-ocurrencia para cada relación espacial según la orientación deseada. En la tabla 3 se muestra la relación espacial (1,0), es decir, la orientación horizontal.

Esta matriz expresa que el pixel de referencia es 0 y su vecino a la derecha es también 0 (Primera celda). En la tercera fila y tercera columna, tres veces el pixel de referencia es 2 y su vecino a la derecha es 2. Por lo tanto, se cuenta cada pixel de referencia con su vecino a la derecha. Si el cálculo se realiza solo de este modo, usando sólo una dirección, la matriz no es simétrica respecto a la diagonal (Presutti, 2004).

Para el correcto cálculo de la matriz de co-ocurrencia es necesaria la simetría. El método más sencillo para lograr la simetría de la matriz es sumándole su respectiva matriz transpuesta (Jiménez Guerrero, 2016). Sumando la transpuesta de la matriz de la tabla 3, el resultado quedaría como lo expresa la tabla 4.

Tabla 3.

Matriz de co-ocurrencia de la relación espacial (1,0) para la figura 7.

2	2	1	0
0	2	0	0
0	0	3	1
0	0	0	1

Tabla 4.

Matriz de co-ocurrencia simétrica de la relación espacial (1,0) para la figura 7.

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

El siguiente paso una vez obtenida la matriz simétrica, es expresar esta matriz como probabilidad. La definición más simple de probabilidad es: el número de veces que un evento ocurre, dividido por el número total de posibles eventos (Presutti, 2004). La ecuación para su cálculo es:

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}} \quad (1)$$

Dónde:

$i \rightarrow$ Es el número de filas y j el número de columnas

$V \rightarrow$ Es el valor de la celda i,j en la ventana

$P_{i,j} \rightarrow$ Es la probabilidad en la celda i,j

$N \rightarrow$ Es el número de filas o columnas.

Considerando la dimensión 4x4 de la imagen de ejemplo, y la relación (1,0), el número total de posibles pares es de 12 y para una relación horizontal (derecha mas izquierda) ese número se duplica (24) (Presutti, 2004). La ecuación (1) transforma la matriz de co-ocurrencia en una aproximación de tabla de probabilidad. Este proceso se denomina normalización de la matriz. Normalizando la matriz simétrica de la tabla 4 se tiene:

Tabla 5.
Matriz de co-ocurrencia normalizada de la relación espacial (1,0) para la figura 7.

0.166	0.083	0.042	0
0.083	0.166	0	0
0.042	0	0.250	0.042
0	0	0.042	0.083

Es importante referir algunos aspectos importantes con respecto a la matriz de co-ocurrencia simétrica y normalizada (Franklin, Moskal, & Presutti, 2001):

- ✓ Los elementos de la diagonal indican pares de pixeles que no tienen diferencias en su nivel de gris.
- ✓ Sumando los valores de la diagonal se tiene la probabilidad que un pixel tenga el mismo nivel de gris que su vecino.
- ✓ Las líneas paralelas a la diagonal separadas una celda, representan los pares de pixeles con una diferencia de un nivel de gris.
- ✓ Sumando los valores de estas diagonales paralelas obtenemos la probabilidad que un pixel tenga 1, 2, 3, etc., niveles de grises de diferencia con su vecino.

Hasta este punto, se ha detallado como se crea una matriz de co-ocurrencia normalizada, expresada como probabilidad, para una determinada relación espacial entre dos píxeles vecinos. De esta matriz se derivan las ocho características de co-ocurrencia texturales utilizadas para la extracción de características del patrón venoso (Jiménez Guerrero, 2016) (Presutti, 2004) (Escobar Riosalido, 2015).

2.4.1.1. Medidas de textura

Contraste

Esta medida indica las variaciones locales en los tonos de gris de una imagen. Cuanto mayor es la variación de los tonos de gris, mayor es el contraste:

$$C1 = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad 2)$$

Dónde:

$N \rightarrow$ Es el número de niveles de gris, ya que se trabaja con imágenes en escala de gris.

Energía

Esta medida da valores altos cuando en la matriz de co-ocurrencia tiene pocas entradas de gran magnitud, y es baja cuando todas las entradas son similares.

$$C2 = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (3)$$

Entropía

Este descriptor mide la aleatoriedad de la imagen. Cuando los elementos de la matriz de co-ocurrencia son iguales, alcanza su máximo valor.

$$C3 = \sum_{i,j=0}^{N-1} (-P_{i,j} \ln(P_{i,j})) \quad (4)$$

Homogeneidad

Esta medida es opuesta al contraste. Cuanto mayor es el número de la diagonal en la matriz de co-ocurrencia, más homogénea es la textura en esa sección de la imagen que se está analizando.

$$C4 = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2} \quad (5)$$

Probabilidad Máxima

Esta medida encuentra la probabilidad máxima del valor de la celda que se está analizando.

$$C5 = \max\{P_{i,j}\} \quad (6)$$

Tonalidad

Lo primero que se debe calcular es la media de la matriz de co-ocurrencia, tanto para las filas como para las columnas.

$$Mx = \sum_{i,j=0}^{N-1} i \cdot P_{i,j} \quad My = \sum_{i,j=0}^{N-1} j \cdot P_{i,j} \quad (7)$$

Se define entonces la tonalidad como:

$$C6 = \sum_{i,j=0}^{N-1} (i - Mx + j + My)^3 \cdot P_{i,j} \quad (8)$$

Importancia

Medida similar a la tonalidad cuya expresión viene dada por:

$$C7 = \sum_{i,j=0}^{N-1} (i - Mx + j + My)^4 \cdot P_{i,j} \quad (9)$$

Correlación

Es la dependencia lineal de los tonos de gris en la imagen. Si la correlación es 0 entonces no existe correlación lineal entre los niveles de gris. Este descriptor mide la semejanza de la imagen consigo misma desplazada.

$$C8 = \sum_{i,j=0}^{N-1} \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \quad (10)$$

Dónde:

$\mu_i, \mu_j \rightarrow$ Media de la muestra

$\sigma_i^2, \sigma_j^2 \rightarrow$ Varianza. Utiliza la combinación entre el pixel de referencia y el pixel vecino. Se definen de la siguiente manera:

$$\sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j} (i - \mu_i)^2 \quad (11)$$

$$\sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j} (j - \mu_j)^2 \quad (12)$$

2.5. Procedimientos de clasificación

En líneas generales, la clasificación consiste en asignar cada patrón de entrada a una sola categoría o clase de entre todas las clases. Para este trabajo de trabaja con 106 clases. Los procedimientos de clasificación han sido desarrollados hace varios años. Sin embargo, estos han sido mejorados y optimizados representando actualmente las ramas más importantes en clasificación, tanto a nivel aplicado como en las áreas de investigación.

2.5.1. Análisis discriminante lineal

Una muestra u observación se clasifica en una clase si la distancia al cuadrado es la mínima desde la observación hasta el centro de la clase. El supuesto del que se parte en este análisis es que las matrices de covarianzas son iguales para todas las clases (Universidad Juárez Autónoma de Tabasco, 2006). Para cualquier observación, la clase con la distancia al cuadrado mínima tiene la función

discriminante lineal más grande y la observación se clasifica en esa clase. El análisis del discriminante lineal cumple la propiedad de la distancia al cuadrado simétrica (Minitab, 2016). El aprendizaje de este clasificador consiste en elegir una función discriminante lineal por clase; es decir, un vector de pesos y un peso umbral por clase. Cada muestra a clasificar se asigna a la clase cuya función discriminante es maximizada (Universidad Politécnica de Valencia, 2006). La función discriminante lineal se ajusta a una densidad normal multivariada a cada grupo, con una estimación de covarianza agrupada (MathWorks, 2017).

La metodología aplicada al análisis del discriminante lineal se describe a continuación (Castrillón, Giraldo, & Sarache, 2008):

- a) Primero se calcula el vector de medias.
- b) Se calcula la matriz de covarianza
- c) Se calculan los coeficientes de las funciones discriminantes para cada una de las clases, con una base de datos de entrenamiento, previamente definidas. (Duda, Hart, & Stork, 2012).
- d) Se definen las funciones discriminantes lineales para cada una de las clases.
- e) Los valores de características seleccionadas y pertenecientes a la base de datos de validación son evaluados sobre cada una de las funciones discriminantes de probabilidad generadas. Se supone que la muestra evaluada, pertenece a aquella cuya función de probabilidad general el máximo valor.
- f) Los resultados obtenidos son comparados con una etiqueta previamente establecida, la cual indica el valor real de la clase, a la que pertenece la muestra. Como resultado de la comparación se define el porcentaje de aciertos y errores que el sistema proporciona en la identificación, de las clases a las cuales pertenece la muestra.

2.5.2. Análisis discriminante cuadrático

Para el análisis discriminante cuadrático se parte de la idea que las matrices de covarianzas no son iguales para todas las clases. El nombre de análisis discriminante cuadrático radica en que la distancia cuadrática no se simplifica en una función lineal

(Rodríguez, 2014). Una observación se clasifica en una clase que tiene la distancia elevada al cuadrado más pequeña. En este análisis, la distancia cuadrática no es simétrica (Minitab, 2016). El aprendizaje de este clasificador consiste en elegir una función discriminante cuadrática por clase; es decir, una matriz de pesos, un vector de pesos y un peso umbral por clase. Cada muestra a clasificar se asigna a la clase cuya función discriminante es maximizada (Universidad Politécnica de Valencia, 2006). La función discriminante cuadrática se ajusta a densidades normales multivariadas con estimaciones de covarianza estratificada por grupo (MathWorks, 2017).

La metodología aplicada al análisis del discriminante cuadrático es el mismo que se describió en el análisis del discriminante lineal, con la diferencia que en numeral d), se define las funciones discriminantes cuadráticos para cada una de las clases.

El análisis discriminante lineal y análisis discriminante cuadrático son dos clasificadores clásicos, con, como sus nombres sugieren, una superficie de decisión lineal y cuadrática, respectivamente. Estos clasificadores se ven como una red o que calcula funciones discriminantes y selecciona la categoría correspondiente al discriminante con máximo valor. En la figura 8 se indica una estructura funcional de un clasificador de patrones que incluye x_i entradas y g_i funciones discriminantes.

Por ejemplo, se considera clasificar a un individuo procedente de una de las dos poblaciones Π_1 y Π_2 . Se asume que para los individuos de Π_i ($i = 1,2$), la observación X se distribuye normalmente con una media μ_i desconocida y la matriz de covarianza Γ_i , y se dispone de una muestra de tamaño N_i para estimar estos parámetros. Sea \bar{X}_i y S_i la media de la muestra y la matriz de covarianza de la muestra, respectivamente, y sea S la matriz de covarianza de la muestra agrupada; es decir, $S = k_1 S_1 + k_2 S_2$, donde $k_i = n_i/n$ and $n = n_1 + n_2$ con $n_i = N_i - 1$. Las funciones discriminante lineal y discriminante cuadrático son (Duda, Hart, & Stork, 2012):

$$L(X) = \{X - \frac{1}{2}(\bar{X}_1 + \bar{X}_2)\}^T S^{-1}(\bar{X}_2 - \bar{X}_1), \quad (11)$$

$$Q(X) = (X - \bar{X}_1)^T S_1^{-1} (X - \bar{X}_1) - (X - \bar{X}_2)^T S_2^{-1} (X - \bar{X}_2) + \log|S_2^{-1} S_1|, \quad (12)$$

respectivamente. Una observación X es clasificada, por ejemplo dentro de Π_1 para valores negativos de estas funciones. Si $\Gamma_1 \neq \Gamma_2$, $Q(X)$ es mejor que $L(X)$ para muestras grandes, ya que $Q(X)$ es un estimador consistente de la función discriminante cuadrático mientras que $L(X)$ es no coherente. Pero para muestras pequeñas $Q(X)$ no siempre es mejor que $L(X)$. Para una clasificación multiclase, se realiza el mismo proceso comparando entre todas las funciones discriminantes.

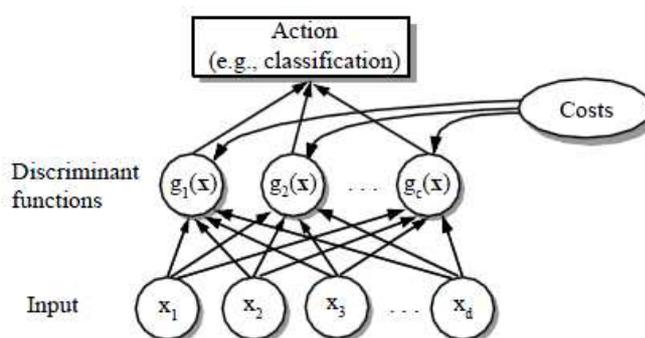


Figura 8. Estructura funcional de un clasificador basado en funciones discriminantes

Fuente: (Duda, Hart, & Stork, 2012)

2.5.3. Método del vecino más cercano

k- Nearest Neighbors (kNN) es un método de clasificación supervisado. Sirve para estimar la función de densidad $F(x/C_j)$ de los predictores x para cada clase C_j (García Laencina, y otros, 2008). Este es un método de clasificación no paramétrico, que estima el valor de la función de densidad de probabilidad o directamente la probabilidad a posteriori de que un elemento x pertenezca a la clase C_j a partir de la información proporcionada por el conjunto de prototipos. Este método no hace ninguna suposición en el proceso de aprendizaje, acerca de la distribución de las variables predictoras (Aler Mur, 2010). También se le conoce como algoritmo de aprendizaje basado en instancias. Su funcionamiento es muy simple: se almacenan los ejemplos de entrenamiento de datos históricos y cuando se requiere clasificar a un nuevo objeto, se extraen los objetos más parecidos y se usa su clasificación para clasificar al nuevo objeto (Morales, 2009) (Orea, Vargas, & Alonso, 2005). Los

vecinos más cercanos a una instancia se obtienen, para el caso de los atributos continuos, utilizando la distancia euclidiana sobre los n posibles atributos. El resultado de la clasificación por medio de este algoritmo puede ser discreto o continuo. En el caso discreto, el resultado de la clasificación es la clase más común de los k vecinos (Flórez, España, & Torres, 2014).

La regla de clasificación del vecino más cercano se utiliza en muchas tareas por su simplicidad y eficacia. El algoritmo más sencillo para implementar dicha regla es conocido como exhaustivo. Este algoritmo calcula todas las distancias de la muestra desconocida a los prototipos del conjunto de entrenamiento y asigna a la muestra la clase de aquel prototipo cuya distancia sea mínima (Carratalá & Seco, 2004). Durante el entrenamiento, solo almacena las instancias, no construye ningún modelo.

Un parámetro importante es el valor de k vecinos más cercanos se escoja. De este valor escogido dependerá el desempeño del clasificador. Históricamente, el valor de k óptimo para la mayoría de conjuntos de datos ha estado entre 3 – 10 (Carratalá & Seco, 2004). Hay un valor de k para el cual, el clasificador se estabiliza y entrega tasas de desempeño semejantes.

Por ejemplo, se quiere clasificar la estrella como se observa en la figura 9. Para $k = 3$ esta es clasificada en la clase B, ya que hay dos círculos morados y solo un amarillo dentro del círculo que los contiene. Para $k = 6$, la estrella es clasificada en la clase A, ya que hay cuatro círculos amarillos y solo dos morados en el círculo que los contiene. En caso de empate, se puede añadir alguna regla heurística como puede ser el vecino más próximo (Cárdenas Montes, 2015).

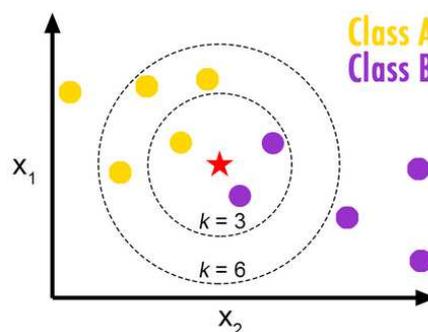


Figura 9. Ejemplo de clasificador k-NN para dos clases
Fuente: (Lewis, 2017)

CAPÍTULO III

DESCRIPCIÓN DEL SISTEMA DE BIOMETRÍA VASCULAR

En este capítulo se detalla el modelo implementado para un sistema biométrico vascular, obedeciendo la disposición de los diagramas de bloques indicados en la figura 1. Cabe recalcar que en este trabajo se ha desarrollado el software del sistema. Es decir, las etapas de pre-procesamiento, extracción de características y clasificación. Para ello se toma 6 muestras del dedo índice de la mano izquierda de 106 individuos. Es decir, 106 clases. Esto da un total de 636 observaciones. El objetivo es determinar a qué clase pertenece cada observación. En la figura 10 se indica el diagrama de flujo del sistema.

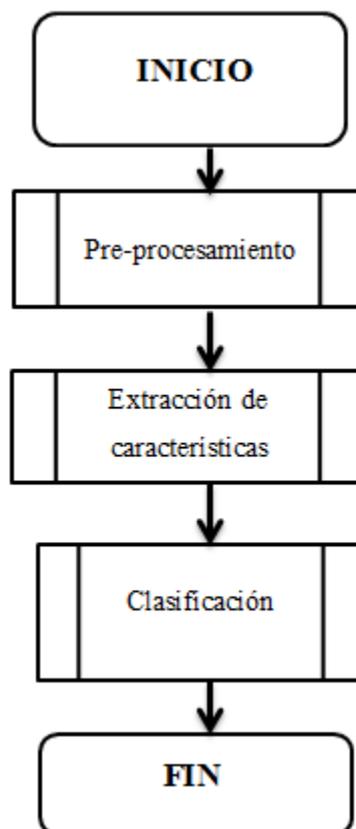


Figura 10. Diagrama de flujo del sistema biométrico vascular

3.1. Módulo de adquisición de imágenes

En general, el proceso que se emplea para la adquisición de una imagen para un sistema biométrico vascular consiste en utilizar la luz infrarroja para capturar los vasos sanguíneos. Este sistema se ilustra en la figura 11.

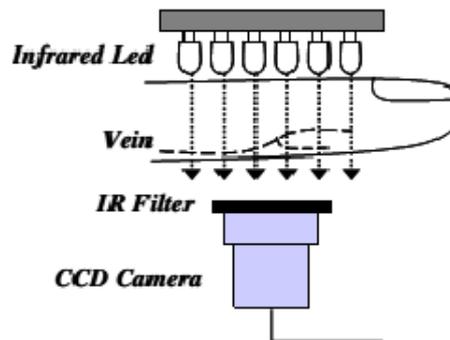
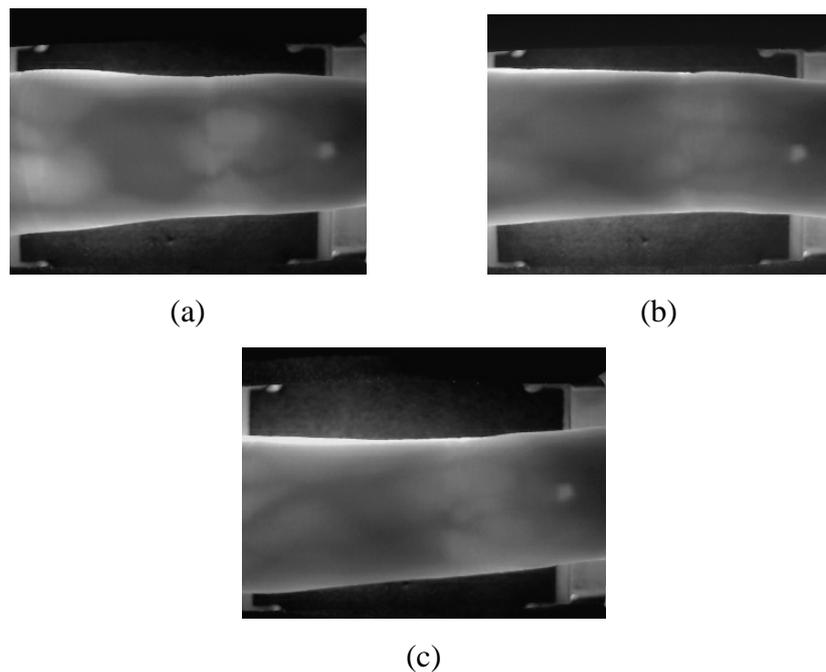


Figura 11. Sistema de adquisición de imágenes para un sistema biométrico vascular
Fuente: (Beltrán, Gómez, Flórez, & Vasquez , 2016).

En esta etapa existen ciertos factores que influyen en la calidad de la muestra. Los factores más importantes son el nivel de oxígeno en la sangre, el tipo de sensor infrarrojo y la iluminación. Por esta razón y con el fin de estandarizar las muestras, se ha trabajado con una base de datos pública llamada SDUMLA-HMT (Liu, Yin, & Sun, 2011). SDUMLA-HMT es una base de datos biométrica multimodal con muestras de 106 individuos, 61 hombres y 45 mujeres entre 17 y 31 años de edad. El dispositivo usado para la captura de las venas de los dedos fue diseñado por Joint Lab para inteligencia informática y sistemas inteligentes de la Universidad Wuhan. En el proceso de captura, cada individuo proporcionó imágenes de su dedo índice, medio y anular en 6 oportunidades. Algunas de las muestras se ilustran en la figura 12.

3.2. Módulo de pre procesamiento y extracción de características

La imagen muestra que captura el sistema, es una imagen infrarroja. Esta imagen contiene la información del patrón vascular único del individuo muestreado.



**Figura 12. (a) Muestra vascular del dedo índice de la mano izquierda
 (b) Muestra vascular del dedo medio de la mano izquierda
 (c) Muestra vascular del dedo anular de la mano izquierda
 Fuente: SDUMLA-HMT multimodal biometric database.**

La función pre-procesamiento indicada en la figura 10, realiza la lectura de las imágenes muestras de cada usuario y las ordena en dos matriz de 3 dimensiones. La primera matriz llamada *Muestras_Imagen* almacena las imágenes tomadas para el entrenamiento y en la segunda matriz llamada *Test_Imagen* se guardan las muestras tomadas para la evaluación. La dimensión de estas matrices es de $240 \times 320 \times \text{MuestrasParaEntrenamiento}$ y de $240 \times 320 \times \text{MuestrasParaEvaluación}$, respectivamente. Posterior a esto se realiza la descomposición utilizando la transformada *wavelet*. El diagrama de flujo de la función pre-procesamiento se indica en la figura 15.

Se ha escogido dos tipos de wavelets con el fin de comparar sus resultados. Estas *wavelet* madre son la *wavelet spline* biortogonal de orden 2 y la *wavelet haar*. (Aldroubi, Eden, & Unser, 1993).

Las características de la *wavelet spline* biortogonal son (MathWorks, 2017):

- ✓ Familia biortogonal.
- ✓ Wavelet biortogonal con fase lineal

- ✓ Soporte compacto. Sólo difieren de cero en un intervalo finito y permiten algoritmos de transformación más rápidos.
- ✓ Simetría y reconstrucción exacta con filtros FIR.
- ✓ N_r y N_d orden de los filtros FIR de reconstrucción y descomposición respectivamente.
- ✓ Tamaño del soporte: $2N_r+1$ para reconstrucción, $2N_d+1$ para descomposición.
- ✓ Longitud del filtro: $\max(2N_r, 2N_d)+2$.
- ✓ Transformada *wavelet* discreta es posible.
- ✓ Transformada *wavelet* continua es posible.
- ✓ Momentos de desvanecimiento: N_r . Esta propiedad está relacionada con la compresión de información y eliminación del ruido.

Las características de la *wavelet haar* son (MathWorks, 2017):

- ✓ Familia *Haar*.
- ✓ Wavelet Ortogonal con fase lineal.
- ✓ Soporte compacto.
- ✓ Transformada *wavelet* discreta es posible.
- ✓ Transformada *wavelet* continua es posible.
- ✓ Tamaño del soporte: 1.
- ✓ Longitud de filtro: 2.
- ✓ No es continua.
- ✓ Simétrica.
- ✓ Momentos de desvanecimiento: 1.

La *wavelet spline* biortogonal tiene la propiedad de fase lineal, que es necesaria para la reconstrucción de la señal y la imagen. Esta familia de *wavelet* es continua y utiliza dos funciones, la una para descomposición y la otra para reconstrucción como se indica en la figura 13 (MathWorks, 2017).

Por otro lado, la *wavelet haar* también tiene la propiedad de fase lineal, que es necesaria para la reconstrucción de la señal y la imagen. A diferencia de la familia de

wavelet spline, esta familia de *wavelet* es discontinua y utiliza una sola función para descomposición y reconstrucción como se indica en la figura 14 (MathWorks, 2017).

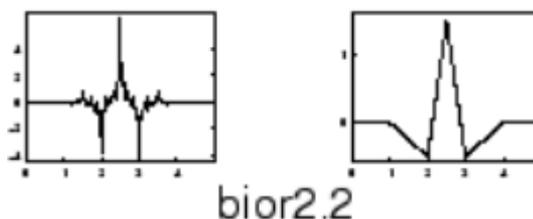


Figura 13. Wavelet spline biortogonal. Función para descomposición (lado izquierdo) y función para reconstrucción (lado derecho)

Fuente: (MathWorks, 2017)

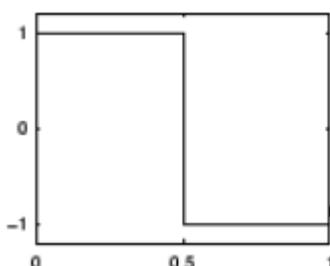


Figura 14. Wavelet haar. Función para descomposición y reconstrucción.

Fuente: (MathWorks, 2017)

Por esta razón se ha escogido este tipo de wavelets con el fin de comparar sus resultados con estas diferencias en sus características. Con cada una de estas *wavelets* madre escogida, se realiza la descomposición en su tercer y cuarto nivel. Los niveles de descomposición están estrictamente relacionados con el tamaño de la imagen a ser procesada. Para las imágenes utilizadas en este trabajo es posible llegar a un cuarto nivel de descomposición. Esto se realiza para comparar los resultados entre las wavelets escogidas y para analizar los resultados en cada nivel de descomposición procesado.

De cada nivel de descomposición se obtienen 3 imágenes de detalle en vertical, horizontal y diagonal. De estas imágenes de detalle, las características de textura de co-ocurrencia forman los vectores de características de cada muestra u observación. Estos vectores de características de co-ocurrencia son los que se han comparado para determinar la clase a la que pertenece la observación correspondiente.

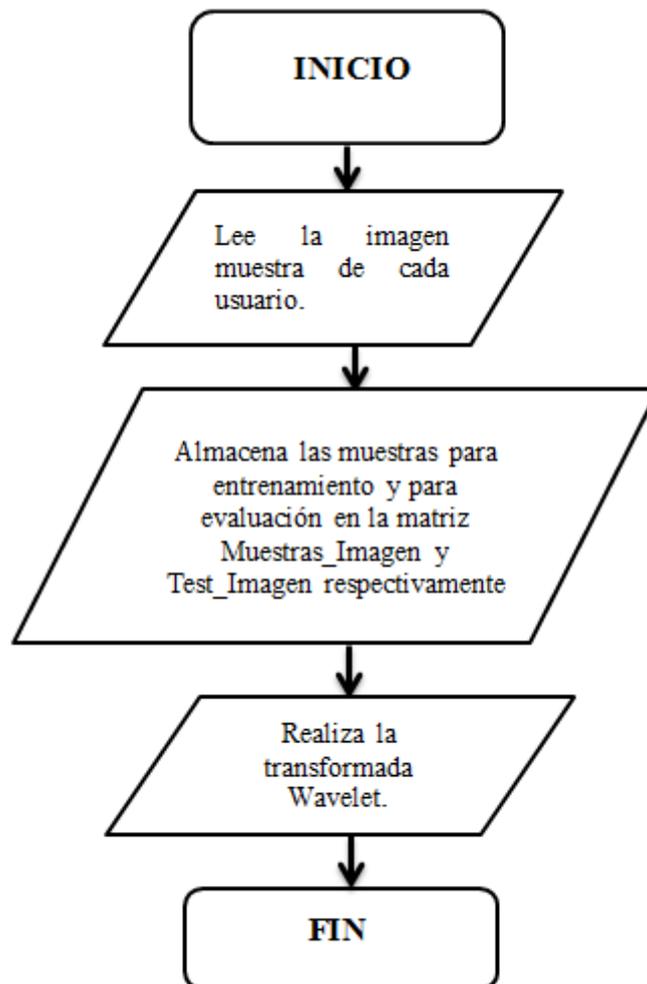


Figura 15. Diagrama de flujo de la función pre-procesamiento

La siguiente etapa es extraer las características de las muestras obtenidas. La función extracción de características de la figura 10, recibe como parámetro de entrada las matrices de detalle horizontal, vertical y diagonal de cada uno de los niveles de descomposición. Como resultado devuelve las características de co-ocurrencia de cada muestra de la matriz de muestras para entrenamiento y para evaluación. Estas características son almacenadas en dos matrices diferentes. Estas matrices son las que almacena las características de co-ocurrencia de todos los usuarios procesados. En la figura 16 se indica el diagrama de flujo de la función extracción de características.

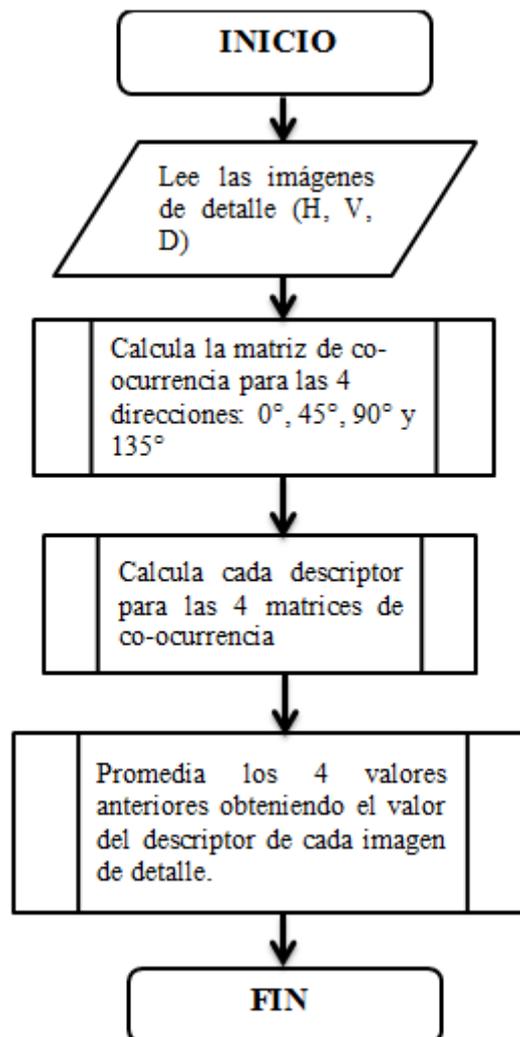


Figura 16. Diagrama de flujo de la función extracción de características

Para realizar la transformada *wavelet spline* biortogonal se ha empleado las siguientes instrucciones en Matlab:

❖ $[RF, DF] = \text{biorwavf}('biorNr.Nd');$

Dónde:

$\text{biorwavf} \rightarrow$ Hace referencia a *biothogonal spline wavelet filtros*. Computa el filtro escalado de reconstrucción (RF) y el filtro escalado de descomposición (DF).

$('biorNr.Nd') \rightarrow$ Identifica al conjunto de *wavelets spline* biortogonales. Nr y Nd son números que especifican el orden de los filtros. En este caso $Nr = Nd = 2$, ya que la transformada es de orden 2.

❖ $[Lo_D, Hi_D, Lo_R, Hi_R] = \text{biorfilt}(DF, RF);$

Dónde:

biorfilt → Hace referencia a biorthogonal wavelet filter set. Computa el filtro paso bajo y paso alto de descomposición y de reconstrucción.

Lo_D, Hi_D → Filtro de descomposición paso bajo y paso alto respectivamente.

Para obtener las sucesivas sub-imágenes de detalle así como la imagen tendencia, existe una instrucción en Matlab que permite obtenerlas con facilidad:

❖ $[A, H, V, D] = \text{swt2}(I, N, Lo_D, Hi_D);$

Dónde:

I → Imagen a procesar

swt2 → Hace referencia a *discrete stationary wavelet transform 2-D*. Proporciona la imagen wavelet estacionaria con N niveles de descomposición.

Para realizar la transformada *wavelet haar* se ha empleado las siguientes instrucciones en Matlab:

❖ $[Lo_D, Hi_D, Lo_R, Hi_R] = \text{wfilters}('haar');$

Dónde:

wfilters → Hace referencia al cálculo de los filtros paso bajo y paso alto de descomposición y de reconstrucción.

Lo_D, Hi_D → Filtro de descomposición paso bajo y paso alto respectivamente.

Para obtener las sucesivas sub-imágenes de detalle así como la imagen tendencia, existe una instrucción en Matlab que permite obtenerlas con facilidad:

❖ $[A, H, V, D] = \text{swt2}(I, N, Lo_D, Hi_D);$

Dónde:

I → Imagen a procesar

swt2 → Hace referencia a *discrete stationary wavelet transform 2-D*. Proporciona la imagen wavelet estacionaria con N niveles de descomposición.

Tanto para la *wavelet spline* biortogonal como para la *wavelet haar*, se ha realizado una representación overcomplete. Esta representación no realiza el sub-muestreo de las imágenes de detalle en cada paso. De esta forma se logra una representación más estable (Fan & Laine, 1996). De este modo las sub-matrices de detalle y de aproximación obtenidas tendrán el mismo tamaño que la imagen muestra procesada.

A partir de las sub-matrices de orientación (vertical, horizontal y diagonal) que contienen todas las sub-imágenes *wavelet* de la imagen muestra en los N niveles de descomposición, se calcula las matrices de co-ocurrencia simétrica y normalizada para todas las direcciones propuestas (0°, 45°, 90° y 135°). Se calculan los ocho descriptores para cada una de estas matrices. Por lo tanto se obtendrán cuatro valores de cada descriptor. Estos valores serán promediados para obtener un único valor para cada descriptor. Este proceso se realiza para cada imagen de detalle en cada uno de los niveles de descomposición. Además, se trabaja con imágenes en escala de gris, por lo tanto, se procesa la matriz de co-ocurrencia con 256 niveles de gris. Es decir, 8 bits de profundidad por pixel. La siguiente instrucción en Matlab facilita este paso:

❖ `glcm = graycomatrix(imdet, 'NumLevels', n, 'offset', [v h]);`

Dónde:

`graycomatrix` → Crea la matriz de co-ocurrencia de una imagen en escala de grises.

`imdet` → Hace referencia a la imagen de detalle.

`'NumLevels', n` → Ayuda a discretizar la imagen con $n = 256$ niveles de gris ya que se trabaja con imágenes en escala de gris.

`'offset', [v h]` → Indica el ángulo del píxel de interés respecto a su vecino.

La matriz obtenida es una matriz asimétrica por lo que es necesario convertirla en una matriz simétrica. Para esto se le suma su matriz transpuesta. Posteriormente, la matriz simétrica obtenida se le divide para la suma de todos los coeficientes de la matriz simétrica. Esto con el fin de obtener la matriz simétrica normalizada. Una vez obtenida la matriz de co-ocurrencia simétrica y normalizada para cada nivel de descomposición, se calcula las ocho características texturales propuestas. Estos son: contraste, energía, entropía, homogeneidad, probabilidad máxima, tonalidad, importancia y correlación.

Cada observación evaluada posee 3 sub-imágenes de detalle por cada nivel de descomposición. En la tabla 6 se detallan cuantas características se procesarán por cada nivel de composición evaluado.

Tabla 6.

Detalle de características para cada nivel evaluado

Nivel de descomposición	Total de sub-imágenes de detalle	Características de textura calculados	Total de características
N = 3	9	8	72
N = 4	12	8	96

La matriz que contiene las características de co-ocurrencia será:

a) Para N = 3:

Utilizando 5 imágenes para entrenamiento y 1 imagen para evaluación, la matriz de características de las muestras para entrenamiento es de orden 72x530 y, la matriz de características de las muestras para evaluación es de orden 72x106. Las filas representan las características y las columnas las observaciones.

b) Para N = 4:

Utilizando 4 imágenes para entrenamiento y 2 imágenes para evaluación, la matriz de características de las muestras para entrenamiento es de orden 96x424 y, la matriz de características de las muestras para

evaluación es de orden 96×212 . Las filas representan las características y las columnas las observaciones.

3.3. Módulo de clasificación

La función Clasificación expuesta en la figura 10, se encarga del reconocimiento de patrones propiamente dicho. Esta función recibe como entrada las matrices de características de co-ocurrencia tanto de entrenamiento como de evaluación. Además, recibe como parámetro de entrada un vector que indica la clase a la que pertenece cada vector de características de la matriz de características de las muestras para entrenamiento.

Antes de proceder con la clasificación, estas matrices de características deben normalizarse para que los coeficientes con mayor varianza no sean las que dominen la clasificación. La normalización se la realiza obteniendo la mediana y la desviación media de cada característica y a cada una se le resta la mediana y se divide entre la desviación media. Además, de estas matrices se calcula su respectiva matriz transpuesta y con estas matrices transpuestas y normalizadas se trabaja para la clasificación. Es decir, ahora las filas corresponden a las observaciones y las columnas indican el número de características.

Para la validación se ha empleado dos métodos. El primer método es utilizando 5 imágenes para entrenamiento y 1 imagen para evaluación. En el segundo método es utilizando 4 imágenes para entrenamiento y 2 imágenes para evaluación.

Primeramente se realiza la clasificación con el método del vecino más cercano. Se determina la proximidad de cada vector de características con cada patrón de clase. Para esto se trabaja con la distancia euclídea. En Matlab existe una instrucción que permite implementar fácilmente este método:

❖ `Class = knnclassify(Sample, Training, Group, k, distance, rule);`

Dónde:

Sample → Hace referencia al vector de observación que se desea clasificar.

Training → Es la matriz con el resto de vectores de observación.

Group → El vector con las etiquetas de los respectivos grupos de la matriz training.

k → Número de vecino más cercanos a tomar en cuenta.

distance → Se ha escogido la opción por defecto. La distancia euclídea.

rule → en caso de empate, la regla que se ha escogido es aquella que se queda con el más cercano.

Para realizar la clasificación con el análisis discriminante cuadrático, Matlab posee una instrucción que es:

❖ `Class = classify(Sample, Training, Group, 'diagQuadratic');`

Dónde:

Sample → Hace referencia al vector de observación que se desea clasificar.

Training → Es la matriz con el resto de vectores de observación.

Group → El vector con las etiquetas de los respectivos grupos de la matriz training.

'diagQuadratic' → Hace referencia al análisis discriminante cuadrático.

Por último, para realizar la evaluación con el discriminante lineal se utiliza la siguiente instrucción:

❖ `Class = classify(Sample, Training, Group, 'diagLinear');`

Dónde:

Sample → Hace referencia al vector de observación que se desea clasificar.

Training → Es la matriz con el resto de vectores de observación.

Group → El vector con las etiquetas de los respectivos grupos de la matriz training.

'diagLinear' → Hace referencia al análisis discriminante lineal.

Los resultados de cada clasificador se almacenan en un vector y estos son comparados con las clases a los que realmente pertenece cada observación de la matriz de características de evaluación. Esta comparación proporciona la tasa correcta de clasificación. Finalmente, en un vector de resultados se guarda la tasa correcta de clasificación. En la figura 17 se presenta el diagrama de flujo de la función clasificación.

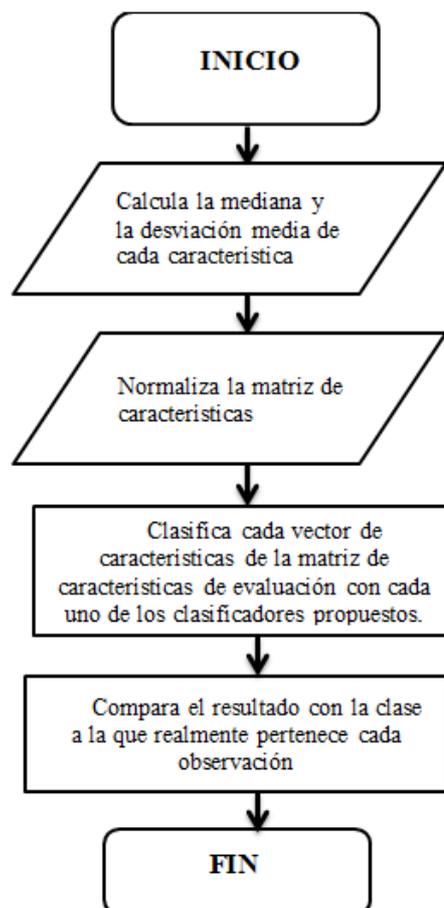


Figura 17. Diagrama de flujo de la función Clasificación

CAPÍTULO IV

PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

4.1. Presentación de los resultados obtenidos

Las pruebas se han hecho para las 636 observaciones de los 106 usuarios (6 muestras del dedo índice de la mano izquierda por cada usuario), con los dos tipos de *wavelets* escogidas variando los niveles de descomposición N de cada *wavelet* utilizando los tres clasificadores propuestos. Para la validación se ha empleado dos métodos de la siguiente manera:

- ✓ Validación A: 5 imágenes para entrenamiento, 1 imagen para evaluación.
- ✓ Validación B: 4 imágenes para entrenamiento, 2 imágenes para evaluación.

Para cada una de las pruebas realizadas se ha obtenido la matriz de confusión y de esta se puede calcular la tasa de falsa aceptación (FAR) y la tasa de falso rechazo (FRR) en cada prueba. Las figuras 18 y 19 muestran una visión parcial de la matriz de confusión de los primeros 20 usuarios, utilizando la validación A y la validación B respectivamente. Las matrices de confusión completa se pueden encontrar en el anexo 1 y 2, respectivamente. Estas matrices corresponden a los mejores resultados obtenidos en todas las pruebas realizadas.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1																			
2		1																		
3			1																	
4				1																
5					1															
6						1														
7							1													
8								1												
9									1											
10										1										
11											1									
12												1								
13													1							
14														1						
15															1					
16																1				
17																	1			
18																		1		
19																			1	
20																				1

Figura 18. Matriz de confusión realizando la validación A.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0																			
2		2																		
3			2																	
4				1																
5					0															
6						2														
7							1											1		
8								2												
9									2											
10										1										
11											2									
12												2								
13													2							
14					1									2						
15															0					
16																2				
17																	2			
18																		1		
19																			2	
20																				1

Figura 19. Matriz de confusión realizando la validación B.

4.1.1. Resultados utilizando el clasificador discriminante lineal

4.1.1.1. Validación A: 5 imágenes de entrenamiento y 1 imagen para evaluación

La tabla 7 muestra los resultados obtenidos aplicando la *wavelet spline* biortogonal y la *wavelet haar*, utilizando como clasificador al discriminante lineal.

Tabla 7.

Resultados del clasificador discriminante lineal y validación A.

	N=3			N=4		
	FAR (%)	FRR (%)	Desempeño (%)	FAR (%)	FRR (%)	Desempeño (%)
Wavelet spline biortogonal	7,62	38,10	62,26	6,10	30,48	69,81
Wavelet haar	4,95	24,76	75,47	4,00	20,00	80,19

4.1.1.2. Validación B: 4 imágenes de entrenamiento y 2 imágenes para evaluación

La tabla 8 muestra los resultados obtenidos aplicando la *wavelet spline* biortogonal y la *wavelet haar*, utilizando como clasificador al discriminante lineal.

Tabla 8.

Resultados del clasificador discriminante lineal y validación B.

	N=3			N=4		
	FAR (%)	FRR (%)	Desempeño (%)	FAR (%)	FRR (%)	Desempeño (%)
Wavelet spline biortogonal	23,33	46,67	7,55	18,48	46,19	8,49
Wavelet haar	23,21	46,43	8,02	18,19	45,47	9,91

4.1.2. Resultados utilizando el clasificador discriminante cuadrático**4.1.2.1. Validación A: 5 imágenes de entrenamiento y 1 imagen para evaluación**

La tabla 9 muestra los resultados obtenidos aplicando la *wavelet spline* biortogonal y la *wavelet* haar, utilizando como clasificador al discriminante cuadrático.

Tabla 9.

Resultados del clasificador discriminante cuadrático y validación A.

	N=3			N=4		
	FAR (%)	FRR (%)	Desempeño (%)	FAR (%)	FRR (%)	Desempeño (%)
Wavelet spline biortogonal	14,09	70,47	30,19	14,09	70,47	30,19
Wavelet haar	11,05	55,24	45,28	10,29	51,43	49,06

4.1.2.2. Validación B: 4 imágenes de entrenamiento y 2 imágenes para evaluación

La tabla 10 muestra los resultados obtenidos aplicando la *wavelet spline* biortogonal y la *wavelet* haar, utilizando como clasificador al discriminante cuadrático.

Tabla 10.

Resultados del clasificador discriminante cuadrático y validación B.

	N=3			N=4		
	FAR (%)	FRR (%)	Desempeño (%)	FAR (%)	FRR (%)	Desempeño (%)
Wavelet spline biortogonal	24,52	49,05	2,83	19,52	48,81	3,30
Wavelet haar	24,05	48,09	4,72	19,05	47,62	5,66

4.1.3. Resultados utilizando el clasificador *k*-nearest neighbor (k-NN)

Dentro de la evaluación del clasificador k-NN, se realizó la clasificación para diferentes valores de *k*. El objetivo de esta evaluación es determinar el valor de *k* para el cual el sistema estabiliza su desempeño. Esta evaluación se la ha realizado en un solo escenario empleando la *wavelet* spline biortogonal en su cuarto nivel de descomposición y el método de validación A. La figura 20 y la tabla 11 muestran los resultados obtenidos.

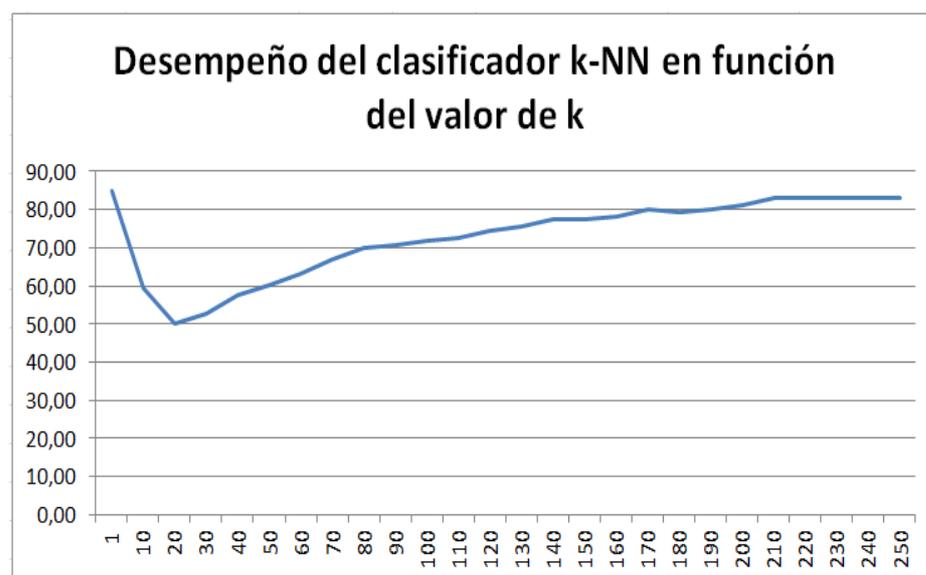
**Figura 20. Desempeño del clasificador k-NN en función del valor de *k*.**

Tabla 11.

Evaluación de desempeño del clasificador k-NN en función de k, con el método de validación A.

k	Desempeño (%)
1	84,91
10	59,43
20	50,00
30	52,83
40	57,55
50	60,38
60	63,21
70	66,98
80	69,81
90	70,75
100	71,70
110	72,64
120	74,53
130	75,47
140	77,36
150	77,36
160	78,30
170	80,19
180	79,25
190	80,19
200	81,13
210	83,02
220	83,02
230	83,02
240	83,02
250	83,02

Se puede observar que el clasificador presenta su más alto desempeño para un valor de $k=1$. Sin embargo, este valor decae para valores de k que oscilan a partir de 2 hasta 20, obteniendo su peor desempeño en $k=20$. A partir de $k=20$, el desempeño del clasificador empieza a mejorar llegando a estabilizarse a partir de $k=210$. Por lo tanto, este es el valor que se ha escogido para realizar las siguientes pruebas con el clasificador k-NN.

4.1.3.1. Validación A: 5 imágenes de entrenamiento y 1 imagen para evaluación

La tabla 12 muestra los resultados obtenidos aplicando la *wavelet spline* biortogonal y la *wavelet haar*, utilizando el método de clasificación k-NN.

Tabla 12.

Resultados del clasificador k-NN y validación A.

	N=3			N=4		
	FAR (%)	FRR (%)	Desempeño (%)	FAR (%)	FRR (%)	Desempeño (%)
Wavelet spline biortogonal	4,00	20,00	80,19	3,43	17,14	83,02
Wavelet haar	4,00	20,00	80,19	3,43	17,14	83,02

4.1.3.2. Validación B: 4 imágenes de entrenamiento y 2 imágenes para evaluación

La tabla 13 muestra los resultados obtenidos aplicando la *wavelet spline* biortogonal y la *wavelet haar*, utilizando el método de clasificación k-NN.

Tabla 13.

Resultados del clasificador k-NN y validación B. Imagen 1.

	N=3			N=4		
	FAR (%)	FRR (%)	Desempeño (%)	FAR (%)	FRR (%)	Desempeño (%)
Wavelet spline biortogonal	21,19	42,38	16,04	16,67	41,67	17,45
Wavelet haar	21,67	43,33	14,15	17,14	42,86	15,09

4.2. Análisis de los resultados obtenidos

Hablando únicamente de las pruebas realizadas con la *wavelet haar*, se puede decir que:

El clasificador discriminante cuadrático, discriminante lineal y el clasificador k-NN, tienen su mejor desempeño en el cuarto nivel de descomposición de la *wavelet haar*. La mejor tasa de acierto porcentual es del clasificador k-NN con $k = 210$ vecinos más cercanos presentando un 83,02%. Estos valores se han obtenido utilizando 5 imágenes para el entrenamiento y 1 imagen para validación. Es decir, de los 106 usuarios para la evaluación, 88 son correctamente clasificados.

Para el caso de utilizar 4 imágenes para el entrenamiento y 2 imágenes para la validación, los clasificadores tienen su mejor desempeño en el cuarto nivel de descomposición de la *wavelet haar*. El clasificador k-NN con $k = 210$ vecinos más cercanos es el que presenta la mejor tasa de acierto, siendo de 15,09%. Es decir, de las 212 imágenes procesadas, 16 imágenes son clasificadas correctamente.

Por otro lado, haciendo referencia a las pruebas obtenidas utilizando la *wavelet spline* biortogonal, se puede mencionar que:

Los tres clasificadores empleados presentan sus mejores tasas de desempeño para el cuarto nivel de descomposición de la *wavelet spline* biortogonal. El valor más alto lo presenta el clasificador k-NN para un valor de $k = 210$ vecinos más cercanos. Esta tasa de acierto es del 83,02%. Esto es utilizando 5 imágenes para el entrenamiento y 1 imagen para la evaluación.

Utilizando 4 imágenes para el entrenamiento y 2 imágenes para la validación, el comportamiento es similar. El clasificador k-NN es el que mejor desempeño tiene en el cuarto nivel de descomposición de la *wavelet spline* biortogonal. La tasa de acierto es de 17,45%. Es decir, 18 imágenes son clasificadas correctamente del conjunto de 212 imágenes.

Haciendo una comparación entre las dos *wavelets* utilizadas se puede mencionar que:

En general, los tres clasificadores implementados entregan mejores tasas de acierto al realizar el pre-procesamiento con *wavelet haar* en comparación a los

obtenidos utilizando la *wavelet spline* biortogonal. Únicamente, utilizando 4 imágenes para entrenamiento y 2 imágenes para evaluación, se obtienen mejores resultados utilizando la *wavelet spline* biortogonal.

Haciendo referencia a los clasificadores implementados, se puede mencionar que:

El clasificador discriminante cuadrático es el que presenta las peores tasas de desempeño en todos los escenarios analizados. El clasificador discriminante lineal presenta tasas de desempeño entre el 62,26% y 80,19%. El clasificador k-NN es el mejores tasas de desempeño presenta para los dos tipos de *wavelets* empleadas en su tercer y cuarto nivel de descomposición. Todos estos valores son analizados al utilizar como método de validación 5 imágenes para entrenamiento y 1 para la evaluación. Mientras que, utilizando 4 imágenes para entrenamiento y 2 para evaluación, los resultados son bastante malos. Las tasas de desempeño con este método de validación oscilan entre 2,83% con el clasificador discriminante cuadrático y 17,45% con el clasificador k-NN.

En todas las pruebas realizadas se puede evidenciar que las tasas de falsa aceptación y falso rechazo disminuyen mientras crece el nivel de descomposición de la *wavelet haar* o *wavelet spline* biortogonal.

CAPÍTULO V

CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS

5.1. Conclusiones

Los niveles de descomposición o profundidad de las *wavelets* utilizadas influyen radicalmente en el desempeño de la clasificación. Es mejor el desempeño del clasificador en el nivel más profundo analizado en este trabajo del par de *wavelets* escogidas. La profundidad de la *wavelet* depende del tamaño de la imagen que se está procesando. Esto sucede ya que en estos niveles de descomposición se detallan mucho mejor los cambios locales en la imagen original.

Dentro del procesamiento wavelet es mucho mejor trabajar con una misma función para descomposición y para reconstrucción. Esto permite obtener mejores tasas de acierto del sistema en comparación al trabajar con funciones para descomposición y para reconstrucción diferentes.

El tipo de *wavelet* aplicada al pre-procesamiento, el número de características analizadas, el clasificador empleado y los métodos de validación ejecutados, influyen directamente en la tasa de falsa aceptación y la tasa de falso rechazo. Varias pruebas han sido realizadas con el objetivo de reducir estas tasas. Al lograr reducir al mínimo posible la tasa de falsa aceptación y la tasa de falso rechazo, se puede garantizar la seguridad del sistema biométrico vascular planteado.

En el clasificador k-NN, el valor de k vecinos más cercanos, influye a la hora de evaluar el desempeño del clasificador. Se ha realizado varias pruebas determinando que a partir de un valor de $k=210$ el desempeño del clasificador se estabiliza. El correcto escogimiento de este valor depende de los datos que se estén evaluando. Por lo que se hace necesario realizar varias pruebas para determinar correctamente este valor.

El clasificador *k-nearest neighbor* (k-NN) utilizando una *wavelet haar* o una *wavelet spline* biortogonal para la descomposición de la imagen, es el que mejor tasas de acierto presenta. Además, presenta bajas tasas de falsa aceptación y falso rechazo en el cuarto nivel de descomposición del par de *wavelets* utilizadas. Si se quiere definir el mejor sistema de biometría vascular sería el que emplea: en el pre-procesamiento la *wavelet haar* o *wavelet spline* biortogonal en su cuarto nivel de descomposición, clasificador k-NN con $k=210$ vecinos más cercanos, utilizando 5 imágenes para el entrenamiento y 1 imagen para validación. Este sistema presenta una tasa de acierto del 83,02%, una tasa de falsa aceptación de 3,43% y una tasa de falso rechazo de 17,14%.

En referencia a los dos métodos de validación empleados se observa que los mejores resultados son empleando 5 imágenes para el entrenamiento y 1 imagen para la validación. Esto lleva a concluir que: mientras más grande sea el conjunto de datos de entrenamiento, el clasificador tiene un mejor desempeño presentando una alta tasa de acierto además de, una baja tasa de falsa aceptación y falso rechazo.

5.2. Recomendaciones

Se recomienda realizar el pre-procesamiento con una *wavelet* que tenga una misma función para la descomposición y para la reconstrucción.

Es de vital importancia realizar el análisis para escoger un valor de k adecuado para el clasificador k-NN, ya que este valor es determinante en el desempeño del clasificador.

Si la etapa de adquisición de rasgos es diferente a la utilizada en este proyecto, se recomienda trabajar con la *wavelet haar* o *wavelet spline* biortogonal en su cuarto nivel de descomposición, clasificador k-NN con $k=210$, utilizando 5 imágenes para el entrenamiento y 1 imagen para la validación.

5.3. Trabajos futuros

Como principal proyección a este trabajo sería interesante trabajar y desarrollar un prototipo que realice la etapa de adquisición de imágenes. Los resultados obtenidos en el sistema propuesto dependen mucho de la calidad de muestra obtenida. Sería importante evaluar el software implementado en conjunto con el hardware necesario para tomar la muestra vascular del usuario.

Ya que se cuenta con una base de datos multimodal como lo es SDUMLA-HMT, sería importante implementar un sistema biométrico multimodal. Es decir, un sistema biométrico que se complemente con otras técnicas de biometría para asegurar al usuario mucha mayor seguridad. Por ejemplo, se podría integrar estratégicamente sistemas biométricos que integren técnicas de reconocimiento basadas en patrones vasculares y además el reconocimiento de la huella dactilar.

Los sistemas biométricos en general presentan algunos inconvenientes relacionados con la seguridad y la privacidad. En comparación a los sistemas basados en contraseñas, que se pueden renovar sin necesidad de que se comprometa la seguridad del sistema, los rasgos biométricos están asociados al usuario de forma permanente. Por lo tanto, si el rasgo vascular de algún usuario es robado de la base de datos de una aplicación de biometría vascular, todas las aplicaciones en las que este rasgo se utilice se verían comprometidas. Por lo tanto, se propone utilizar biometría cancelable dentro del sistema propuesto para evitar el robo de los patrones vasculares.

REFERENCIAS

- Aglío Caballero, A. (2016). *Implementación y evaluación de un sistema biométrico de reconocimiento de venas de las manos mediante descriptores locales de textura*. Obtenido de http://oa.upm.es/43449/1/PFC_ALICIA_AGLIO_CABALLERO_2016.pdf
- Alberto, J., Mateos, M. T., Pizarro, S. (2005). *Tecnologías biométricas aplicadas a la seguridad*. Ra - Ma.
- Albertos Cabañas, M. (2011). *Monitorización y clasificación de defectos en rodamientos de bolas: estudio transitorio*. Madrid - España.
- Aldroubi, A., Eden, M., Unser, M. (1993). A family of polynomial spline wavelet transform. *Signal Process*, 141-162.
- Aler Mur, R. (2010). *Universidad Carlos III Madrid - Ingeniería Informática*. Obtenido de kNn y Prototipos: <http://ocw.uc3m.es/ingenieria-informatica/analisis-de-datos/transparencias/KNNyPrototipos.pdf>
- Badawi, A. M. (2007). Hand vein biometric verification Prototype: A testing performance and patterns similarity. *International Journal of Biomedical Sciences*, 141 - 148.
- Bansal, K., Kaur, S. (2015). Finger Vein Recognition Using Minutiae Extraction and Curve Analysis. *International Journal of Science and Research*, 2402 - 2405.
- Barrera, Y., Beltrán, C., Camacho, M. (2016). *Prototipo de un Sistema biométrico integrado para la identificación de personas - fase I*. Obtenido de DOI: 10.13140/RG.2.1.3267.2881
- Beltrán, C., Gómez, M., Flórez, M., Vasquez, Y. (2016). Identificación de personas por medio de imágenes de venas en el dedo índice utilizando las transformadas SIFT y Correlación. *ResearchGate*.
- Bolle, R., Pankati, S. (1998). *Biometrics: Personal Identification in Networked Society*. Kluwer Academic Publisher Norwell.
- Cárdenas Montes, M. (Noviembre de 2015). *Centro de Investigaciones Energéticas Medioambientales y Tecnologías*. Obtenido de k-Nearest Neighbors: http://wwwae.ciemat.es/~cardenas/docs/curso_MD/knn.pdf
- Carrasco García, E., Fuentes García, D. A., García Benítez, C. D., López Hernández, F. (2016). *Sistemas de verificación biométrico vascular*. Obtenido de <http://tesis.ipn.mx/bitstream/handle/123456789/17378/Reporte%20te%CC%81cnico%20TT2012-A008.pdf?sequence=1>
- Carratalá, J., Seco, F. (2004). Clasificadores eficaces basados en algoritmos rápidos de búsqueda del vecino más cercano. *Departamento de lenguajes y sistemas informáticos, Universidad de Alicante*.

- Carratalá, J., Seco, F. (2004). Clasificadores eficaces basados en algoritmos rápidos de búsqueda del vecino más cercano. *Universidad de Alicante. Departamento de lenguajes y sistemas informáticos.*
- Castrillón, O., Giraldo, J., Sarache, W. (2008). Sistema de clasificación Bayesiano basado en múltiples clases. *Sistemas, Cibernética e Informática.*
- Centro de Procesamiento y Análisis Infrarrojo (IPAC). (22 de Febrero de 2001). *Cercano, Mediano y Lejano Infrarrojo.* Obtenido de <http://legacy.spitzer.caltech.edu/espanol/edu/ir/regions/irregions.html>
- Choi, H. S., Chung, C. K., Im, S. K., Kim, S. W., Park, H. M. (2000). Improved Vein Pattern Extracting Algorithm and its Implementation. *Proc. de IEEE ICCE 2000*, 2 - 3.
- Dinstein, I., Haralick, R. M., Shanmugan , K. (1973). Textural features for image classification. *IEEE transaction on System, Man and Cybernetics*, 610-621.
- Duda, R., Hart, P., Stork , D. (2012). *Pattern classification.* New York: John Wiley & Sons.
- E-Cultura Group. (26 de enero de 2015). *definiciona.com.* Obtenido de Definición y etimología de biometría.: <https://definiciona.com/biometria/>
- Escobar Riosalido, J. (2015). *Identificación de imágenes por análisis de texturas utilizando aprendizaje estadístico.* Madrid: Universidad Complutense de Madrid.
- Fan, J., Laine, A. (1996). Frame representation for texture segmentation. *IEEE Transactions on Image Processing*, 771 - 780.
- Flórez, J., España, G., Torres, H. (2014). Estrategia de regresión basada en el método de los k vecinos más cercanos para la estimación de la distancia de falla en sistemas radiales. *Revista Facultad de Ingeniería*, 100 - 108.
- Franklin, S. E., Moskal, L. M., Presutti, M. E. (2001). Supervised classification of multisource satellite image spectral and texture data for agricultural crop mapping in Buenos Aires Province. *Canadian Journal of Remote Sensing*, 679-684.
- Fujitsu Laboratories Ltd. (2003). *Fujitsu.* Obtenido de Fujitsu: <http://www.fujitsu.com/global/about/resources/news/press-releases/2003/0331-05.html>
- Fujitsu Laboratories Ltd. (2006). Whitepaper. Palm Vein Authentication Technology. *Biometric Products.*
- García Laencina , P. J., Gómez, S., Larrey Ruiz, J., Luis, J., Morales Sánchez, J., Verdú Monedero, R. (2008). Algoritmo kNN basado en información mutua para clasificación de patrones con valores perdidos. *Universidad Politécnica de Cartagena.*

- Gonzalez, G. (2015). *catarina.udlap.mx*. Obtenido de Capítulo III: Revisión de la teoría de wavelets: http://catarina.udlap.mx/u_dl_a/tales/documentos/mel/gonzalez_g_ra/capitulo_3.pdf
- Hidalgo, B., Marcotti, A., Mathé, L. (2012). *Método de detección no invasivo utilizando luz infrarroja*. Córdoba.
- Hu, Y.-p., Wang, Z.-y., Yang, X.-p. (2014). Hand vein recognition based on the connection lines of reference point and feature point. *Infrared Physics & Technology*, 110 - 114.
- Im, S. K., Park, H. M., Kim, S. W. (2001). A Biometric Identification System by Extracting Hand Vein Patterns. *Journal of the Korean Physical Society*, 268 - 272.
- ISO/IEC. (2011). *ISO/IEC 19794-9:2011 - Biometric data interchange formats Part 9: Vascular image data*. Obtenido de http://www.iso.org/iso/catalogue_detail.htm?csnumber=56163
- Jain, A. K., Maio, D., Maltoni, D., Wayman, J. L. (2006). Performance evaluation of fingerprint verification systems. *IEEE transactions on pattern analysis and machine intelligence*, 3 - 18.
- Jiménez Guerrero, M. P. (2016). *Extracción de características de textura basada en la transformada Wavelet Discreta*. Sevilla.
- Kono, M., Ueki, H., Umemura, S. (2002). Near-infrared finger vein patterns for personal identification. *Appl. Opt.*, 7429 - 7436.
- Leedham, G., Wang, L. (2006). Near and Far Infrared Imaging for Vein Pattern Biometrics. *Proceedings of the IEEE International Conference on Video and Signal Based Surveillance*, 1 - 6.
- Lewis, N. (22 de enero de 2017). *Classification using k-Nearest Neighbors*. Obtenido de [Classification using k-Nearest Neighbors: http://en.proft.me/2017/01/22/classification-using-k-nearest-neighbors-r/](http://en.proft.me/2017/01/22/classification-using-k-nearest-neighbors-r/)
- Liu, L., Yin, Y., Sun, X. (2011). SDUMLA_HMT: A Multimodal Biometric Database. *Lecture Notes in Computer Science*, 260 - 268.
- Mallat, S. G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Rrepresentatio. *IEEE transactions on pattern analysis and machine intelligence*, 674-693.
- MathWorks. (2017). *MathWorks Documentation*. Obtenido de Statistics and Machine Learning Toolbox - classify: <https://es.mathworks.com/help/stats/classify.html>

- MathWorks. (2017). *MathWorks Documentation*. Obtenido de Wavelet Toolbox - Wavelet Families: <https://es.mathworks.com/help/wavelet/ug/wavelet-families-additional-discussion.html>
- Minitab. (2016). *Análisis multivariado*. Obtenido de ¿Qué es análisis discriminante cuadrático?: <http://support.minitab.com/es-mx/minitab/17/topic-library/modeling-statistics/multivariate/discriminant-analysis/what-is-quadratic-discriminant-analysis/>
- Minitab. (2016). *Análisis Multivariado*. Obtenido de ¿Qué es análisis discriminante lineal?: <http://support.minitab.com/es-mx/minitab/17/topic-library/modeling-statistics/multivariate/discriminant-analysis/what-is-linear-discriminant-analysis/>
- Miura, N., Miyatake, T., Nagasaka, A. (2004). Feature Extraction of finger vein patterns based on repeated line tracking and its Application to Personal Identification. *Machine Vision and Applications*, 194 - 203.
- Morales , E. (2009). *ccc.inaoep.mx*. Obtenido de Descubrimiento de conocimiento en base de datos: <http://ccc.inaoep.mx/~emorales/Cursos/KDD/principal.html>
- Motato Toro, Ó. F., Loaiza Correa, H. (2009). Identificación biométrica utilizando imágenes infrarrojas de la red vascular de la cara dorsal de la mano. *Revista Ingeniería e Investigación*, 90 - 100.
- Orea , S., Vargas, A., Alonso, M. (2005). Minería de datos: predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos. *Universidad Tecnológica de Izúcar de Matamoros*, 33.
- Ortega, J. (2015). *Universidad Autónoma de Madrid*. Obtenido de Sistemas biométricos e identificación personal: http://arantxa.ii.uam.es/~jortega/Evaluacion_Biometricos_v1.pdf
- Presutti, M. (2004). La matriz de co-ocurrencia en la clasificación multispectral: Tutorial para la enseñanza de medidas texturales en cursos de grado universitario. *4a Jornada de Educacao em Sensoramiento Remoto no Âmbito do Mercosul*.
- Rodríguez, O. (2014). *oldemarrodriguez*. Obtenido de Aprendizaje Supervisado : http://www.oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentaci%C3%B3n_-_Análisis_Discriminante.26282559.pdf
- Schaffer, R. (2005). *Biometric identification using subcutaneous vein patterns*. Obtenido de Biotech-weblog: http://www.biotech-weblog.com/50226711/biometric_identification_using_subcutaneous_vein_patterns.php
- Shimizu, K. (1992). Optical trans-body imaging: feasibility of non-invasion CT and functional imaging of living body. *Jpn. J. of Medicine Philosophica*, 620 629.

- Silveyra, J. O. (2006). Investigación científica del delito. Sistemas de identificación humana.
- Tocoma, G. (2015). *Introducción al reconocimeinto de objetos*. Obtenido de <http://alojamientos.us.es/gtocomap/pid/tema7.pdf>
- UNICEN. (s.f.). *Universidad Nacional del Centro de la Provincia de Buenos Aires*. Obtenido de Introducción a la transformada wavelet - Descomposición de señales: <http://www.exa.unicen.edu.ar/escuelapav/cursos/wavelets/apunte.pdf>
- Universidad Juárez Aunónoma de Tabasco. (2006). *Avances en informática y sistemas computacionales - Tomo I*. Villahermosa, Tabasco: CONAIS.
- Universidad Politécnica de Valencia. (Diciembre de 2006). *user.dsic.upv.es*. Obtenido de Clasificadores: <http://users.dsic.upv.es/~ajuan/oracle/classC.html>
- Willmore, M. R. (1994). Infrared imaging and patten recognition system. *USPatent No 5,351,303*.