

U

P

T

**Análisis de la distribución de venas
en la mano a partir de imágenes IR
para el reconocimiento de
personas**

por

Raúl Castro Ortega

Tesis sometida como requisito parcial para
obtener el grado de

MAESTRO EN COMPUTACIÓN ÓPTICA

en la

**UNIVERSIDAD POLITÉCNICA DE
TULANCINGO**

Mayo 2015

Tulancingo de Bravo, Hidalgo.

Supervisada por:

Dra. Carina Toxqui Quitl

Dr. José Francisco Solís Villarreal

©UPT

El autor otorga a la UPT el permiso de reproducir y
distribuir copias en su totalidad o en partes de esta tesis.



Dedicatoria

La presente Tesis está dedicada principalmente a mis padres José Delfino Castro Ortíz y María Carmen Guadalupe Ortega Islas, porque siempre estuvieron a mi lado brindándome su apoyo y sus consejos para hacer de mí una mejor persona y que me han impulsado a seguir adelante.

A mis amigos y compañeros *Martín, Edwin, Victoria, Laura y Jorge*, por su amistad y compañía durante la maestría y tesis con quienes compartí buenos momentos.

A todas aquellas personas que creyeron en mí.

A mis profesores.

A todos,

¡Muchas Gracias!

*Raúl Castro Ortega.
Mayo, 2015*

Agradecimientos

Al Laboratorio de Óptica y Visión por Computadora de la Universidad Politécnica de Tulancingo por las facilidades brindadas para desarrollar este proyecto de Tesis de Maestría.

Al Consejo Nacional de Ciencia y Tecnología CONACyT por el apoyo a través de la beca nacional 303672 y la beca mixta 290842.

A mis Asesores, Dra. Carina Toxqui Quitl y Dr. José Francisco Solís Villarreal por haber dirigido este proyecto de Tesis.

Al Dr. Gabriel Cristóbal Pérez del Instituto de Óptica del departamento de Imagen y Visión del Consejo Superior de Investigaciones Científicas (CSIC) de Madrid, por el apoyo durante mi estancia de investigación dentro de su grupo de trabajo.

Al Dr. José Víctor Marcos Martín por su apoyo, consejos y aportaciones a esta investigación.

A mis sinodales, Dr. Alfonso Padilla Vivanco y Dr. Juan Carlos Valdiviezo Navarro por las valiosas sugerencias para mejorar este trabajo.

Al Maestro Gerardo Téllez Reyes, Rector de la Universidad Politécnica de Tulancingo por el apoyo otorgado para la realización de esta tesis.



Prefacio

El área de biometría ha crecido vertiginosamente en las últimas décadas, desde el uso de la huella dactilar, hasta emplear varias medidas físicas o de comportamiento. Estas características son usadas como datos de identificación personal, además, al ser más fáciles de usar ya que el usuario no tiene nada que recordar, cambiar o perder.

Por otro lado, dada la creciente necesidad de mejorar la seguridad, la familia biométrica ha comenzado a desplazar a los sistemas de identificación clásicos por nuevos y mejores sistemas de alto rendimiento. Por ejemplo, los sistemas de identificación basados en huellas dactilares, han sido complementados o sustituidos, debido a la facilidad con la que pueden ser copiadas o alteradas y, por lo tanto, pueden ser usadas para suplantación de la identidad. Entre los nuevos sistemas de identificación basados en imágenes, está el que utiliza la estructura interna de las venas de la palma o dorso de la mano.

El patrón de venas es utilizado como una biometría para propósitos de acceso, identificación y autenticación, el cual es más seguro que los métodos clásicos de identificación. Ya que es más difícil de suplantar debido a que el patrón se consigue en presencia de la persona y se requiere que este viva. Además, estos patrones pueden ser útiles en el campo de la salud. En este campo, existen personas que tienen ocultas sus venas, y, cuando los médicos realizan una venopunción, les toma cierta cantidad de tiempo para encontrar las venas y suministrar el medicamento o conseguir una muestra sanguínea. Esto, causa un alto nivel de estrés y ansiedad tanto a pacientes como médicos.

Este trabajo se encuentra organizado de la siguiente manera: en el Capítulo 1, se enlistan los objetivos del proyecto de Tesis, los antecedentes, el estado del arte del reconocimiento de personas a través de las venas de la mano. En el Capítulo 2, se expone la teoría con respecto a las técnicas de procesamiento digital de imágenes con propósitos de mejoramiento, realce y segmentación. En el Capítulo 3, se describen los métodos de extracción de características y las técnicas de clasificación. En el Capítulo 4, se muestra el sistema de visión por computadora implementado, así como su interfaz para la adquisición y procesamiento de imágenes en el infrarrojo cercano. En el Capítulo 5 se muestran los resultados de clasificación para 4 bases de datos diferentes. Las conclusiones generales se enuncian en el Capítulo 6.

Resumen

Se implementó un sistema de visión por computadora para la adquisición y análisis de imágenes en el infrarrojo cercano de la palma y dorso de la mano, con propósitos de reconocimiento y autenticación de la identidad de una persona. El sistema de adquisición de imágenes, está compuesto por una cámara 2CCD Multispectral de la marca jAi con sensibilidad en el visible e infrarrojo y una fuente de iluminación con un pico de emisión en los 880 *nm*. Esta radiación puede penetrar y ser absorbida mayormente por la desoxihemoglobina que se encuentra presente en la sangre de las venas.

Los algoritmos para el mejoramiento de imágenes están basados en: a) Ecuilibración del Histograma, b) Contraste Limitado de la Ecuilibración Adaptativa del Histograma y c) Máscara de Desenfoque. Así mismo, se usó el filtraje espacial con el propósito de remover información innecesaria como pliegues y vellos de la palma y dorso respectivamente. Las técnicas de binarización adaptativa y las operaciones de morfología matemática fueron implementadas con el objetivo de obtener únicamente la distribución del patrón de venas.

En la etapa de extracción de características se utilizaron los métodos de a) Momentos Geométricos, b) Invariantes de Hu y c) el denominado Descriptor de Venas, propuesto en este trabajo. La tarea de clasificación se realiza mediante los métodos de K-Vecinos Próximos (K-nn) y Redes Neuronales Artificiales (RNA).

Cuatro bases de datos son usadas para el análisis del rendimiento de los algoritmos. La primera de ellas, pertenece a la Universidad Politécnica de Hong Kong, la segunda, al Instituto de Control e Ingeniería de la Información de la Universidad Tecnológica de Polonia, la tercera, al Instituto de Automatización de la Academia de Ciencias China, y la cuarta, una base de datos propia obtenida en el Laboratorio de Óptica y Sistemas de Visión de la Universidad Politécnica de Tulancingo.

Finalmente, se presenta un prototipo de sistema para la adquisición y extracción del patrón de venas de la mano. La interfaz que controla la cámara y procesa las imágenes adquiridas fue programada en MATLAB.

Abstract

An image acquisition system was implemented in order to acquire digital images in the near infrared from the palm and back of the hand with purposes of recognition and authentication of the identity of the people. The image acquisition system consists of a Multispectral 2CCD camera from the jAi trademark and a light source with a peak emission in the 880 *nm*. This radiation can penetrate and can be strongly absorbed by the desoxyhemoglobin that is present in the blood of the veins.

The algorithms for image enhancement are based on: a) Histogram Equalization, b) Contrast Limited Adaptive Histogram Equalization and c) Unsharp Masking. Likewise, spatial filtering is used in order to remove unnecessary information from the folds and hair in the palm and back of the hand respectively. Adaptive binarization techniques and mathematical morphology operations were implemented in order to obtain only the vein pattern distribution.

In the stage of feature extraction were used the methods of a) Geometric Moments, b) Hu Invariants and the called Vein Descriptor, proposed in this work. The classification task is performed by the methods of K-Nearest Neighbors (K-nn) and Artificial Neural Networks (ANN).

Four databases are used for the performance analysis of the algorithms. The first database used is owned of the Hong Kong Polytechnic University, the second database at the Institute of Control and Information Engineering of the Technological University of Poland, the third, at the Institute of Automation in the Chinese Academy of Sciences, and the fourth, own database obtained in the Laboratory of Optics and Vision Systems at the Polytechnic University of Tulancingo.

Finally, a prototype system is presented for acquisition and extraction of the vein patterns of the hand. The interface controls the camera for image acquisition. It was developed in MATLAB.

Índice general

1. Introducción	1
1.1. Introducción	1
1.2. Planteamiento del problema	4
1.3. Justificación	5
1.4. Objetivos	7
1.5. Antecedentes	8
1.6. Estado del arte	12
1.7. Aportaciones	16
2. Procesamiento Digital de Imágenes	21
2.1. Introducción	21
2.2. Mejoramiento de imágenes	23
2.3. Histograma de una imagen	26
2.3.1. Ecualización del Histograma	27
2.3.2. Ecualización Adaptativa del Histograma (AHE)	29
2.3.3. Contraste Limitado de la Ecualización Adaptativa del Histograma (CLAHE)	31
2.3.4. Máscara de desenfoque (Unsharp Masking)	33
2.4. Filtrado Espacial	35
2.4.1. Filtro pasa-bajas	36
2.5. Segmentación de imágenes	37
2.5.1. Umbralización	38
2.5.2. Binarización	40
2.5.3. Binarización Adaptativa	41
2.6. Región de Interés (ROI)	42
2.7. Morfología matemática	44
2.7.1. Dilatación	44
2.7.2. Adelgazamiento	46
2.8. Conclusiones	47

3. Técnicas de extracción de características y métodos de clasificación	51
3.1. Introducción	51
3.2. Patrones y Patrón de clases	53
3.3. Momentos geométricos	54
3.4. Descriptores de venas	57
3.5. Clasificadores	58
3.6. Redes Neuronales Artificiales (RNA)	59
3.6.1. La neurona biológica	59
3.6.2. Neurona artificial	60
3.6.3. El perceptrón	61
3.6.4. Tipos de redes neuronales artificiales	62
3.6.5. Función de activación	63
3.6.6. Regla de aprendizaje del perceptrón	65
3.7. La red backpropagation	67
3.7.1. Regla delta generalizada	68
3.8. K-Vecinos Próximos (KNN)	70
3.8.1. Algoritmo	71
3.8.2. Elección del k	73
3.9. Conclusiones	74
4. Sistema de adquisición de imágenes	76
4.1. Introducción	76
4.2. Sistema para la captura de imágenes en el NIR	82
4.2.1. Interfaz para la captura de imágenes	85
4.3. Imágenes adquiridas con el sistema de la UPT	90
4.3.1. Base de datos UPT preprocesada	91
4.4. Imágenes adquiridas con el sistema de la UPHK	96
4.4.1. Base de datos PolyU Multispectral Palmprint Database preprocesada	98
4.5. Imágenes adquiridas con el sistema del CIE de la UTP	100
4.5.1. Base de datos Vein Pattern Database preprocesada	102
4.6. Imágenes adquiridas con el sistema CASIA	106
4.6.1. Base de datos CASIA Multi-Spectral Palmprint Image Database preprocesada	107
4.7. Conclusiones	112
5. Análisis de resultados	115
5.1. Introducción	115
5.2. Resultados de clasificación para la base de datos UPT	116
5.3. Resultados de clasificación para la base de datos de la UPHK	119
5.4. Resultados de clasificación para la base de datos VPD	121

5.5. Resultados de clasificación para la base de datos CASIA	124
5.6. Conclusiones	128
6. Conclusiones	130
6.1. Trabajos derivados de la Tesis	139

Índice de figuras

1.1.	Huellas dactilares impresas en a) Tablet, b) Copa de vidrio, c) Espejo, d) Vaso de vidrio, e) Acercamiento a una huella sobre una superficie y f) Huella copiada sobre un dedo falso para suplantación de identidad.	4
1.2.	Aplicaciones basadas en el reconocimiento del patrón de venas. a) Inicio de sesión en computadoras y b) pago de servicios.	5
1.3.	Venopunción. a) Método tradicional, b) AccuVein, c) Visor de venas portátil, d) VeinViewer Flex, e) Venopunción por AccuVein y f) Venopunción por VeinViewer Flex.	6
1.4.	Ficha de filiación: Sir Francis Galton fotografiado por Bertillon en 1893 [17]. . .	9
1.5.	Ejemplos de rasgos biométricos que pueden ser usados para identificación. Los rasgos físicos (estáticos) incluyen las huellas dactilares, iris, geometría de cara, mano u oreja. Los rasgos de comportamiento (dinámicos), incluyen la firma, dinámica de teclado y marcha [8].	11
1.6.	Ejemplo de imágenes del patrón de venas de los dedos de la mano [23].	12
1.7.	Ejemplo de imágenes del patrón de venas de la palma de la mano [30].	12
1.8.	Ejemplo de imagen del patrón de venas del dorso de la mano [33].	13
2.1.	Un ejemplo del proceso de adquisición de una imagen digital. (a) Fuente de iluminación (energía), (b) Un elemento de una escena, (c) Sistema de adquisición de imágenes, (d) Proyección de la escena en el plano imagen y (e) Imagen digitalizada.	22
2.2.	Una vecindad de 3×3 sobre un punto en una imagen.	24
2.3.	Ejemplo de transformaciones básicas en niveles de gris en el que se observa un incremento del contraste de la imagen. (a) Imagen Original, (b) Negativo, (c) Logaritmo, (d) Inversa de logaritmo, (e) Potencia y (f) Raíz.	25
2.4.	Cuatro tipos básicos de imágenes: (a) oscura, (b) brillante, (c) bajo contraste y (d) alto contraste con sus histogramas correspondientes.	26
2.5.	(a) Imágenes de la Figura 2.4, (b) Resultados de aplicar ecualización del histograma y (c) Histograma correspondiente.	28
2.6.	Función del método AHE.	29

2.7. Resultado de aplicar Ecuación Adaptativa del Histograma con una vecindad de tamaño 5 píxeles. Nótese que en la imagen (b) el fondo ha sido realzado pero hay presencia de ruido.	30
2.8. Equilibrio del límite de corte en el histograma.	31
2.9. Resultado de aplicar Contraste Limitado de la Ecuación Adaptativa del Histograma con una vecindad de tamaño 5 píxeles. Nótese que en la imagen (b) el fondo ha sido realzado y la presencia de ruido ha sido disminuida.	32
2.10. Ejemplo de mejoramiento de imágenes usando una máscara de desenfoque. (a) Imagen original, (b) Imagen nítida con valores de Cantidad y Radio de 10 y 5 píxeles (c) Imagen más nítida con valores de Cantidad y Radio de 5 y 10 píxeles y (d) Imagen muy nítida con valores de Cantidad y Radio de 10 y 20 píxeles.	34
2.11. Mecanismo del filtraje espacial. La imagen grande muestra una máscara de 3×3 y la sección de la imagen bajo esta; la sección de la imagen se muestra desplazada debajo de la máscara para facilitar su legibilidad.	35
2.12. Filtro pasa bajas de tamaño 3×3 y 5×5	36
2.13. Ejemplo de filtrado pasa bajas (promedio). (a) Imagen original, (b) imagen filtrada con una máscara de 3×3 píxeles, (c) imagen filtrada con una máscara de 5×5 píxeles y (d) imagen filtrada con una máscara de 10×10 píxeles.	37
2.14. (a) Histograma ideal y (b) Histograma no ideal.	38
2.15. Ejemplo de umbralización basado en el histograma. (a) Imagen original y (b) imagen umbralizada.	39
2.16. Ejemplo de binarización (a) Imagen original, (b) imagen con binarización local y (c) imagen con binarización global.	40
2.17. Ejemplo de binarización adaptativa.	41
2.18. Ejemplo de extracción de la ROI. (a) Imagen original, (b) Centroides de la imagen, (c) Cálculo del cuadro que conformará la ROI y (d) Región de Interés extraída.	43
2.19. Ejemplo de dilatación. (a) Imagen original y (b) imagen dilatada con un elemento estructural de tamaño 3×3	45
2.20. Ejemplo de dilatación. (a) Imagen original y (b) imagen adelgazada con un elemento estructural de tamaño 3×3	46
3.1. De izquierda a derecha para (a) Persona 1 y (b) Persona 2: imagen original, escalada, rotada y desplazada y (c) gráfica de los descriptores invariantes de Hu de 8 imágenes para 2 personas diferentes.	56
3.2. Extracción de descriptores del patrón de venas de la (a) Persona 1, (b) Persona 2.	57
3.3. Elementos de una neurona biológica.	59
3.4. Representación gráfica del perceptrón.	61
3.5. Clases linealmente separables.	62
3.6. Clasificación de las Redes Neuronales Artificiales [9].	63
3.7. Función Sigmoidal.	64
3.8. Mapas de decisión para las operaciones (a) AND y (b) OR	66

3.9. En la función XOR no es posible obtener una recta que separe las dos clases. . .	66
3.10. Solución del problema de la función XOR.	67
3.11. Algoritmo K -nn para la clasificación de las clases círculo, cuadrado y triángulo. .	73
4.1. Imágenes de las características que afectan la visibilidad de las venas en las manos: (a) Pliegues de la piel de la palma, (b) Vello del dorso, (c) Cicatrices, (d) Lesiones, lunares o verrugas, (e) Pigmentación de la piel y (f) Falta de visibilidad a simple vista.	77
4.2. Imágenes en el infrarrojo cercano de: (a) Palma, (b) Dorso y (c) Muñeca de la mano.	78
4.3. Niveles de absorción de la oxihemoglobina, desoxihemoglobina y el agua.	79
4.4. Profundidad de penetración de la luz en la piel a diferentes longitudes de onda.	79
4.5. Sistema de adquisición de imágenes del dorso de la mano.	80
4.6. Sistema de adquisición de imágenes del dorso de la mano.	81
4.7. Cámara 2CCD Multiespectral jAi de escaneo progresivo.	82
4.8. Respuesta espectral total incluyendo prisma y sensor para los canales (a) color y (b) infrarrojo.	83
4.9. Diagrama conceptual del prisma dicróico que separa dos canales (a) Visible, (b) Infrarrojo y (c) Ubicación del prisma en el interior de la cámara.	83
4.10. Sistema de adquisición de imágenes de la palma de la mano.	84
4.11. Sistema de adquisición de imágenes del dorso de la mano.	85
4.12. Inicio de la Interfaz para la adquisición y procesamiento de las imágenes.	86
4.13. Previsualización del área para adquisición del patrón de venas.	86
4.14. Imagen capturada de la distribución del patrón de venas de una persona.	87
4.15. Visualización de las técnicas disponibles de preprocesamiento para la imagen capturada.	87
4.16. Proceso de Delimitación, Mejoramiento y Segmentación de una imagen utilizando las Técnicas Clásicas.	88
4.17. Proceso de Delimitación, Mejoramiento y Segmentación de una imagen utilizando la técnica de Contraste Limitado de la Ecuación Adaptativa del Histograma (CLAHE).	89
4.18. Proceso de Delimitación, Mejoramiento y Segmentación de una imagen utilizando la técnica de Máscara de Desenfoque.	89
4.19. Imágenes de la palma de la mano adquiridas con el sistema de la Figura 4.10. . .	90
4.20. Imágenes del dorso de la mano adquiridas con el sistema de la Figura 4.11. . . .	90
4.21. (a) Imagen original, (b) Centroides y detección de la región de interés (ROI), (c) Extracción de ROI, (d) Filtrado espacial, (e) Ecuación del histograma, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación e imagen final con el patrón de venas segmentado.	92

4.22. (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Contraste limitado de la ecualización adaptativa del histograma, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.	93
4.23. (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Máscara de desenfoque, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.	94
4.24. Imágenes mejoradas usando: (a) Ecualización del Histograma, (b) Contraste Limitado de la Ecualización Adaptativa del Histograma y (c) Máscara de desenfoque.	95
4.25. Sistema de adquisición de imágenes multispectrales de la palma de la mano [7].	97
4.26. Imágenes de la palma de la mano de 2 personas diferentes adquiridas en los canales: (a) Rojo (R), (b) Verde (G), (c) Azul (B) y (d) Infrarrojo cercano (NIR) [7].	97
4.27. (a) Imagen original, (b) Filtrado espacial, (c) Ecualización del histograma, (d) Binarización adaptativa, (e) Adelgazamiento y (f) Dilatación e imagen final con el patrón de venas segmentado.	98
4.28. (a) Imagen original, (b) Contraste limitado de la ecualización adaptativa del histograma, (c) Filtrado espacial, (d) Binarización adaptativa, (e) Adelgazamiento y (f) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.	98
4.29. (a) Imagen original, (b) Máscara de Desenfoque, (c) Filtrado espacial, (d) Binarización adaptativa, (e) Adelgazamiento y (f) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.	99
4.30. Imágenes mejoradas usando: (a) Ecualización del Histograma, (b) Contraste Limitado de la Ecualización Adaptativa del Histograma y (c) Máscara de desenfoque.	99
4.31. Dispositivo utilizado para adquirir imágenes de la palma y muñeca de la mano. .	100
4.32. Imágenes de la palma de la mano de 4 personas diferentes [8].	101
4.33. Imágenes de la muñeca de la mano de 4 personas diferentes [8].	101
4.34. (a) Imagen original, (b) Centroides y detección de la región de interés (ROI), (c) Extracción de ROI, (d) Filtrado espacial, (e) Ecualización del histograma, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación e imagen final con el patrón de venas segmentado.	102
4.35. (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Contraste limitado de la ecualización adaptativa del histograma, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.	103
4.36. (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Máscara de desenfoque, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.	104

4.37. Imágenes mejoradas usando: (a) Ecualización del Histograma, (b) Contraste Limitado de la Ecualización Adaptativa del Histograma y (c) Máscara de desenfoque. 105

4.38. Diseño conceptual del dispositivo multispectral de adquisición de imágenes de la palma de la mano [9]. 106

4.39. Imágenes de la palma de la mano adquiridas en longitudes de onda de: (a) 460 nm, (b) 630 nm, (c) 700 nm, (d) 850 nm, (e) 940 nm y (f) Luz blanca [9]. 107

4.40. (a) Imagen original, (b) Centroides y detección de la región de interés (ROI), (c) Extracción de ROI, (d) Filtrado espacial, (e) Ecualización del histograma, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación e imagen final con el patrón de venas segmentado. 108

4.41. (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Contraste limitado de la ecualización del histograma, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado. 109

4.42. (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Máscara de desenfoque, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado. 110

4.43. Imágenes mejoradas usando: (a) Ecualización del Histograma, (b) Contraste Limitado de la Ecualización Adaptativa del Histograma y (c) Máscara de desenfoque. 111

6.1. Ejemplo de aumento de personas y disminución del porcentaje de clasificación. . . 133

6.2. Porcentajes de clasificación obtenidos para las 4 bases de datos con K Vecinos Próximos. 134

6.3. Porcentajes de clasificación obtenidos para las 4 bases de datos con Redes Neuronales Artificiales. 134

6.4. Tiempos de cómputo para la base de datos PolyU Multispectral Palmprint. . . . 135

6.5. Tiempos de cómputo para la base de datos UPT con selección de la ROI. 136

6.6. Tiempos de cómputo para la base de datos UPT sin selección de la ROI. 136

6.7. Tiempos de cómputo para la base de datos Vein Pattern Database con selección de la ROI. 137

6.8. Tiempos de cómputo para la base de datos Vein Pattern Database sin selección de la ROI. 137

6.9. Tiempos de cómputo para la base de datos CASIA con selección de la ROI. . . . 138

6.10. Tiempos de cómputo para la base de datos CASIA sin selección de la ROI. . . . 138

Índice de tablas

1.1. Comparativa de porcentajes de clasificación de diversos trabajos relacionados al reconocimiento de personas a través de la distribución del patrón de venas.	15
4.1. Resumen de las características más importantes de las bases de datos analizadas. Simbología: (DD) Dorso Derecho, (DI) Dorso Izquierdo, (PD) Palma Derecha, (PI) Palma Izquierda, (***) Sin Datos.	113
5.1. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	116
5.2. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).117	
5.3. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	117
5.4. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).117	
5.5. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	118
5.6. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).118	
5.7. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	119
5.8. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).119	
5.9. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	120
5.10. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).120	
5.11. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	120
5.12. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).120	
5.13. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	121
5.14. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).122	
5.15. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	122

5.16. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).	122
5.17. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	123
5.18. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).	123
5.19. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	125
5.20. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).	125
5.21. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	126
5.22. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).	126
5.23. Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.	127
5.24. Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).	127
5.25. Resumen de las características más importantes de las bases de datos analizadas. Simbología: (DD) Dorso Derecho, (DI) Dorso Izquierdo, (PD) Palma Derecha, (PI) Palma Izquierda, (***) Sin Datos.	128
5.26. Comparativa de porcentajes de clasificación de diversos trabajos relacionados al reconocimiento de personas a través de la distribución del patrón de venas. .	129

Capítulo 1

Introducción

1.1. Introducción

Actualmente, los sistemas empleados para el reconocimiento de personas han basado su funcionamiento en sensores digitales. Estos dispositivos pueden adquirir información de números de identificación personal (PIN's), tarjetas electrónicas, contraseñas, llaves, entre otros. Estos métodos de identificación ofrecen a los usuarios seguridad limitada debido a que pueden ser fácilmente violados [1-4].

Otros dispositivos, utilizan las características fisiológicas o de comportamiento de las personas para propósitos de identificación. Éstas características son útiles para análisis biométrico, debido a que tienen un alto nivel de seguridad [4], [5]. Además, este tipo de identificación puede minimizar las desventajas con respecto a los sistemas clásicos de identificación [1].

Una ventaja es que la identificación biométrica de una persona es única e inalterable, haciéndola más exacta, cómoda y segura para el usuario. Por otra parte, la huella digital, el iris del ojo, la forma de la cara y la oreja, la voz, la geometría de la mano, la marcha, la dinámica de teclado y la firma, son usados por muchos sistemas como una biometría de referencia inmediata.

Sin embargo, este tipo de sistemas tienen algunos inconvenientes cuando se necesita de un riguroso método de identificación de las personas debido a no existe todavía un algoritmo capaz de reconocer a una persona con un grado máximo de precisión [1], [2], [4]. Esto se debe a que, conforme se incrementan las medidas y técnicas de seguridad, también se incrementan las técnicas para ocultar o suplantar la identidad.

Dada la creciente necesidad de mejorar la seguridad y controles de acceso, la familia biométrica ha comenzado a desplazar los sistemas de identificación clásicos por nuevos y mejores sistemas de alto rendimiento basados en el proceso de autenticación de nuevos rasgos.

Algunos sistemas de identificación basados en imágenes, utilizan la estructura interna de las venas de la palma o dorso de la mano. Estas características son usadas como datos de identificación personal ya que son indicadores biométricos fisiológicos.

Un número de características biométricas se utilizan en varias aplicaciones. Cada biometría tiene sus ventajas y desventajas y, por lo tanto, la elección de un rasgo biométrico para una aplicación en particular depende de una variedad de cuestiones, además del rendimiento del algoritmo de clasificación.

Jain et al. [6], tiene identificados 7 factores que determinan el rasgo físico o de comportamiento apropiado para ser usado en una aplicación de biometría [1], [7].

- a) **Universalidad:** Todas las personas que tengan acceso a la aplicación deben de poseer ese rasgo.
- b) **Unicidad:** El rasgo debe de ser lo suficientemente diferente entre los individuos que comprenden la población.
- c) **Permanencia:** el rasgo biométrico de una persona deber ser lo suficientemente invariante a lo largo de un periodo de tiempo con respecto al algoritmo de clasificación. Un rasgo que cambia significativamente con el tiempo, no es una biometría útil.
- d) **Medibilidad:** Esto debe de ser posible para adquirir y digitalizar el rasgo biométrico usando dispositivos adecuados que no causen inconvenientes para los individuos. Además, los datos adquiridos deben de ser adecuados para procesamiento con el objetivo de extraer conjuntos de características representativas.
- e) **Rendimiento:** La precisión del reconocimiento y los recursos requeridos para lograr la precisión deben cumplir con las restricciones impuestas por la aplicación.
- f) **Aceptabilidad:** Las personas de la población que utilizarán la aplicación deben estar dispuestos a presentar su rasgo biométrico al sistema.
- g) **Falsificación:** Se refiere a la facilidad con la cual el rasgo de una persona pueda ser imitado utilizando artefactos (por ejemplo, dedos falsos), en el caso de rasgos físicos, e imitación, en el caso de los rasgos de comportamiento.

No se espera que una sola biometría responda efectivamente a todos los requerimientos impuestos por todas las aplicaciones. En otras palabras, ninguna biometría es *ideal*, pero **un conjunto de ellas son admisibles**.

La relevancia de una biometría en específico para una aplicación, esta establecida dependiendo de la naturaleza y requerimientos de la aplicación, y las propiedades de las características propias de la biometría [8].

Debido a que la distribución de venas en la palma y dorso de la mano de cada persona son diferentes, el método biométrico ofrece ventajas para encontrar muchas características, las cuales las hacen únicas. Este hecho es usado para adquirir imágenes de la mano donde el flujo sanguíneo es registrado.

Con esto, es imposible suplantar la identidad a través del uso de técnicas de suplantación. Además, es difícil de duplicar dado que las venas se encuentran ocultas debajo de la piel y son en su mayoría difíciles de observar a simple vista, además, son menos susceptibles a daños y contaminación [2], [9].

Además, algunos estudios muestran que la distribución de venas en la mano son únicas para cada persona, incluso para el caso de gemelos, igualmente, esta distribución es diferente en ambas manos. También, el patrón de venas no cambia con el crecimiento, únicamente se amplía manteniendo el mismo patrón [5].

La luz en el infrarrojo cercano (700 - 900 *nm*) se utiliza a menudo en sistemas de adquisición de imágenes de las venas de la palma de la mano debido a que puede penetrar la piel y ser absorbida en gran medida por la desoxihemoglobina presente en las venas [1], [6].

En anatomía, las venas de la palma forman una red a lo largo de la palma y esta red no puede estar rotas a menos que algunas venas sufran una ruptura [7] a causa de un accidente.

Algunas de las etapas típicas de un sistema biométrico basado en el reconocimiento a partir del patrón de venas son:

- Adquisición de imágenes.
- Mejoramiento de imágenes.
- Segmentación.
- Extracción de características.
- Clasificación / Validación.

1.2. Planteamiento del problema

La identificación de personas se ha vuelto hoy en día una prioridad para algunas organizaciones y empresas, debido a la facilidad con la que se puede llegar a suplantar la identidad de una persona. Los sistemas biométricos actuales tienen ciertos problemas, más cuando se necesita una identificación rigurosa de los individuos, ya que no existe aún un algoritmo capaz de reconocer con un grado máximo de exactitud a las personas.

Estos sistemas utilizan para su funcionamiento sensores digitales para el reconocimiento de personas, estos dispositivos pueden adquirir información contenida al ingresar un número de identificación, una clave o password o utilizar algún medio físico como llaves o tarjetas electrónicas. Sin embargo, estos métodos para autenticar la identidad de una persona ofrecen una seguridad limitada debido a su facilidad de ser violados, copiados y suplantados.

Algunos dispositivos, usan datos de comportamiento o fisiológicos de las personas como medio de identificación siendo útiles para realizar reconocimiento biométrico ya que presentan un alto nivel de seguridad disminuyendo las desventajas de los sistemas tradicionales o clásicos de identificación. Este tipo de sistemas presentan inconvenientes más cuando se necesita de una identificación rigurosa dado que no hay aún un algoritmo para identificar con una precisión máxima a una persona.

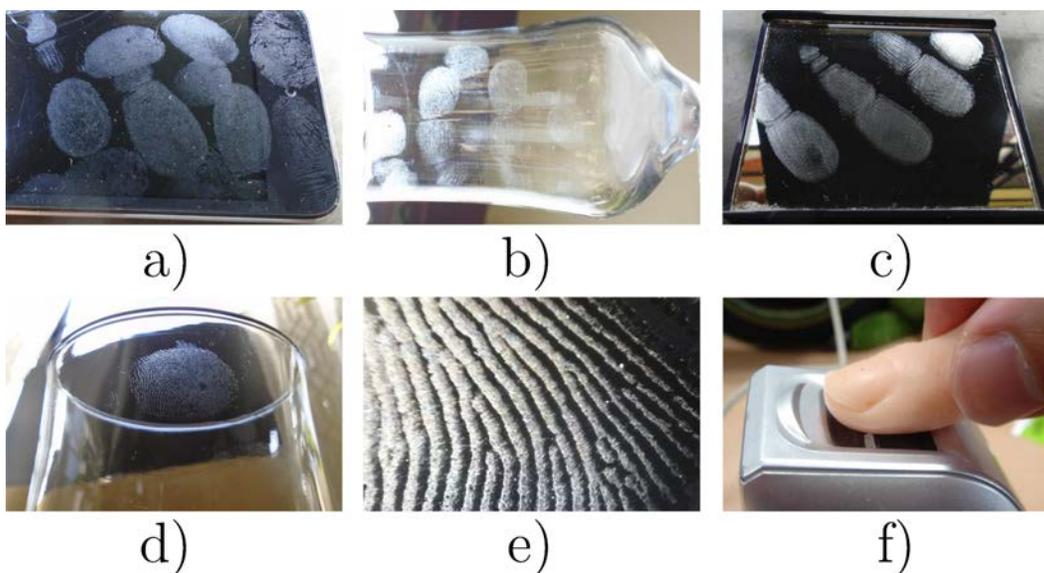


Figura 1.1: Huellas dactilares impresas en a) Tablet, b) Copa de vidrio, c) Espejo, d) Vaso de vidrio, e) Acercamiento a una huella sobre una superficie y f) Huella copiada sobre un dedo falso para suplantación de identidad.

Por ejemplo, las huellas dactilares pueden ser fácilmente falsificadas o copiadas debido a que cuando las personas tocan una superficie, éstas pueden dejar impresas sus huellas sobre dicha superficie y mediante técnicas o métodos de falsificación copiar y suplantar la identidad de una persona. Esto se puede observar en la Figura 1.1 en las que intencionalmente se imprimieron las huellas de una persona sobre una tablet, un vaso, una copa de vidrio y un espejo.

1.3. Justificación

Identificar a una persona por sus características físicas (biometría estática) es un área que ha crecido en las últimas décadas, el uso de dispositivos que identifican a las personas por sus huellas digitales, palmas, reconocimiento de retina, entre otras, se ha difundido en gran parte de las empresas y/u organizaciones en todo el mundo. Sin embargo, aún con estos sistemas biométricos, se tienen problemas de confianza cuando se necesita una identificación rigurosa de los individuos.

Por ejemplo, en el momento de realizar una transacción bancaria de manera autónoma, es decir, un sistema que certifique con un alto índice de certeza a un individuo para poder realizar una acción importante. Otros ejemplos incluyen el acceso a computadoras personales así como sistemas de pago más seguros como los que se muestran en la Figura 1.2.



Figura 1.2: Aplicaciones basadas en el reconocimiento del patrón de venas. a) Inicio de sesión en computadoras y b) pago de servicios.

El reconocimiento a partir del patrón de venas es uno de los enfoques principales en el campo de la identificación biométrica en años recientes. Dicha técnica apareció en 1990, y ha

sido popular desde el año 2000 por las ventajas de seguridad que ofrece. Aparte de los propósitos de seguridad, la distribución del patrón de venas puede ser usada en el campo de la salud.

En este campo, existen personas quienes tienen ocultas sus venas, o, cuando los médicos realizan una venopunción a estas personas, les toma cierta cantidad de tiempo para encontrar las venas y suministrar el medicamento o conseguir una muestra sanguínea.

Lo anterior causa un alto nivel de estrés y ansiedad tanto a pacientes como médicos. Estas aplicaciones en medicina se pueden observar en la Figura 1.3.

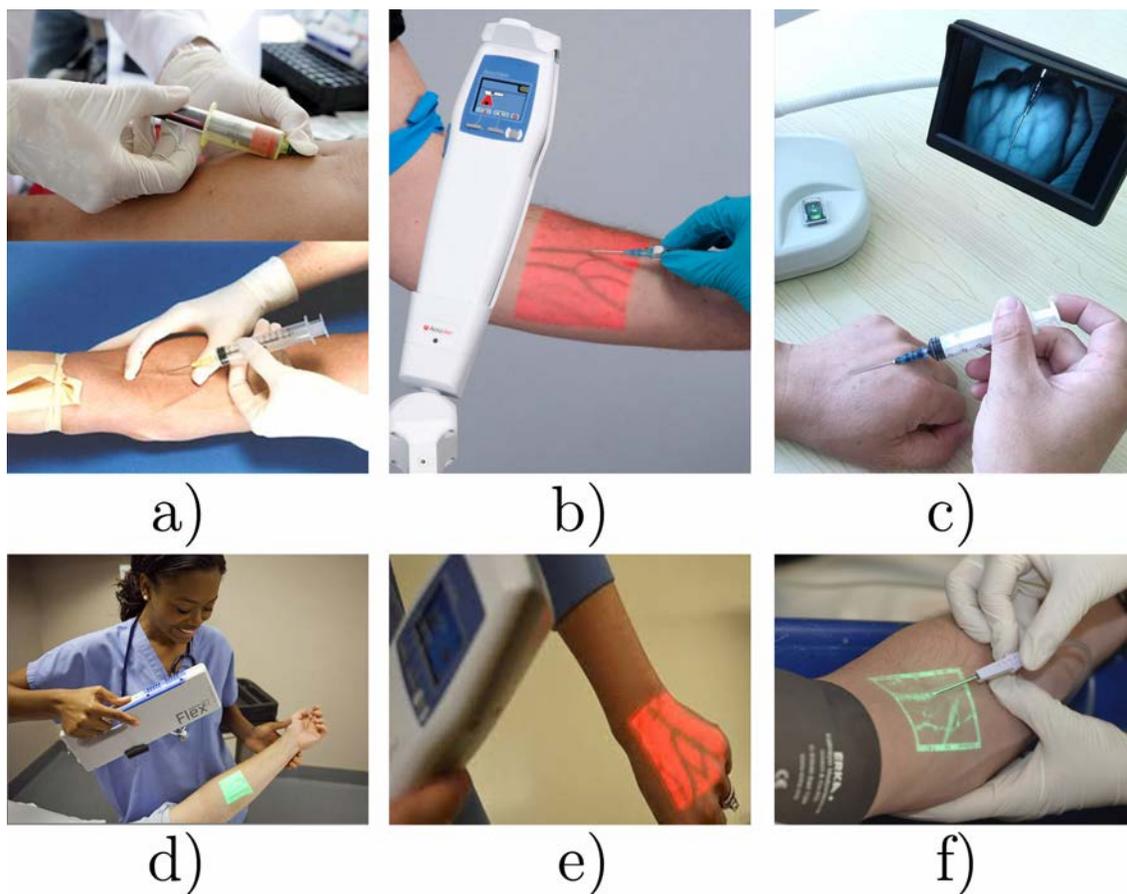


Figura 1.3: Venopunción. a) Método tradicional, b) AccuVein, c) Visor de venas portátil, d) VeinViewer Flex, e) Venopunción por AccuVein y f) Venopunción por VeinViewer Flex.

1.4. Objetivos

Objetivo general

Implementar un sistema de visión por computadora para la adquisición y análisis de imágenes en el infrarrojo cercano de la palma y dorso de la mano con propósitos de reconocimiento y autenticación de la identidad de una persona.

Objetivos principales

- Implementar un sistema de adquisición de imágenes IR de la palma y dorso de la mano. El sistema consta principalmente de una fuente de iluminación con emisión en los 880 *nm* y una cámara Multiespectral 2CCD Visible e Infrarrojo.
- Implementar algoritmos para el mejoramiento y segmentación de imágenes usando técnicas de procesamiento, basadas principalmente en a) Ecualización del histograma, b) Contraste Limitado de la Ecualización del Histograma y c) Máscara de desenfoque. Además de otras técnicas clásicas como filtraje espacial, binarización y morfología matemática.
- Implementar algoritmos para la extracción de características de las imágenes binarias del patrón de venas de una persona. Las técnicas implementadas están basadas en Momentos Geométricos, Momentos de Hu y Descriptor de Venas.
- Implementar algoritmos para la clasificación de personas a partir de los descriptores obtenidos del patrón de venas. Estos métodos están basados en K Vecinos Próximos (k-nn) y Redes Neuronales Artificiales (RNA).
- Implementar una interfaz en Matlab para la adquisición y procesamiento de imágenes IR de la mano de una persona.
- Análisis del rendimiento de los algoritmos de preprocesamiento, extracción de características y clasificación en bases de datos públicas. Las bases de datos analizadas son: a) PolyU Multispectral Palmprint Database, b) Vein Pattern Database y c) CASIA Multi-Spectral Palmprint Image Database. Todas ellas con características propias.

1.5. Antecedentes

La biometría es el estudio de los métodos automáticos para el reconocimiento único de humanos basados en uno o más rasgos conductuales o físicos intrínsecos. El término biometría proviene de las voces griegas "bios", que significa vida y "metron" que significa medida [14].

La biometría se puso en práctica en las culturas occidentales hasta finales del siglo XIX, pero era utilizada en China desde al menos el siglo XIV. Un explorador y escritor que respondía al nombre de Joao de Barros escribió que los comerciantes chinos estampaban las impresiones y las huellas de la palma de las manos de los niños en papel con tinta. Los comerciantes hacían esto como método para distinguir entre los niños jóvenes [15].

En Occidente, la identificación confiaba simplemente en la "memoria eidética" o "memoria fotográfica" definida como la capacidad de recordar cosas oídas y vistas con un nivel de detalle casi perfecto, hasta que Alphonse Bertillon (1853-1914), jefe del departamento fotográfico de la Policía de París, desarrolló el sistema antropométrico o como se le conocería más tarde como Bertillonage en 1883 [16].

El sistema consistía simplemente en dar una aplicación práctica a los trabajos antropológicos de su padre. Partiendo del principio del que no existen dos individuos absolutamente idénticos, pensó que tomando cierta cantidad de medidas susceptibles de no variar en el curso de una vida se conseguiría identificar a un individuo de un modo prácticamente infalible, cualesquiera que fuesen su edad y el medio que empleara para modificar su apariencia física.

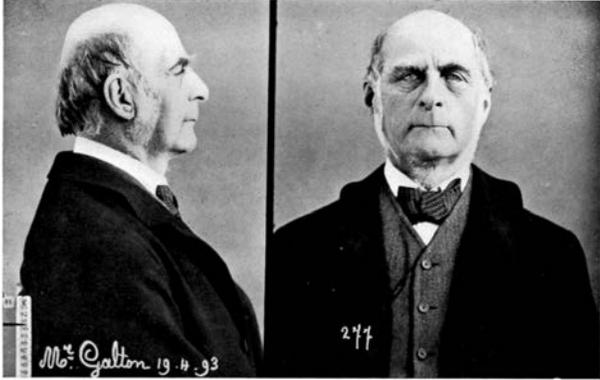
Bertillon creó por tanto una ficha de filiación como la de la Figura 1.4, en la cual hizo figurar cierta cantidad de medidas invariables para un individuo que ha concluido su desarrollo: la dimensión de la cabeza, la de la mano, etc. Añadió naturalmente un apartado para las señales particulares [17].

Éste era el primer sistema preciso, ampliamente utilizado científicamente para identificar a criminales y convirtió a la biometría en un campo de estudio. Funcionaba midiendo de forma precisa ciertas longitudes y anchuras de la cabeza y del cuerpo, así como registrando marcas individuales como tatuajes y cicatrices. El sistema de Bertillon fue adoptado extensamente en occidente hasta que aparecieron defectos en él. Después de esto, las fuerzas policiales occidentales comenzaron a usar la huella dactilar.

El área de biometría ha crecido en las últimas décadas la cual inicio desde usar simplemente la huella dactilar, a emplear muchos métodos distintos teniendo en cuenta varias medidas físicas y de comportamiento. Las aplicaciones de la biometría también han aumentado desde el reconocimiento de personas hasta los controles de acceso, la seguridad informática y la medicina.

Taille 1 ^m	Long ^m	Pied g.	N ^o de cl.	Agé de
Voite	Larg ^m	Médus g.	Aur ^m	né le
Enverg 1 ^m	Long ^m	Auric ^m g.	Part ^m	a
Buste 0,	Larg ^m	Coudée g.	Part ^m	dep ^m
				Agé app ^m

(Réduction photographique 1/7.)



Inclin ^m	Racine (cavité)	Bord. o. s. p. f.	Burbe	plig ^m
Haut ^m	Dis Base	Lob. c. a. m. d.	Cheveux	Cou ^m sang ^m
Larg ^m	Haut ^m Skallie Larg ^m	A. trg. i. p. r. d.	Car	Couit.
Part ^m	Part ^m	Pli. f. s. h. E.	Autres traits caractéristiques :	
		Part.	Sig ^m dressé par M.	

Figura 1.4: Ficha de filiación: Sir Francis Galton fotografiado por Bertillon en 1893 [17].

La idea para usar patrones de iris como método de identificación fue propuesta en 1936 por el oftalmólogo Frank Burch. Para los años 80's la idea ya había aparecido en películas de James Bond, pero permanecía siendo ciencia ficción. En 1985 los doctores Leonard Flom y Aran Safir retomaron la idea. Su investigación les concedió una patente en 1987 [18] ya que propusieron el concepto de que no hay dos iris semejantes.

En 1989 Flom y Safir recurrieron a John Daugman para crear algoritmos para el reconocimiento de iris. Estos algoritmos, patentados por Daugman en 1994 [19] y que son propiedad de Iridian Technologies, son la base para todos los productos de reconocimiento de iris. En 1993 la Agencia Nuclear de Defensa empezó a trabajar con IriScan, Inc. para desarrollar y probar un prototipo, 18 meses después el primer prototipo se completó y está disponible comercialmente.

Hoy en día, existen tecnologías que permiten automatizar el proceso de comparación de datos, acelerando el tiempo de verificación de la identidad. Para la identificación de personas en entornos físicos como el papel, se han ido incorporando tecnologías biométricas automatizadas que siguen estándares internacionales, facilitando el rastreo de datos.

Actualmente, la tecnología biométrica esta presente en multiples aplicaciones tanto públicas como privadas. Los principales estudios y aplicaciones de la biometría estática están basados en la medición de las huellas dactilares, el iris, la retina, la geometría de la mano, la

forma de la cara, la forma de la oreja, la quiromancia, las venas del dorso y palma de la mano, la temperatura corporal, la forma del cuerpo, entre otras [20].

La medición de las características del comportamiento de las personas es conocida como biometría dinámica. Estas están basadas en el análisis de los patrones de modulación de voz, la firma, el reconocimiento por la forma de caminar, el análisis gestual, la dinámica de teclado, entre otras [20]. Los ejemplos antes mencionados se pueden observar en la Figura 1.5.

En general, no importa el tipo de biometría que se use, este tipo de tecnología siempre permitirá un control eficiente y preciso de la identidad de las personas, prácticamente eliminando los riesgos de suplantación o robo de identidad. Y, como es de prever, sus usos y aplicaciones aumentarán progresivamente, pues existen posibilidades ilimitadas para este tipo de tecnología.

Por otro lado, la biometría también ha sido ampliamente usada en aplicaciones forenses como identificación de criminales y la seguridad en prisiones. En general, la tecnología biométrica está evolucionando rápidamente y tiene un fuerte potencial que puede ser ampliamente usado en aplicaciones civiles como por ejemplo, la banca y el comercio electrónico, los controles de acceso, seguridad informática.

Lo anterior debido al rápido crecimiento del número y uso de medios y dispositivos electrónicos, estos se están convirtiendo en una de las más importantes aplicaciones de la biometría. Estas aplicaciones incluyen tarjetas de crédito, seguridad en cajeros automáticos, transferencia de efectivo, transacciones en línea y acceso a la web.

Para esto, existen diversos dispositivos que son utilizados para almacenar los datos biométricos. Dichos dispositivos, se apoyan en distintos factores, de acuerdo al nivel de seguridad que requiera la aplicación. Comúnmente se considera como factor de autenticación, algo que se sabe (clave o PIN), que se tiene (llaves y tarjetas electrónicas) o que se es (biometría). La utilización de estos factores de autenticación da seguridad al proceso.

Las aplicaciones de control de acceso físico, acceso remoto y de acceso a datos han usado tradicionalmente la autenticación basada en claves, estas aplicaciones han comenzado a usar biometría para la autenticación de personas. El uso de la biometría estará aún más extendida en los siguientes años a medida que la tecnología avanza y esta se va haciendo más confiable [21]. Otras aplicaciones biométricas incluyen:

- Control de migración.
- Identificación nacional.
- Votación.
- Control de asistencia.

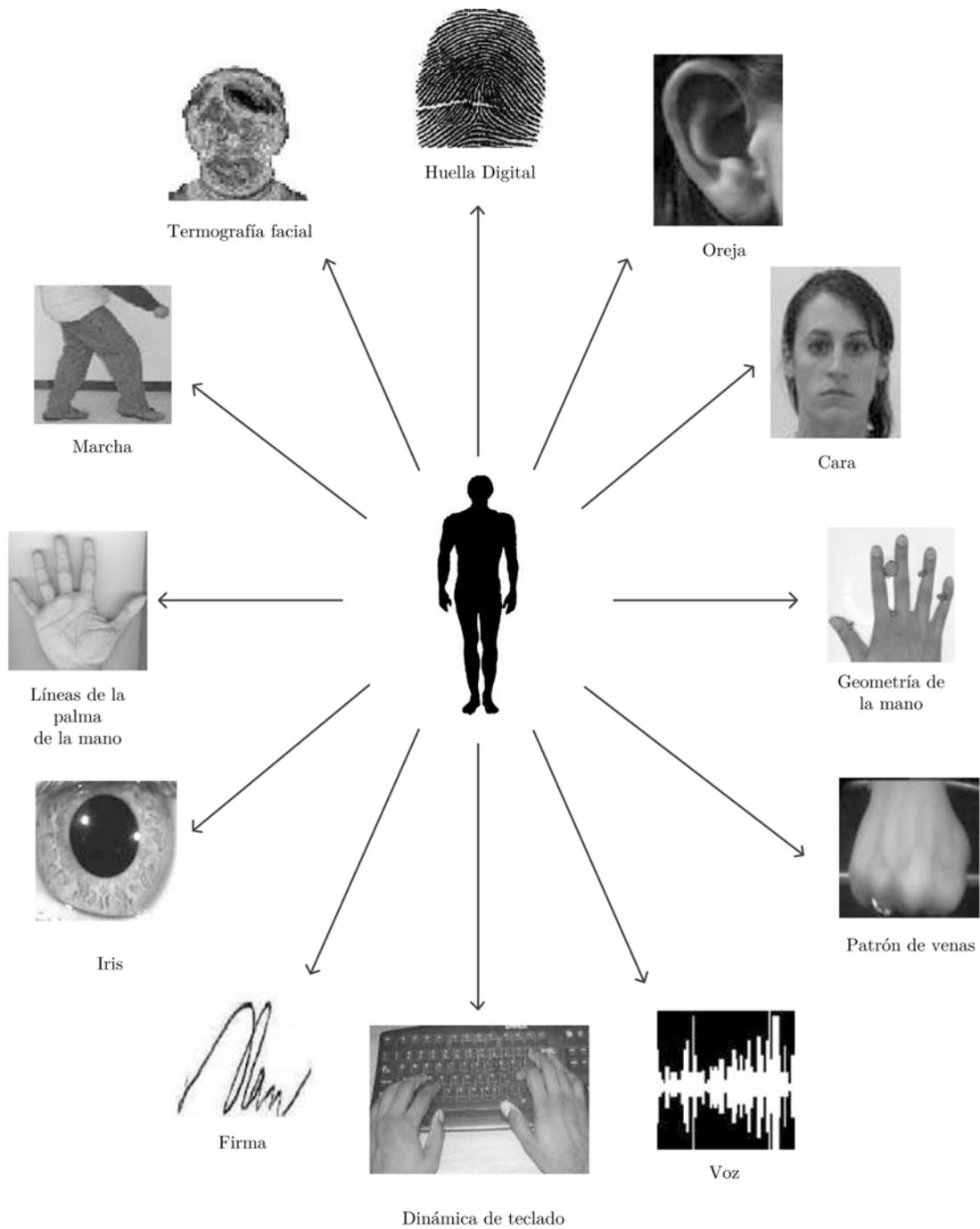


Figura 1.5: Ejemplos de rasgos biométricos que pueden ser usados para identificación. Los rasgos físicos (estáticos) incluyen las huellas dactilares, iris, geometría de cara, mano u oreja. Los rasgos de comportamiento (dinámicos), incluyen la firma, dinámica de teclado y marcha [8].

1.6. Estado del arte

El reconocimiento de personas a partir de imágenes adquiridas en el infrarrojo cercano de los patrones de venas de una persona, ha sido objeto de estudio por muchos investigadores.

Algunos como, Zhi Liu *et al.* [2], Jian-Da Wu *et al.* [3], Jinfeng Yang *et al.* [7] y [22], Eui Chul Lee *et al.* [23] y Byung Jun Kang *et al.* [24] y [25], utilizan imágenes de la distribución del patrón de venas de los dedos (Figura 1.6) de la mano como medio de identificación.

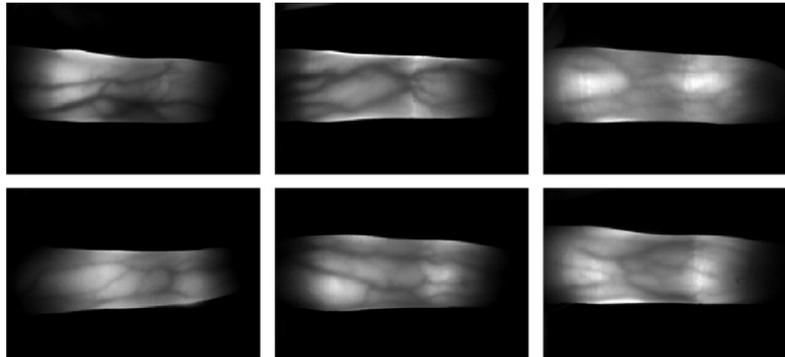


Figura 1.6: Ejemplo de imágenes del patrón de venas de los dedos de la mano [23].

Algunos otros, utilizan imágenes de la palma de la mano (Figura 1.7) en el infrarrojo para propósitos de reconocimiento. Ejemplo de ello, son los trabajos desarrollados por Jen-Chun Lee [1], Masaki Watanabe *et al.* [5], Suleyman Malki *et al.* [26], David Zhang *et al.* [27], Ishani Sarkar *et al.* [28], Jixing Wang *et al.* [29], Wei-Yu Han *et al.* [30], G. S. Badrinath *et al.* [31] y Jian-Gang Wang *et al.* [32].

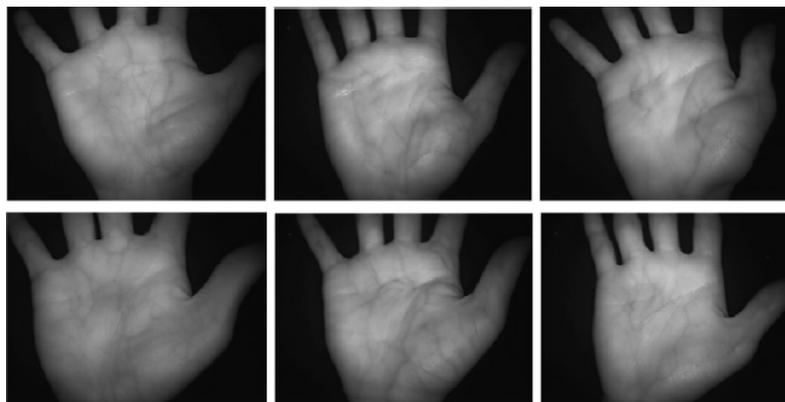


Figura 1.7: Ejemplo de imágenes del patrón de venas de la palma de la mano [30].

Existen trabajos que, de igual manera, utilizan imágenes del patrón vascular del dorso de la mano (Figura 1.8). Como ejemplos de estos trabajos se pueden mencionar los realizados por Oscar Fernando *et al.* [1], Deepak Prasanna. R *et al.* [5], Aída Marcotti, *et al.* [6], Zahra Honarpisheh *et al.* [33], Ahmed M. Badawi [34], Lingyu Wang *et al.* [35] y Wenxiong Kang [36].

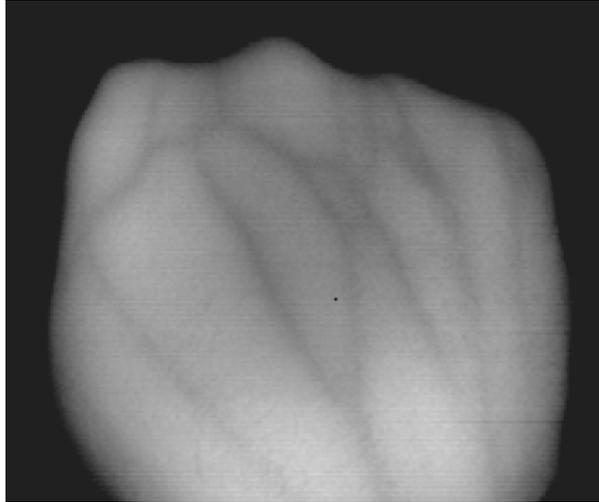


Figura 1.8: Ejemplo de imagen del patrón de venas del dorso de la mano [33].

Otros como Qiushi Zhao *et al.* [4], Maleika Heenaye *et al.* [7] y Goh Kah Ong Michael *et al.* [37] y [38], utilizan una combinación de imágenes tanto de la palma, dorso o de los dedos tanto en el visible como en el infrarrojo cercano.

De los trabajos anteriores, con el objetivo de reducir los tiempos de cómputo, se determinan zonas o regiones de interés. Estas regiones de interés disminuyen la carga de trabajo a la computadora, más cuando se trata de analizar y procesar grandes cantidades de imágenes. Muestra de ello son los trabajos realizados por [1-5], [9], [6], [7] y [22], [27], [29-32] y [34-38].

El filtraje espacial es utilizado para remover información innecesaria de los pliegues y vellos de la palma y dorso de la mano, respectivamente. Este tipo de información innecesaria, genera ruido a las imágenes, por lo que es necesario eliminarlo como lo hacen en [5], [9], [6], [29], [34] y [35].

Debido a que las imágenes adquiridas cuentan con un bajo contraste, algunos autores como [5] y [33], utilizan algoritmos para el mejoramiento y realce del contraste de las imágenes del patrón de venas, basados en la información que se encuentra contenida en el histograma de la imagen.

Algunos de estos algoritmos utilizan la Ecualización del histograma con el objetivo de redistribuir los valores de intensidad de gris y con ello, realzar el contraste de la imagen adquirida. Otros como [3], utilizan un valor de Gamma para el realce de la imagen.

En la etapa de segmentación y binarización, las técnicas más comunes son por el método de Otsu como en [1], [7], [30] y [32]. Utilizando el método de Niblack en [9] y [36] realizan la segmentación de las imágenes y de los patrones de venas. Empleando la técnica de binarización adaptativa que combina tanto umbralización local como global, se muestran en [1], [5], [24] y [35]. La técnica de binarización basada en un umbral es utilizada en [4], [6], [23] y [26-29] y [31], [33] y [34].

Algunas de las técnicas para la extracción de características se exponen a continuación. Óscar Fernando *et al.* [1] utiliza un submuestreo del esqueleto del patrón vascular para extraer la información cualitativa de la distribución de venas.

Jen-Chun Lee [1], Jinfeng Yang *et al.* [22] y Jixing Wang *et al.* [29], utilizan un filtro de Gabor 2-D, el cual es usado para la extracción de características en imágenes con texturas, en otros casos se utiliza un solo filtro de Gabor como lo hace Wei-Yu Han *et al.* [30]. Otros como Jian-Da Wu *et al.* [3], emplean la transformada de Radon como método para la extracción de características de las imágenes de las venas de los dedos.

Qiushi Zhao *et al.* [4], utilizan el algoritmo denominado *Competitive Code* el cual representa la información de la orientación de los píxeles en imágenes de los pliegues de la palma. Por otra parte, para imágenes de las venas, se emplea el algoritmo *Orientation Of Local Binary Patterns*, el cual es usado para extraer las características de la orientación de las venas.

Las técnicas clásicas para la extracción de características de mayor efectividad, están basadas en métodos denominados como puntos de referencia utilizados por Yun-peng Hu *et al.* [9], Suleyman Malki *et al.* [26]. Sin embargo, para que éstas técnicas funcionen, es necesario que la etapa de preprocesamiento sea la mejor posible. Byung Jun Kang *et al.* [24], Zahra Honarpisheh *et al.* [33] y Lingyu Wang *et al.* [35], utilizan las intersecciones o bifurcaciones, así como los puntos finales de las venas como características de las imágenes analizadas.

Goh Kah Ong Michael *et al.* [37], proponen un nuevo esquema llamado *Directional Coding* para la extracción de características de los pliegues y venas de la palma de la mano.

En la Tabla 5.26, se muestra un resumen con los porcentajes de clasificación así como la cantidad de personas e imágenes que fueron utilizadas. Además se muestra la región de la mano utilizada como por ejemplo el dorso, los dedos o la palma.

Referencia	Número de Imágenes	Número de Personas	Número de Versiones	Porcentaje de Clasificación	Lugar
[1]	100	20	5	95.72	Dorso
[1]	4140	207	***	99.18	Palma
[3]	500	25	10	99.2	Dedos
[4]	10800	90	***	***	Palma y Dorso
[7]	6000	500	12	***	Palma y Dorso
[9]	100	20	5	98	Dorso
[22]	640	64	10	97.187	Dedos
[23]	8160	102	***	62.18	Dedos
[25]	6380	80	***	98.56	Dedos
[27]	6000	500	12	98.8	Palma
[29]	800	80	***	98.8	Palma
[34]	5000	500	10	99.88	Dorso
[35]	141	47	***	***	Dorso
[36]	360	60	6	93.54	Dorso

Tabla 1.1: Comparativa de porcentajes de clasificación de diversos trabajos relacionados al reconocimiento de personas a través de la distribución del patrón de venas.

1.7. Aportaciones

En este proyecto de tesis se tienen las siguientes aportaciones:

1. Se desarrolló una interfaz en MATLAB para la adquisición y procesamiento de imágenes IR de 3 regiones de la mano (palma, dorso y antebrazo). Los algoritmos programados para el procesamiento de imágenes, Delimitan la Región de Interés, Mejoran el contraste y Segmentan el patrón de venas.
2. Se proponen dos procedimientos para la extracción de la distribución del patrón de venas de la mano a partir de imágenes IR. Los procesos de mejoramiento de imágenes están basados en el Contraste Limitado de la Ecuilización Adaptativa del Histograma y la Máscara de Desenfoque.
3. Se propuso una técnica para la generación de descriptores del patrón de venas, la cual se denominó: Descriptor de Venas.

Bibliografía

- [1] Óscar Fernando Motato Toro y Humberto Loaiza Correa, "Identificación biométrica utilizando imágenes infrarrojas de la red vascular de la cara dorsal de la mano," *Revista Ingeniería e Investigación*, vol. 29, pp. 90-100, 2009.
- [2] Jen-Chun Lee, "A novel biometric system based on palm vein image," *Pattern Recognition Letters*, vol. 33, pp. 1520-1528, 2012.
- [3] Zhi Liu, Yilong Yin, Hongjun Wang, Shangling Song and Qingli Li, "Finger vein recognition with manifold learning," *Journal of Network and Computer Applications*, vol. 33, pp. 275-282, 2010.
- [4] Jian-Da Wu and Siou-Huan Ye, "Driver identification using finger-vein patterns with Radon transform and neural network," *Expert Systems with Applications*, vol. 36, pp. 5793-5799, 2009.
- [5] Qiushi Zhao, Wei Bu, Xiangqian Wu, and David Zhang, "Design and implementation of a contactless multiple hand feature acquisition System," *Proc. of SPIE* vol. 8371, 2012.
- [6] Deepak Prasanna. R, Neelamegam. P, Sriram. S and Nagarajan Raju, "Enhancement of vein patterns in hand image for biometric and biomedical application using various image enhancement techniques," *Procedia Engineering*, vol. 38, pp. 1174-1185, 2012.
- [7] A. K. Jain, R. Bolle, and S. Pankanti, editors. "Biometrics: Personal Identification in Networked Society," *Kluwer Academic Publishers*, 1999.
- [8] Maleika Heenaye and Mamode Khan, "A multimodal hand vein biometric based on Score Level Fusion," *Procedia Engineering*, vol. 41, pp. 897-903, 2012.
- [9] Anil K. Jain, Patrick Flynn y Arun A. Ross, editors, "Handbook of Biometrics," *Springer Science+Business Media*, 2008.
- [10] Yun-peng Hu, Zhi-yong Wang, Xiao-ping Yang and Yu-ming Xue, "Hand vein recognition based on the connection lines of reference point and feature point," *Infrared Physics & Technology*, vol. 62, pp. 110-114, 2014.

-
- [11] Masaki Watanabe, Toshio Endoh, Morito Shiohara and Shigeru Sasaki, "Palm vein authentication technology and its applications," *Proceedings of The Biometric Consortium Conference*, 2005.
- [12] Aída Marcotti, Belén Hidalgo and Ladislao Mathé, "Método de detección no invasivo utilizando luz infrarroja," *Universidad Nacional de Córdoba, Argentina, Master Degree Thesis*, 2012.
- [13] Jinfeng Yang and Yihua Shi, "Finger-vein ROI localization and vein ridge enhancement," *Pattern Recognition Letters*, vol. 33, pp. 1569-1579, 2012
- [14] Ricardo Palomares Borja, "Marketing en el punto de venta, 100 ideas clave para vender más," *ESIC Editorial*, ISBN: 9788473568869, 2da. Edición, pags. 509, 2012.
- [15] Ramón Blanco Gonzalo, "Análisis de seguridad en comunicaciones USB," *Universidad Carlos III de Madrid, Escuela Politécnica Superior, Departamento de Tecnología Electrónica*, Proyecto de Fin de Carrera, Noviembre 2009.
- [16] Anna Maria Guasch, "Arte y archivo, 1920-2010: Genealogías, tipologías y discontinuidades," *Ediciones AKAL*, ISBN: 9788446025399, 1ra. Edición, pags. 320, 2011.
- [17] Alphonse Bertillon, "Identification Anthropometrique, Instructions Signalétiques," *Melun, Imprimerie Administrative*, 1893.
- [18] Leonard Flom y Aran Safir, "Iris recognition system," <http://www.google.com/patents/US4641349>, Patente, 3 de Febrero de 1987.
- [19] John G. Daugman, "Biometric personl identification system based on iris analysis," <https://www.google.com/patents/US5291560>, Patente, 1 de Marzo de 1994.
- [20] Luzmila Pró, Juan Carlos Gonzáles, Walter Contreras y Carlos Yañez, "Tecnologías biométricas aplicadas a la seguridad en las organizaciones," *Revista de Ingeniería de Sistemas e Informática*, vol. 6, num. 2, Julio - Diciembre 2009.
- [21] Avril Hernández Briones, "Biometría Informática," <http://redyseguridad.fi-p.unam.mx/proyectos/biometria/index.html>, UNAM - Facultad de Ingeniería.
- [22] Jinfeng Yang and Xu Zhang, "Feature-level fusion of fingerprint and finger-vein for personal identification," *Pattern Recognition Letters*, vol. 33, pp. 623-628, 2012.
- [23] Eui Chul Lee and Kang Ryoung Park, "Image restoration of skin scattering and optical blurring for finger vein recognition," *Optics and Lasers in Engineering*, vol. 49, pp. 816-828, 2011.

-
- [24] Byung Jun Kang and Kang Ryoung Park, "Multimodal biometric authentication based on the fusion of finger vein and finger geometry," *Optical Engineering SPIE*, vol. 48, Septiembre 2009.
- [25] Byung Jun Kang and Kang Ryoung Park, Jang-Hee Yoo, Jeong Nyeo Kim, "Multimodal biometric method that combines veins, prints, and shape of a finger," *Optical Engineering SPIE*, vol. 50, Enero 2011.
- [26] Suleyman Malki and Lambert Spaanenbunrg, "Hand veins feature extraction using DT-CNNs," *VLSI Circuits and Systems III*, vol. 6590, 2007.
- [27] David Zhang, Zhenhua Guo, Guangming Lu, Lei Zhang, Yahui Liu, Wangmeng Zuo, "Online joint palmprint and palmvein verification," *Expert Systems with Applications*, vol. 38, pp. 2621-2631, 2011.
- [28] Ishani Sarkar, Farkhod Alisherov, Tai-hoon Kim and Debnath Bhattacharyya, "Palm vein authentication system: a review," *International Journal of Control and Automation*, vol. 3, Marzo 2010.
- [29] Jixing Wang, Yuqing He, Jiadan Zhu, Xinru Gao and Yongsheng Cui, "Palm vein for efficient person recognition based on 2D Gabor Filter," *Biometric and Surveillance Technology for Human and Activity Identification X*, vol. 8712, 2013.
- [30] Wei-Yu Han and Jen-Chun Lee, "Palm vein recognition using adaptive Gabor filter," *Expert System s with Applications*, vol. 39, pp. 13225-13234, 2012.
- [31] G.S. Badrinath and Phalguni Gupta, "Palmprint based recognition system using phase-difference information," *Future Generation computer Systems*, vol. 28, pp. 287-305, 2012.
- [32] Jian-Gang Wang, Wei-Yun Yau, Andy Suwandy and Eric Sung, "Person recognition by fusing palmprint and palm vein images based on "Laplacianpalm" representation," *Pattern Recognition*, vol. 41, pp. 1514-1527, 2008.
- [33] Zahra Honarpisheh and Karim Faez, "An efficient dorsla hand vein recognition based on Firefly Algorithm," *International Journal of Electrical and Computer Engineering*, vol. 3, pp. 30-41, Febrero 2013.
- [34] Ahmed M. Badawi, "Hand vein biometric verification Prototype: A testing performance and patterns similarity," *International Journal of Biomedical Sciences*, pp. 141-148, 2007.
- [35] Lingyu Wang, Graham Leedham and David Siu-Yeung Cho, "Minutiae feature analysis for infrared hand vein pattern biometrics," *Pattern Recognition*, vol. 41, pp. 920-929, 2008.

- [36] Wenxiong Kang, "Vein pattern extraction based on vectorgrams of maximal intra-neighbor difference," *Pattern Recognition Letters*, vol. 33, pp. 1916-1923, 2012.
- [37] Goh Kah Ong Michael, Tee Connie and Andrew Beng Jin Teoh, Girija Chetty and Jucheng Yang (Editors), "A contactless biometric system using palm print and palm vein features," *Advanced Biometric Technologies*, ISBN: 978-953-307-487-0, pp. 394, 2011.
- [38] Goh Kah Ong Michael, Tee Connie and Andrew Beng Jin Teoh, "A contactless biometric system using multiple hand features," *Journal of visual communication & image representation*, vol. 23, pp. 1068-1084, 2012.

Capítulo 2

Procesamiento Digital de Imágenes

2.1. Introducción

El procesamiento digital de imágenes, incluye un conjunto de técnicas que operan sobre la representación digital de una imagen, con el objetivo de destacar algunos de los elementos que conforman la escena, de modo que se facilite su análisis posterior. En general, las técnicas de procesamiento de imágenes son aplicadas cuando se requiere medir, contrastar o clasificar algún elemento presente en dicha imagen [1].

Existen múltiples formas de adquirir las imágenes, pero el objetivo es el mismo: generar imágenes digitales a partir de datos censados o detectados, ya que la salida en la mayoría de los sensores es una forma de onda de voltaje continua, la amplitud y el comportamiento espacial están relacionados con el fenómeno físico que se está censando o detectando [1].

Para crear una imagen digital, se requieren esencialmente dos elementos de hardware. El primero es un dispositivo sensible a algún tipo de radiación electromagnética (luz visible, luz infrarroja, rayos X, luz ultravioleta, etc.) capaz de traducir esa radiación en una señal eléctrica proporcional al nivel de energía recibido. El segundo es un dispositivo digitalizador, el cual convertirá los datos continuos detectados en forma digital. Esto implica dos procesos: el muestreo y la cuantización [1], [2].

Un sistema de adquisición de imágenes está compuesto por una cámara CCD, la cual genera una señal de video a partir de un sensor semiconductor matricial. Esta es luego cuantificada mediante un conversor analógico/digital (A/D) adaptado a este tipo de señal. Adicionalmente de manera opcional se puede utilizar el pre-procesamiento por hardware que generalmente se trata de dispositivos de filtraje y/o restauración de la señal eléctrica para evitar la alteración de la información contenida en ella [2]. En la Figura 2.1 se muestra un sistema de adquisición de imágenes.

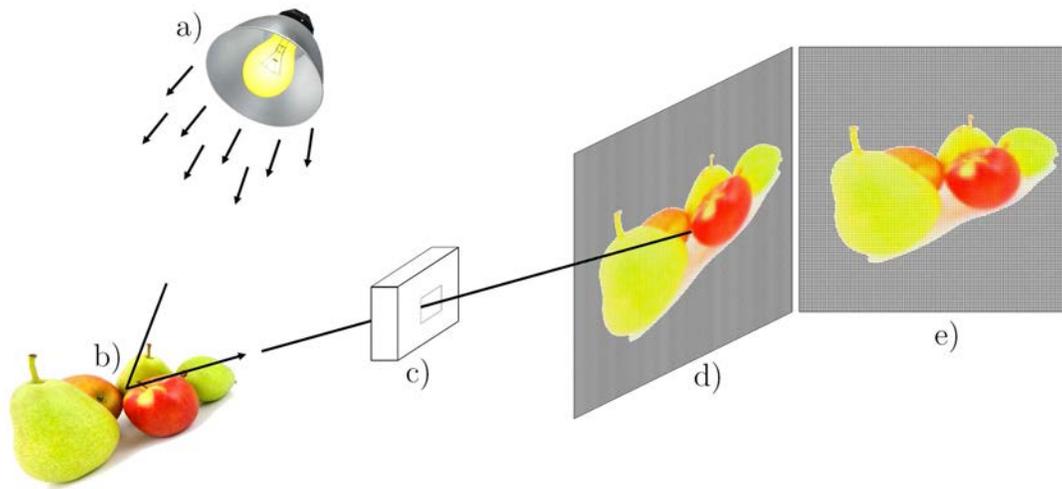


Figura 2.1: Un ejemplo del proceso de adquisición de una imagen digital. (a) Fuente de iluminación (energía), (b) Un elemento de una escena, (c) Sistema de adquisición de imágenes, (d) Proyección de la escena en el plano imagen y (e) Imagen digitalizada.

Por lo tanto, una imagen puede ser definida como una función bidimensional $f(x, y)$, donde x y y son coordenadas espaciales, y la amplitud de f en cualquier par de coordenadas (x, y) son llamadas niveles de intensidad o gris de una imagen en ese punto [2].

A partir de éstas imágenes digitales, es posible obtener información para:

- La interpretación visual, en la cual se utilizan técnicas o métodos de realce de imágenes.
- La interpretación automática, en la que se extrae la información contenida en las imágenes utilizando técnicas como sistemas basados en momentos, Redes Neuronales Artificiales (RNA), análisis de forma, tamaño y posición.

En el presente capítulo, se describen las técnicas de mejoramiento de imágenes basadas en transformaciones en niveles de gris. Así mismo, se describen los métodos que utilizan el histograma para realzar el contraste de las imágenes como la Ecuación del Histograma, y sus variantes, la Ecuación Adaptativa del Histograma (AHE) y el Contraste Limitado por AHE (CLAHE). Se presentan además las técnicas que usan máscaras o kernels de transformación basadas en convolución como filtros promedio o la máscara de desenfoco.

De igual forma se detallan algunos métodos para la segmentación de imágenes como los basados en la información obtenida a partir del histograma de una imagen con el fin de separar el objeto o elemento a analizar del fondo y en la delimitación de una área de interés. A

este proceso se le conoce como binarización de la cual se hablará más adelante, así mismo se detallará el método de binarización adaptativa el cual es una variante de la binarización en la que la iluminación uniforme en una imagen no depende para la obtención de buenos resultados.

Posteriormente se muestran los fundamentos de la morfología matemática así como las dos operaciones (Dilatación y Adelgazamiento) utilizadas en este trabajo de Tesis. Finalmente se describen las conclusiones.

2.2. Mejoramiento de imágenes

El objetivo principal del realce o mejoramiento, es procesar una imagen de manera que el resultado sea el más adecuado que la imagen original para una aplicación específica. La palabra específico es importante, ya que establece en primer lugar, que las técnicas están orientadas a muchos problemas. Así, por ejemplo, un método que resulte ser muy útil para mejorar las imágenes de rayos X puede no ser necesariamente la mejor para realzar características de fotografías de Marte transmitidas por una sonda espacial [1].

Independientemente del método utilizado, el mejoramiento de imágenes es una de las áreas más interesantes y visualmente atractivas para el procesamiento de imágenes.

El mejoramiento de imágenes se divide en dos grandes categorías:

- Métodos en el dominio espacial.
- Métodos en el dominio frecuencial.

El término dominio espacial se refiere al propio plano de la imagen, y esta categoría está enfocada en la manipulación directa de los píxeles de una imagen. Las técnicas de procesamiento de dominio frecuencial se basan en la manipulación de los píxeles en el dominio de la transformada de Fourier de una imagen [1]. En este trabajo únicamente se utilizarán los métodos de mejoramiento de imágenes en el dominio espacial.

Los procesos en el dominio espacial serán denotados por la expresión,

$$g(x, y) = T[f(x, y)] \quad (2.1)$$

donde $f(x, y)$ es la imagen de entrada, $g(x, y)$ la imagen procesada, y T es un operador sobre f , definido sobre alguna vecindad de (x, y) . Adicionalmente, T puede operar en un conjunto de imágenes de entrada, tal como realizar la suma pixel por pixel de K imágenes para reducción de ruido.

La principal aproximación en la definición de una vecindad de un punto (x, y) es el usar un área cuadrada o rectangular como subimagen centrada en (x, y) como se observa en la Figura 2.2. El centro de la subimagen es trasladado al pixel de inicio, por ejemplo, en la esquina superior izquierda [1].

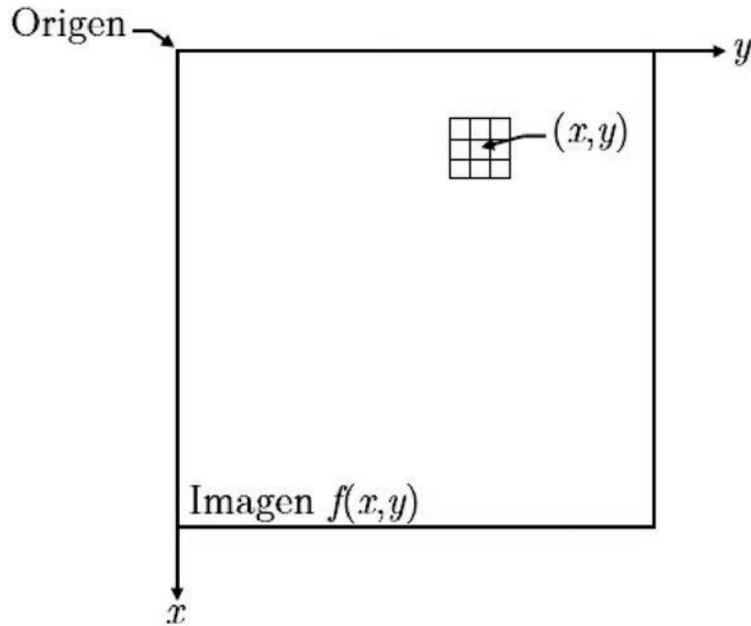


Figura 2.2: Una vecindad de 3×3 sobre un punto en una imagen.

El operador T es aplicado en cada (x, y) para producir la salida, g , en esa ubicación. El proceso utiliza sólo los píxeles en el área de la imagen expandida por la vecindad. A veces, son usadas otras formas de vecindades tales como aproximaciones al círculo, las vecindades cuadradas y rectangulares son los más predominantes debido a su facilidad de implementación.

La forma más simple de T es cuando el tamaño de la vecindad es de 1×1 . En este caso, g , depende solamente del valor de f en (x, y) , y T se convierte en una función de transformación de nivel de gris (también llamado una intensidad) de la forma,

$$s = T(r) \quad (2.2)$$

donde, para simplificar la notación, r y s son variables denotando, respectivamente, el nivel de gris de $f(x, y)$ y $g(x, y)$ en cualquier punto (x, y) .

El tamaño de la vecindad depende de las estructuras o elementos a analizar en una imagen. La aproximación general es utilizar una función de los valores de f en una vecindad prede-

finida de (x, y) para determinar el valor de g en (x, y) . Una principal aproximación en esta formulación esta basada en el uso de las llamadas máscaras (también conocido como filtros, kernels, plantillas o ventanas).

Básicamente, una máscara es un pequeño arreglo 2-D tal como se muestra en la Figura 2.2, las técnicas de realce basadas en este tipo de aproximaciones son conocidos como procesamiento por máscara o filtraje [1]. Existen tres tipos básicos de funciones que se utilizan con frecuencia para el mejoramiento de imágenes:

- Lineal (transformaciones negativo e identidad)
- Logarítmica (transformaciones de logaritmo e inversa del logaritmo)
- Potencias (transformaciones de enésima potencia y enésima raíz)

La función identidad es el caso trivial en el que los valores de intensidades de salida son iguales a las intensidades de entrada. En la Figura 2.3 se muestra un ejemplo de las transformaciones básicas en niveles de gris sobre una función imagen [1].

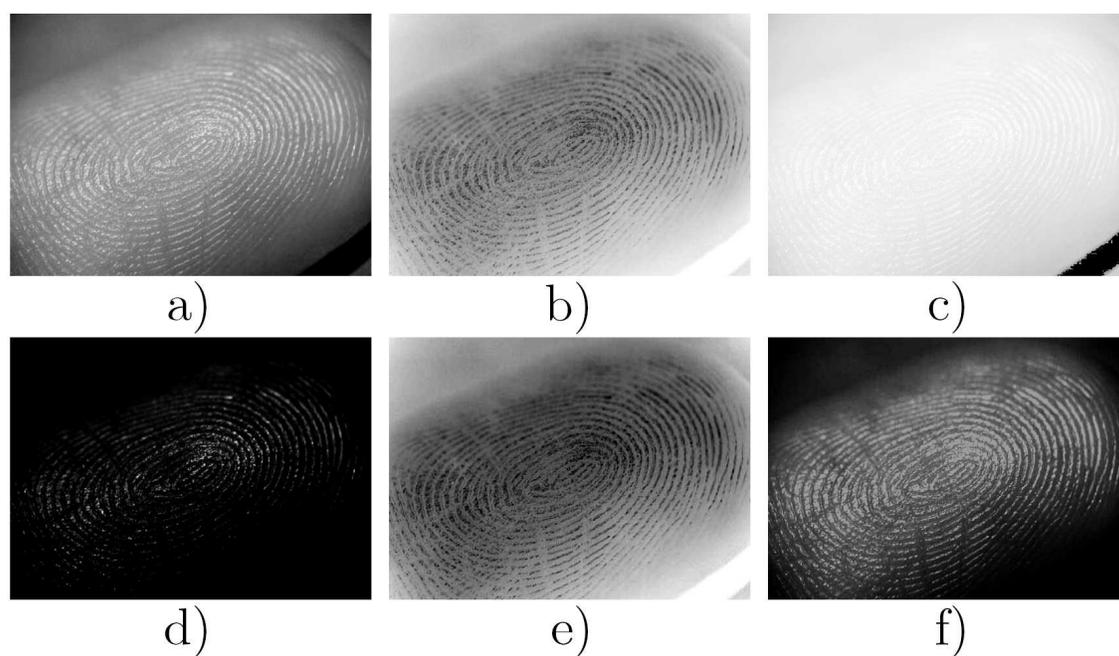


Figura 2.3: Ejemplo de transformaciones básicas en niveles de gris en el que se observa un incremento del contraste de la imagen. (a) Imagen Original, (b) Negativo, (c) Logaritmo, (d) Inversa de logaritmo, (e) Potencia y (f) Raíz.

2.3. Histograma de una imagen

El histograma de una imagen digital es ampliamente utilizado como herramienta tanto cualitativa como cuantitativa. Este corresponde a un gráfico de la distribución de valores de intensidad de los píxeles de una imagen (niveles de gris) o de una porción de la misma [1]. Este gráfico puede proporcionar importante información acerca del brillo y contraste de una imagen así como de su rango dinámico.

El histograma de una imagen con niveles de gris en el rango $[0, L - 1]$ es una función discreta de $h(r_k) = n_k$, donde r_k es el k -ésimo nivel de gris y n_k es el número de píxeles en la imagen que tienen un nivel de gris r_k . Es una práctica común normalizar un histograma dividiendo cada uno de sus valores por el número total de píxeles en la imagen, que estará denotado por $M \times N$. Así, un histograma normalizado está dado por $P(r_k) = \frac{n_k}{M \times N}$, para $k = 0, 1, \dots, L - 1$. Por lo tanto, $p(r_k)$ da una estimación de la probabilidad de ocurrencia de un nivel de gris r_k [1].

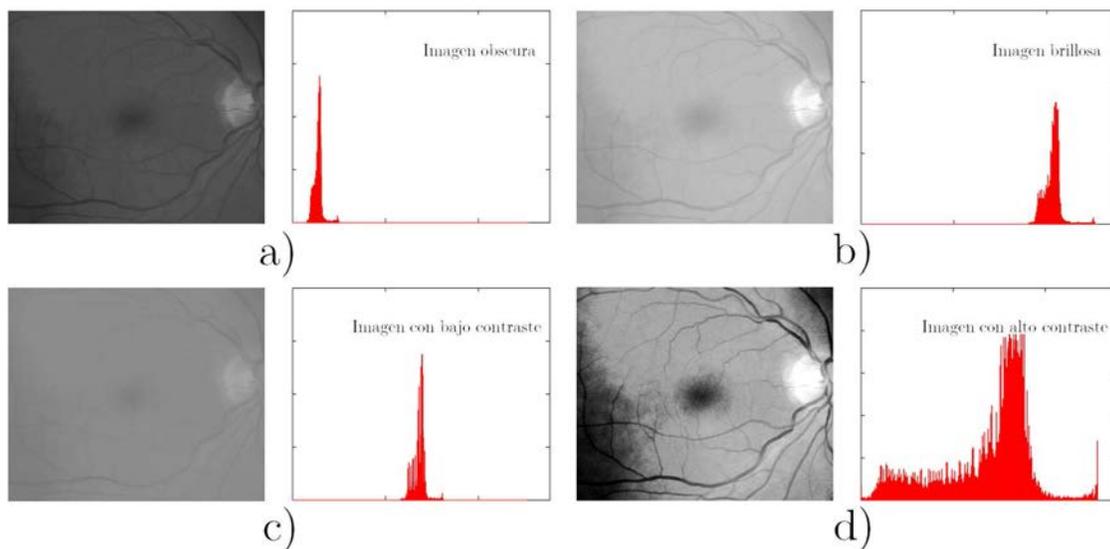


Figura 2.4: Cuatro tipos básicos de imágenes: (a) oscura, (b) brillante, (c) bajo contraste y (d) alto contraste con sus histogramas correspondientes.

Considérese la Figura 2.4 la cual es una imagen que tiene características básicas de niveles de gris: oscura, brillante, bajo contraste y alto contraste. El lado derecho de la Figura muestra las gráficas de los histogramas correspondientes a éstas imágenes. El eje horizontal de cada gráfica de histograma corresponde a los niveles de gris, r_k . El eje vertical corresponde a los valores de $h(r_k) = n_k$ o $p(r_k) = \frac{n_k}{M \times N}$ si los valores están normalizados. Así, esta gráfica de histogramas son simplemente gráficas de $h(r_k) = n_k$ contra r_k o $p(r_k) = \frac{n_k}{M \times N}$ contra r_k .

Como se puede observar en la Figura 2.4, las componentes del histograma están concentrados en valores cercanos a cero (negro) de la escala de grises. De manera similar, las componentes del histograma de la imagen brillante están parcialmente desplazadas hacia valores cercanos a L de la escala de grises.

La imagen con bajo contraste tiene un histograma que se estrecha y estará centrado a la mitad de la escala de grises. Finalmente, se puede observar que las componentes del histograma en la imagen altamente contrastada cubren un amplio rango de escalas de gris y, adicionalmente, esa distribución de píxeles no está demasiado uniforme debido a que tiene muy pocas líneas verticales que son muy superiores a las demás.

Es razonable concluir que una imagen cuyos píxeles tienden a estar distribuidos uniformemente, tendrá un aspecto de alto contraste y exhibirá una gran variedad de tonos de gris. El efecto neto será una imagen que muestra una gran cantidad de detalle de niveles de gris y tendrá un rango dinámico alto [1].

2.3.1. Ecuación del Histograma

La ecuación del histograma es una técnica de transformación no lineal que opera sobre los píxeles de una imagen de entrada y busca producir una imagen de salida con un histograma que se aproxime a un histograma especificado, usualmente constante.

Este, es un procedimiento que redistribuye los niveles de gris del histograma de una imagen, utilizando de la mejor manera posible, el rango disponible de niveles de gris en la imagen de salida. Un caso particular, el cual es utilizado con frecuencia, corresponde a considerar que el histograma de la imagen de salida es constante. Generalmente la transformación T es una función monótona creciente [1].

Como se mencionó anteriormente, la probabilidad de ocurrencia de un nivel de gris r_k en una imagen es,

$$p_r(r_k) = \frac{n_k}{M \times N} \quad k = 0, 1, 2, \dots, L - 1 \quad (2.3)$$

La versión discreta de la función de transformación es,

$$\begin{aligned} s_k = T(r_k) &= \sum_{j=0}^k p_r(r_j) \\ &= \sum_{j=0}^k \frac{n_j}{M \times N} \quad k = 0, 1, 2, \dots, L - 1 \end{aligned} \quad (2.4)$$

Por lo tanto, una imagen procesada (de salida) se obtiene asignando cada pixel con nivel r_k en la imagen de entrada en un pixel correspondiente con nivel s_k en la imagen de salida a través de la Ecuación 2.4. Una gráfica de $p_r(r_k)$ contra r_k se conoce como histograma. La transformación dada en la Ecuación 2.4 se conoce como ecualización del histograma.

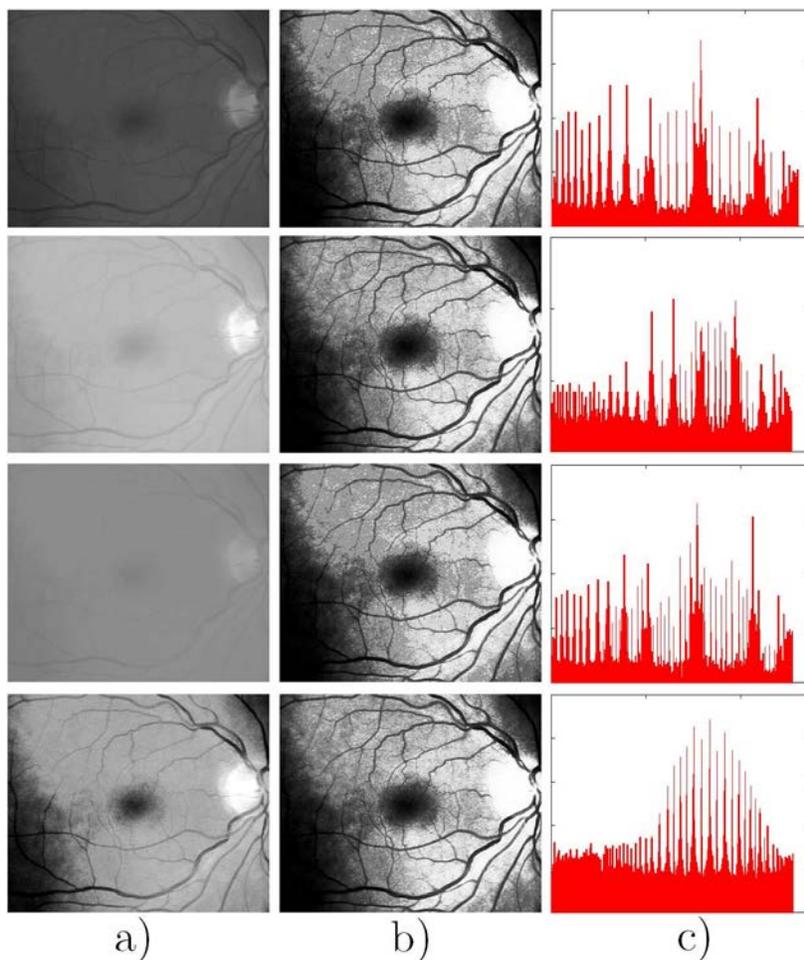


Figura 2.5: (a) Imágenes de la Figura 2.4, (b) Resultados de aplicar ecualización del histograma y (c) Histograma correspondiente.

En la Figura 2.5 se muestran cuatro imágenes, en las cuales se observa el resultado de realizar la ecualización del histograma a cada una. Los tres primeros resultados muestran una mejora significativa. Como era de esperar, la ecualización del histograma no produjo una diferencia visual significativa en la cuarta imagen porque el histograma de esta imagen ya abarcaba todo el espectro de la escala de grises [1].

2.3.2. Ecualización Adaptativa del Histograma (AHE)

Es una técnica de procesamiento de imágenes por computadora usada para mejorar el contraste en las imágenes. Éste método se diferencia de la ecualización del histograma ordinario en que el método adaptativo calcula varios histogramas, cada uno correspondiente a una sección distinta de la imagen, y los utiliza para redistribuir los valores de luminosidad de la imagen. Por lo tanto, es adecuado para el mejoramiento del contraste local de una imagen y obtener mayores detalles [3], [4].

Sin embargo, AHE tiene una tendencia a ampliar el ruido en regiones relativamente homogéneas de una imagen. Una variante de AHE es el llamado Contraste Limitado de la Ecualización Adaptativa del Histograma (CLAHE por sus siglas en inglés) el cual impide que se amplifique el ruido en la imagen [5].

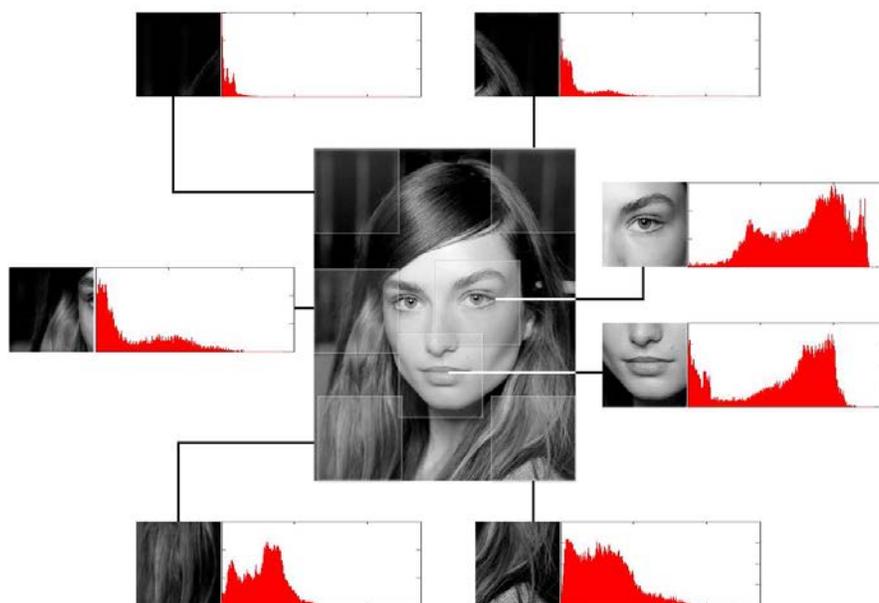


Figura 2.6: Función del método AHE.

La ecualización ordinaria de un histograma utiliza la misma transformación derivada del histograma de la imagen para transformar todos los píxeles. Esto funciona bien cuando la distribución de valores de los píxeles es similar a lo largo de la imagen. Sin embargo, cuando la imagen contiene regiones que son significativamente claras u oscuras que la mayoría de la imagen, el contraste en aquellas regiones no será suficientemente realizado.

AHE mejora esto transformando cada píxel con una función de transformación derivada de una región de vecindad y fue desarrollado por primera vez para su uso en pantallas de la cabina

de los aviones. De forma más simple, cada pixel es transformado basado en el histograma de un cuadro alrededor del pixel [6], esto se muestra en la Figura 2.6.

La derivación de las funciones de transformación de los histogramas es exactamente la misma como para una ecualización del histograma ordinario: La función de transformación es proporcional a la función de distribución acumulativa (CDF por sus siglas en inglés) de valores de pixel en la vecindad. En la Figura 2.7 se muestra el resultado de aplicar AHE.

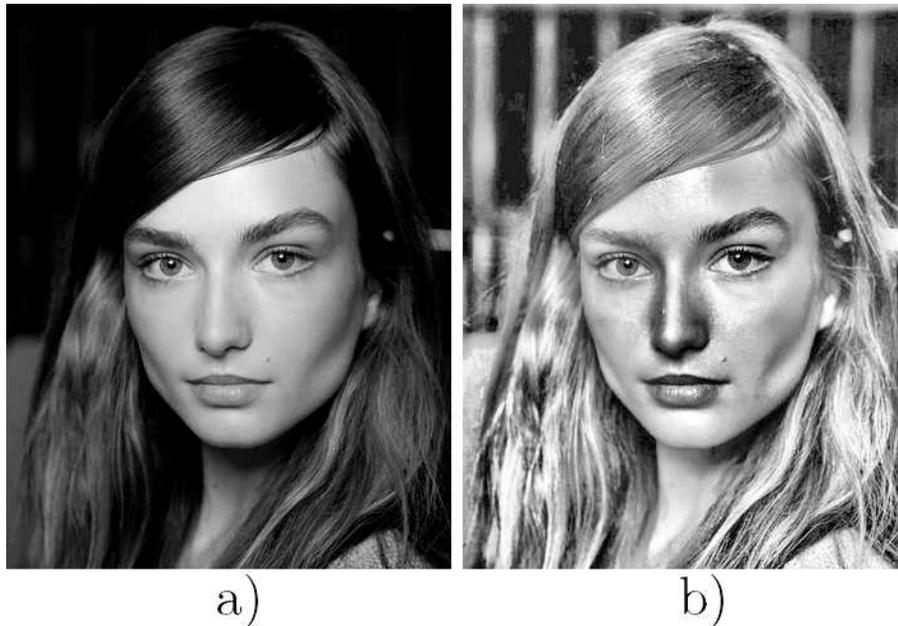


Figura 2.7: Resultado de aplicar Ecualización Adaptativa del Histograma con una vecindad de tamaño 5 pixeles. Nótese que en la imagen (b) el fondo ha sido realzado pero hay presencia de ruido.

Algunas propiedades de la técnica de ecualización adaptativa del histograma son:

- El tamaño de la region de vecindad es un parámetro del método. Esto constituye una característica de escala de longitud: el contraste en pequeñas escalas es realzado, mientras que el contraste en grandes escalas es reducido.
- Debido a la naturaleza de la ecualización del histograma, el valor de resultado de un pixel bajo AHE es proporcional a su rango entre los pixeles en su vecindad. Esto permite una implementación eficiente en hardware especializado que puede comparar el pixel central con los demás pixeles en la vecindad [6].

- Cuando la región de la imagen que contiene la vecindad de un pixel es bastante homogénea, su histograma estará fuertemente elevado (con picos), y la función de transformación mapeará un estrecho rango de valores de pixel a todo el rango de la imagen resultante. Esto causa que AHE amplíe pequeñas cantidades de ruido en regiones muy homogéneas de la imagen [7].

2.3.3. Contraste Limitado de la Ecuación Adaptativa del Histograma (CLAHE)

El contraste limitado de la ecuación adaptativa del histograma difiere de la ecuación adaptativa del histograma ordinario en su limitante de contraste. Esta característica también puede ser aplicada a la ecuación global del histograma, dando lugar a la Ecuación del Histograma con Contraste Limitado (CLHE), el cual es raramente utilizado en la práctica [8].

En el caso de CLAHE, el procedimiento de limitación de contraste tiene que ser aplicado para cada zona de la cual una función de transformación es derivada. CLAHE fue desarrollado para prevenir el aumento de ruido que la ecuación adaptativa del histograma puede llegar a dar [6], [8].

Esto se logra por la limitación del realce de contraste de AHE. La amplificación del contraste en la vecindad de un valor de pixel dado viene dada por la pendiente de la función de transformación.

Esta es proporcional a la pendiente de la función de distribución acumulativa (CDF) de la vecindad y por lo tanto, el valor del histograma en ese valor de pixel. CLAHE limita la amplificación recortando el histograma a un valor predefinido antes de calcular la CDF. Esto limita la pendiente de la CDF y por lo tanto, de la función de transformación.

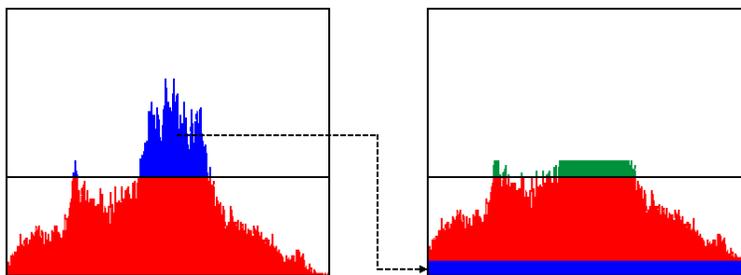


Figura 2.8: Equilibrio del límite de corte en el histograma.

El valor al que el histograma es recortado, llamado límite de corte, depende de la normalización del histograma y por lo tanto en el tamaño de la región de vecindad [8]. Comúnmente los valores límite resultantes de la amplificación están entre 3 y 4. Resulta conveniente no descartar la parte del histograma que supera el límite de corte, ya que esta parte debe redistribuirse de forma equivalente o equilibrada entre todos los intervalos del histograma [6] como se muestra en la Figura 2.8.

La redistribución moverá algunos intervalos sobre el límite de corte nuevamente (región verde en la Figura 2.8), resultando en un límite de corte efectivo que es mayor que el límite prescrito y el valor exacto de la cual depende de la imagen. Si esto no es lo deseable, el procedimiento de redistribución puede ser repetido recursivamente hasta que el exceso sea lo menor posible [9]. A continuación en la Figura 2.9 se muestra el resultado de aplicar CLAHE.

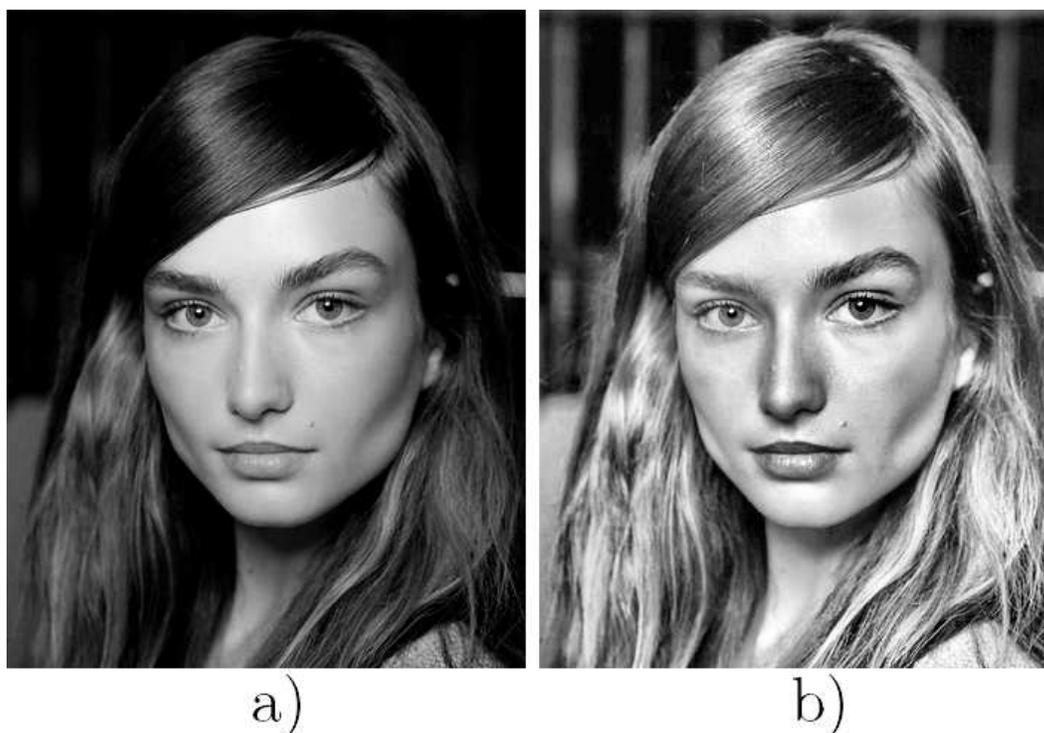


Figura 2.9: Resultado de aplicar Contraste Limitado de la Ecuación Adaptativa del Histograma con una vecindad de tamaño 5 píxeles. Nótese que en la imagen (b) el fondo ha sido realzado y la presencia de ruido ha sido disminuida.

2.3.4. Máscara de desenfoque (Unsharp Masking)

Es una técnica para dar nitidez a una imagen. La máscara de desenfoque es una de las técnicas típicamente usadas para la eliminación de ruido. La operación de desenfoque está dada por:

$$g(x, y) = f(x, y) + \alpha[f(x, y) - \bar{f}(x, y)] \quad (2.5)$$

donde $f(x, y)$ es la imagen original, $\bar{f}(x, y)$ es la versión suavizada de la imagen original, α es un factor del efecto de nitidez, y $g(x, y)$ es la imagen resultante [5], [6]. El término "desenfoque" deriva del hecho de que la técnica usa una imagen borrosa positiva, o "desenfocada" para crear una máscara de la imagen original [7], [14].

Debido a que la imagen positiva ha sido suavizada intencionalmente, solo la información de las bajas frecuencias ha sido cancelada. Además, la máscara reduce efectivamente el rango dinámico del negativo original. Así, si la imagen ampliada resultante se registra en papel fotográfico contrastado, la cancelación parcial enfatiza la información de las altas frecuencias espaciales (detalles finos) en la imagen original, sin pérdida de detalles de iluminación o sombra. La impresión resultante aparece más nítida que una hecha sin la máscara de desenfoque: el resultado es que la nitidez de la imagen ha incrementado.

El mismo principio de diferenciación es usado en la herramienta de máscara de desenfoque en muchos paquetes de software tales como Adobe Photoshop, PicMonkey y GIMP. El software aplica un desenfoque Gaussiano a una copia de la imagen original y esta es comparada con la original. Si la diferencia es mayor que un umbral especificado por el usuario, las imágenes son (en efecto) restadas. El control del umbral limita la nitidez de los elementos de la imagen que difieren entre sí por encima de un tamaño de umbral, de modo que la nitidez de los detalles pueden ser suprimidos.

La máscara de desenfoque digital es una flexible y poderosa forma de incrementar la nitidez especialmente en imágenes escaneadas. Desafortunadamente, esto puede crear efectos no deseados de bordes sobresalientes, o incrementar el ruido en la imagen. Los efectos no deseados pueden ser reducidos usando una máscara (particularmente una creada para detección de bordes) para que sólo sea aplicada en las regiones deseadas.

Típicamente una máscara de desenfoque digital tiene tres ajustes de control:

- Cantidad: Fuerza del efecto de nitidez, se especifica como un valor numérico. Un valor alto conduce a un mayor aumento del contraste de los píxeles. Los valores típicos para este parámetro son en el intervalo de 0 a 2, aunque valores mayores a dos son permitidos. Grandes valores de este parámetro puede llegar a crear efectos indeseables en la imagen de salida.

- **Radio:** Es la desviación estándar del filtro pasa bajas Gaussiano, se especifica como un valor numérico. Este valor controla el tamaño de la región alrededor de los píxeles del borde que serán afectados. Un valor grande realza regiones más amplias alrededor de los bordes, mientras que un valor pequeño realza regiones más estrechas alrededor de los bordes [7].

La Figura 2.10 muestra un ejemplo de la aplicación de la máscara de desenfoque para (a) con diferentes valores de Cantidad y Radio para (b) de 10 y 5 píxeles en el que la imagen se ve nítida, para (c) valores de 5 y 10 píxeles en la que se observa un aumento de contraste y para (d) con valores de 10 y 20 píxeles respectivamente. En ésta última se observa un notable aumento del contraste con respecto a las imágenes anteriores.

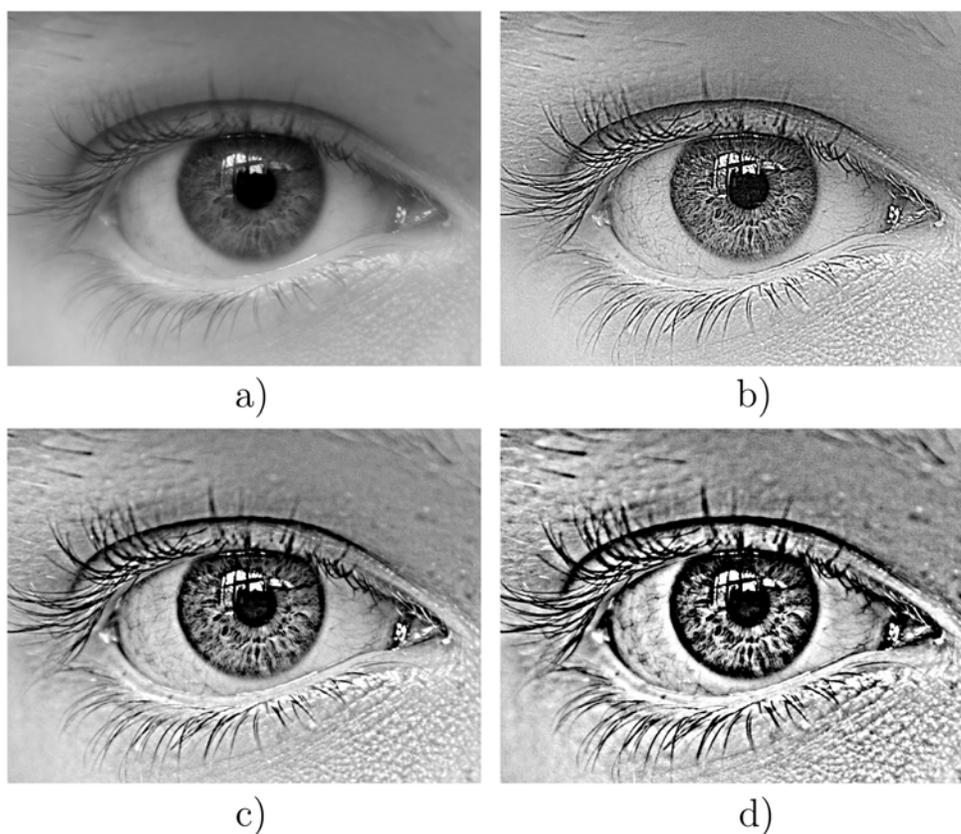


Figura 2.10: Ejemplo de mejoramiento de imágenes usando una máscara de desenfoque. (a) Imagen original, (b) Imagen nítida con valores de Cantidad y Radio de 10 y 5 píxeles (c) Imagen más nítida con valores de Cantidad y Radio de 5 y 10 píxeles y (d) Imagen muy nítida con valores de Cantidad y Radio de 10 y 20 píxeles.

2.4. Filtrado Espacial

El filtrado lineal de una imagen f de tamaño $M \times N$ con una máscara de filtro de tamaño $m \times n$ está dada por la expresión,

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (2.6)$$

donde, $a = \frac{(m-1)}{2}$ y $b = \frac{(n-1)}{2}$. Para generar una imagen filtrada completa esta ecuación debe ser aplicada para $x = 0, 1, 2, \dots, M - 1$ y $y = 0, 1, 2, \dots, N - 1$. De esta manera, aseguramos que la máscara procese todos los píxeles en la imagen.

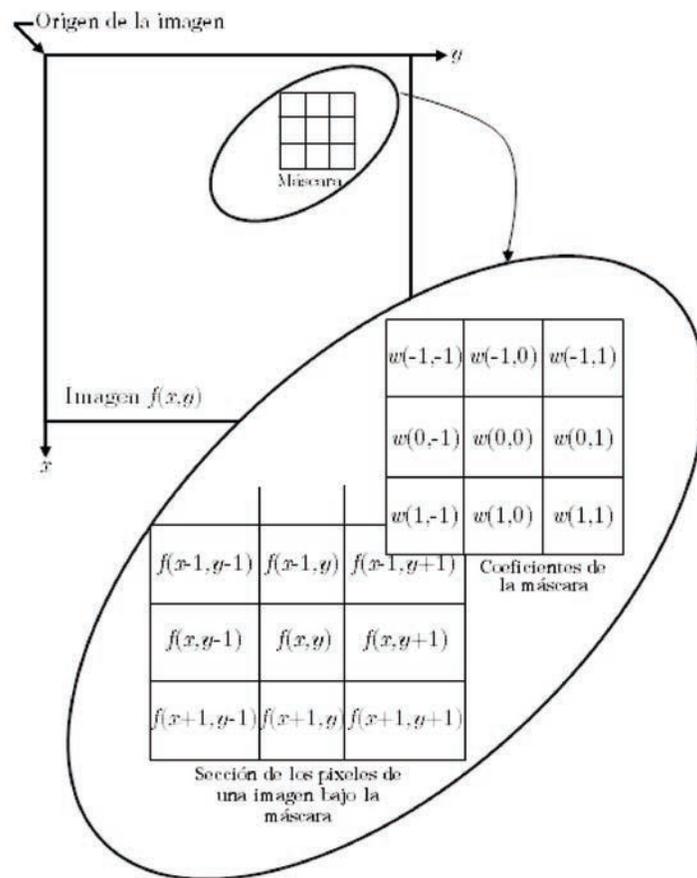


Figura 2.11: Mecanismo del filtrado espacial. La imagen grande muestra una máscara de 3×3 y la sección de la imagen bajo esta; la sección de la imagen se muestra desplazada debajo de la máscara para facilitar su legibilidad.

La mecánica del filtraje espacial se muestra en la Figura 2.11. El proceso consiste simplemente en mover la máscara de filtro de un punto a otro en una imagen. En cada punto (x, y) , la respuesta del filtro en aquel punto es calculado usando una relación predefinida. Para el filtraje espacial lineal, la respuesta es dada por una suma de productos de los coeficientes del filtro y los pixeles de la imagen correspondientes en el área extendida por la máscara de filtro.

El proceso de filtrado lineal dado en la Ecuación 2.6 es similar al concepto del dominio frecuencial llamado convolución. Por esta razón, el filtraje espacial lineal a menudo es referido como “convolución de una máscara con una imagen”. De manera similar, las máscaras de filtros son algunas veces llamadas máscaras de convolución. El término kernel de convolución también es de uso común. [1]

2.4.1. Filtro pasa-bajas

Los filtros pasa bajas son usados para suavizar y reducir el ruido. El suavizado es usado en el preprocesamiento para remover pequeños detalles en la imagen antes de la extracción de objetos. La reducción de ruido puede ser realizada por suavizado con un filtro lineal y también por filtrado no lineal.

La salida del suavizado es simplemente el promedio de los pixeles contenidos en la vecindad del filtro. Este filtro es algunas veces llamado filtro promedio o filtro pasa bajas como se muestra en la Figura 2.12.

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \quad \frac{1}{25} \times \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Figura 2.12: Filtro pasa bajas de tamaño 3×3 y 5×5 .

La idea detrás de los filtros de suavizado es sencillo. Mediante la sustitución de los valores de cada pixel en una imagen por el promedio de los niveles de gris en la vecindad definida por la máscara, este proceso da como resultado una imagen con reducción de transiciones bruscas en niveles de gris. Sin embargo, los bordes (que casi siempre son características deseadas de una imagen) también son caracterizados como transiciones bruscas en niveles de gris, por lo

que los filtros promedio tienen un efecto no deseado pues suavizan los bordes [1].

Uno de los principales usos de los filtros promedio es el de reducción de detalles "irrelevantes" en una imagen. Por "irrelevantes" se refiere a las regiones de píxeles que son pequeños con respecto al tamaño de la máscara. Este tipo de filtros también reducen el contraste en la imagen. En la Figura 2.13 se muestra el resultado de aplicar un filtraje pasa bajas.

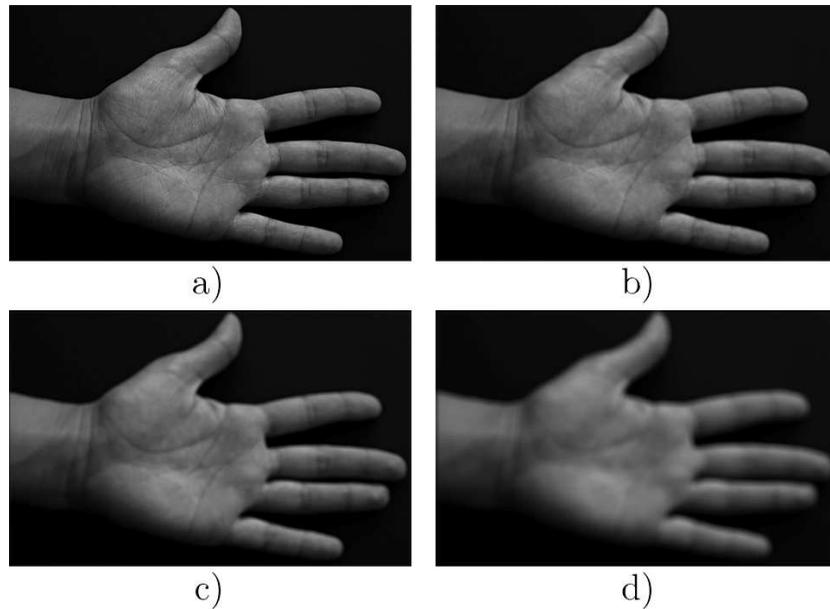


Figura 2.13: Ejemplo de filtrado pasa bajas (promedio). (a) Imagen original, (b) imagen filtrada con una máscara de 3×3 píxeles, (c) imagen filtrada con una máscara de 5×5 píxeles y (d) imagen filtrada con una máscara de 10×10 píxeles.

2.5. Segmentación de imágenes

Para realizar la identificación de estructuras anatómicas presentes en la imagen, se utilizan técnicas de segmentación, las cuales permiten dividir a una imagen en un conjunto de regiones. En muchas ocasiones, dependiendo de la aplicación, el proceso de segmentación tiende a ser uno de los pasos más difíciles y críticos para determinar la geometría de las diversas estructuras que componen la imagen y los valores apropiados para lograr esta determinación de estructuras.

De manera general las técnicas de segmentación tienden a ajustarse a las siguientes reglas:

1. Las regiones resultantes del proceso de segmentación debieran ser uniformes y homogéneas respecto a alguna característica, tal como lo es el nivel de gris o la textura.
2. Las regiones interiores debieran ser simples y no incluir abundantes huecos o estructuras con ruido.
3. Las regiones adyacentes en una segmentación debieran tener valores diferentes con respecto a la característica según la cual son uniformes.
4. Los límites de cada segmento debieran ser lo más simple posibles.

Como se puede observar, lograr que se cumplan estas propiedades resulta a menudo difícil y por lo general lo que logran la mayor parte de métodos, son regiones en las que comúnmente se observa la presencia de huecos y adicionalmente los límites o bordes de las mismas no son simples [1].

2.5.1. Umbralización

La umbralización de imágenes es uno de los métodos que abarca la segmentación de imágenes digitales. Las técnicas de umbralización buscan obtener un valor de umbral que permita binarizar a la imagen separando adecuadamente el fondo (background) y el objeto de interés a separar (foreground).

Muchas de las técnicas de umbralización están basadas en la información estadística que brinda el histograma, sobre todo en aquellas imágenes en donde los objetos tienen una superficie o textura homogénea y el fondo es más o menos uniforme.

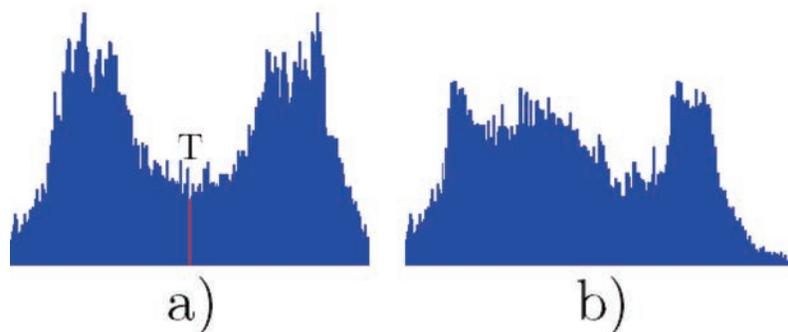


Figura 2.14: (a) Histograma ideal y (b) Histograma no ideal.

El problema de la umbralización es encontrar el valor T (umbral) adecuado entre los valores de grises en las imágenes para que permita una óptima separación entre el objeto y el fondo.

Si el histograma de una imagen posee 2 picos, podemos separar entonces 2 zonas o regiones y el umbral será aquel valor que se encuentra en el valle entre ambas. Este por supuesto es en el caso ideal, en el que el histograma tendría bien marcados los 2 picos, uno para el objeto y otro para el fondo, por lo tanto el umbral óptimo T es aquel que separa ambas regiones. Este ejemplo se ilustra en la Figura 2.14 (a).

Sin embargo, las imágenes siempre presentan ruido y, como consecuencia de esto, sus histogramas son difíciles de analizar como el que se presenta en la Figura 2.14 (b), es decir presentan histogramas sin una marcada separación de regiones. El ruido se debe a que la imagen depende de otros factores como la iluminación, movimiento, etc.

Actualmente existen numerosas técnicas de umbralización que han sido desarrolladas, ampliadas y modificadas por diversos autores pero, aun no hay un método general que se aplique a todas las imágenes y que nos brinde una umbralización perfecta [15].

Generalmente, lo anterior varía dependiendo de la imagen y de otros factores como el ruido, el fondo, los objetos a destacar, el tamaño de la imagen, y la más importante: el tipo de imagen, ya que no es lo mismo una imagen industrial a una imagen médica. En la Figura 2.15 se muestra un ejemplo de umbralización basada en un valor umbral del histograma.

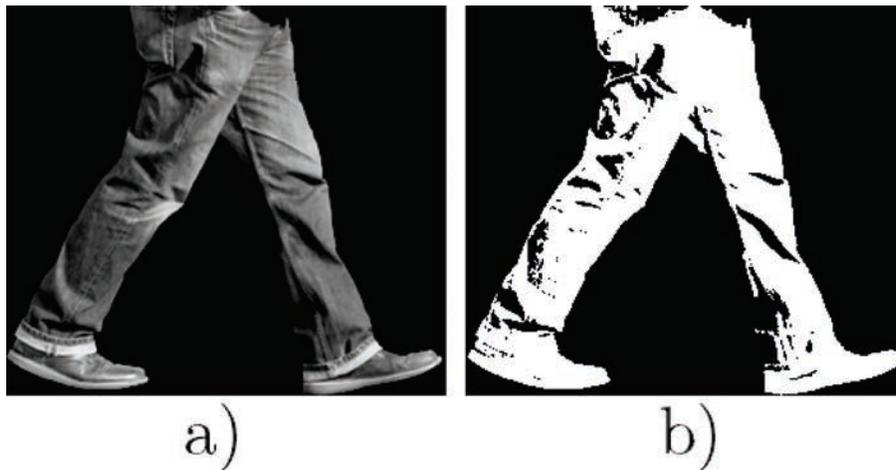


Figura 2.15: Ejemplo de umbralización basado en el histograma. (a) Imagen original y (b) imagen umbralizada.

2.5.2. Binarización

Es una técnica del procesamiento de imágenes que consiste en un proceso de reducción de la información de una imagen digital a dos valores: 0 (negro) y 255 (blanco). Esta técnica consiste en comparar cada píxel de la imagen con un determinado umbral pre-especificado. Los valores de la imagen que sean mayores que dicho umbral toman un valor blanco (255), mientras que el resto de píxeles tomarán un valor negro (0) [16].

En muchos casos, una imagen binaria es el resultado de una segmentación por niveles de gris o de una segmentación por selección de un rango de color determinado. En otros casos, una imagen binaria es simplemente el resultado de una selección interactiva de regiones de interés.

Trabajar con imágenes binarias resulta muy interesante debido a:

- La reducción al mínimo de los datos necesarios para la representación de la imagen permitiendo aprovechar al máximo la potencia computacional.
- Las propiedades geométricas y topológicas de los objetos presentes en la imagen, en las que se basan un gran número de aplicaciones industriales, puede obtenerse rápida y fácilmente a partir de las imágenes binarias.

Desde el punto de vista computacional, las imágenes binarias se procesan de manera mucho más rápidamente. Por este motivo, las primeras aplicaciones de visión artificial emplearon este tipo de imágenes. Hoy en día, se dispone de potentes procesadores para hacer frente a imágenes en niveles de gris y a pesar de eso, siguen siendo aún más numerosas las aplicaciones que trabajan sobre imágenes binarias por su simplicidad y robustez [17].

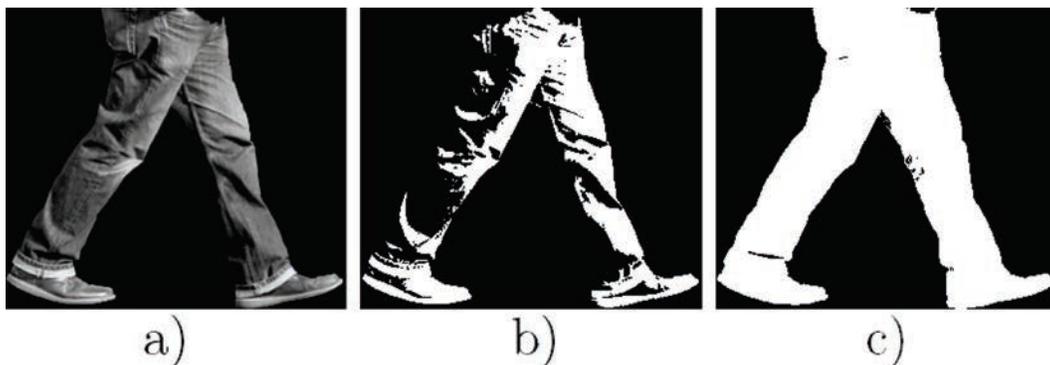


Figura 2.16: Ejemplo de binarización (a) Imagen original, (b) imagen con binarización local y (c) imagen con binarización global.

Los métodos de binarización se pueden dividir en dos tipos: Globales y Locales. Los métodos de la primera categoría tratan de encontrar un umbral para ser aplicado a toda la imagen. Los métodos locales obtienen el umbral para cada pixel en la imagen usando para ello los valores de sus vecinos, generalmente producen un mejor resultado al binarizar la imagen incluso en situaciones donde la iluminación es variable [18]. En la Figura 2.16 se muestra un ejemplo de binarización local y global de una imagen dada.

2.5.3. Binarización Adaptativa

Debido al hecho de que los valores de intensidad de niveles de gris varían a través de la imagen, las técnicas de umbralización globales por sí solas no proporcionan resultados satisfactorios.

La binarización adaptativa es un algoritmo de umbralización que combina el umbral adaptativo global y local: umbral local con reducción global. En esta técnica, se asignan diferentes valores de umbral para cada pixel de la imagen basado en el análisis de los niveles de gris de sus vecinos. Dichos umbrales corresponden al valor promedio de los niveles de gris de una vecindad de $M \times N$ alrededor de cada pixel restado por un umbral de corte global [19], [20]. Así, la función de transformación queda definida como,

$$g(x, y) = \begin{cases} 1 & \text{si } f(x, y) \geq \mu(x, y) - T_g \\ 0 & \text{en otro caso} \end{cases} \quad (2.7)$$

donde $f(x, y)$ es la imagen original, $\mu(x, y)$ es la imagen promedio de $f(x, y)$, T_g es un umbral de corte y $g(x, y)$ es la imagen binaria resultante. El tamaño de la máscara del filtro promedio para obtener $\mu(x, y)$ y el umbral de corte global T_g son elegidos en base a prueba y error ya que depende de la aplicación, la imagen y los elementos u objetos binarizar. Un ejemplo de ésta técnica se muestra en la Figura 2.17.

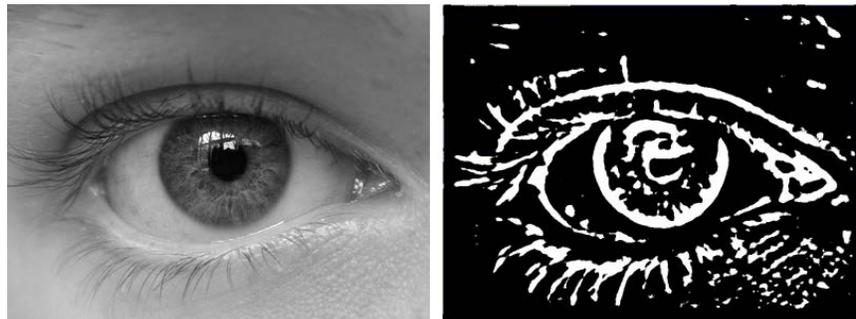


Figura 2.17: Ejemplo de binarización adaptativa.

2.6. Región de Interés (ROI)

Una Región de Interés (ROI) es una selección de muestras en un conjunto de datos idénticos, para un propósito en particular. Generalmente, la selección es un área cuadrada o rectangular especificada por el usuario la cual es usada para limitar ciertos cálculos dentro de los límites de ésta, además la ROI puede ser utilizada como técnica de segmentación de imágenes [21].

Principalmente se utiliza la ROI como un paso intermedio entre la decisión de utilizar la imagen adquirida de manera completa o únicamente un área en particular que comúnmente es la región en donde se encuentra el o los objetos de estudio o la mayor cantidad de información discriminante de una imagen.

A menudo hay situaciones en las que es necesario trabajar sobre toda la imagen y otras en las que utilizar por completo la imagen representan un exceso de carga de trabajo para el CPU que por lo general resulta lento e innecesario.

Al especificar la ROI que rodea sólo una área de menor tamaño, permite que la computadora trabaje de manera más rápida ya que sólo efectuará los cálculos exclusivamente en los pixeles seleccionados, por lo que es uno de los mejores métodos para evitar el exceso de costo computacional.

Generalmente algunos dispositivos de biometría utilizan delimitadores forzando a mantener el area del cuerpo en una posición fija durante la captura haciendo más fácil la obtención de puntos de referencia específicos y delimitar la región. Basado en la imagen original adquirida, la extracción de la ROI se puede hacer a través de los siguientes pasos:

Calcular el centroide de la imagen usando los momentos geométricos centrales con,

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (2.8)$$

Posteriormente se obtienen los puntos de las esquinas que conformarán la región de interés mediante,

$$\begin{aligned} E1 &= \bar{y} - d2 \\ E2 &= \bar{x} - d1 \\ E3 &= \bar{y} + d2 \\ E4 &= \bar{x} + d1 \end{aligned} \quad (2.9)$$

donde \bar{x} y \bar{y} so las coordenadas del centroide de la imagen, $d1$ y $d2$ son las distancias en x y y para el largo y ancho del cuadro delimitador de la ROI.

Una vez calculados los puntos $E1, E2, E3, E4$ se puede obtener la ROI utilizando,

$$ROI = (E1, E3)(E2, E4) \quad (2.10)$$

El proceso de extracción de la ROI de una imagen, es ilustrado en la Figura 2.18.

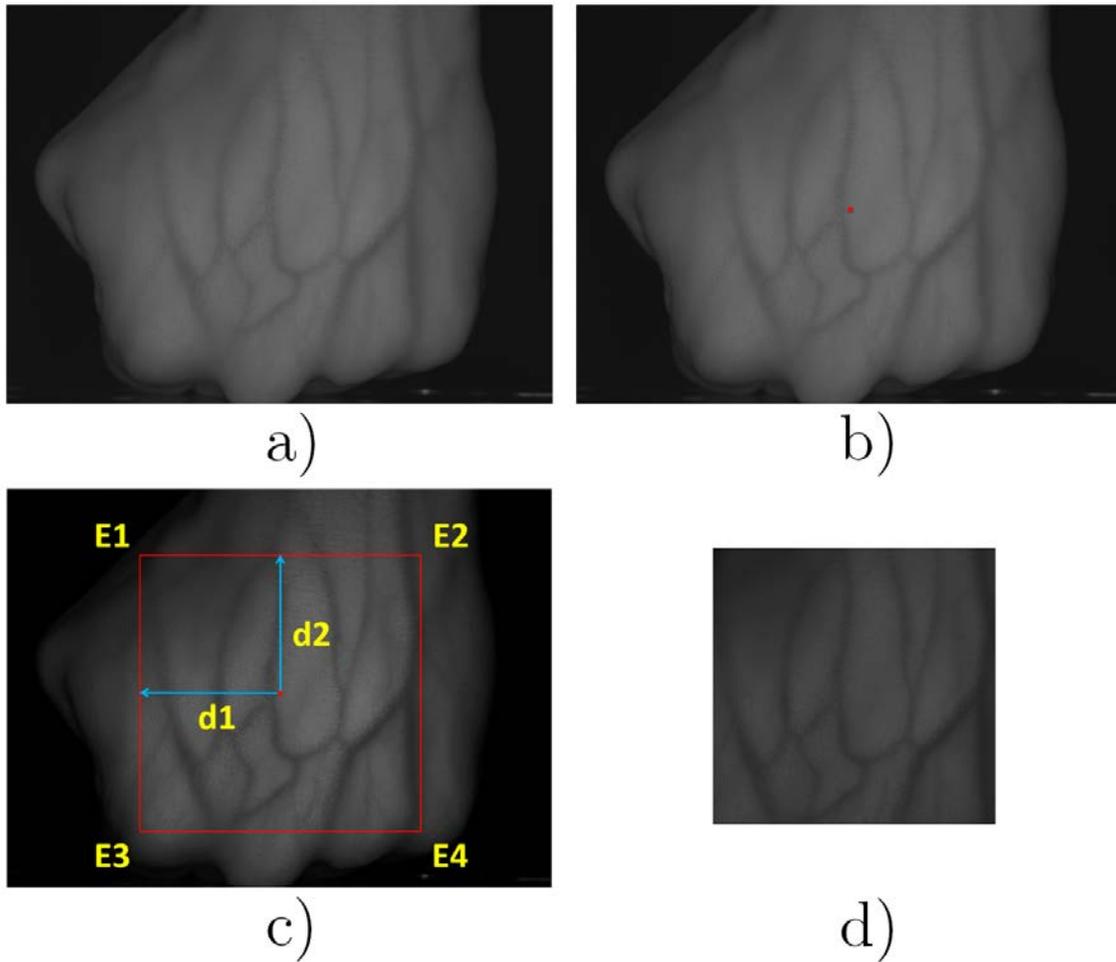


Figura 2.18: Ejemplo de extracción de la ROI. (a) Imagen original, (b) Centroide de la imagen, (c) Cálculo del cuadro que conformará la ROI y (d) Región de Interés extraída.

2.7. Morfología matemática

La morfología matemática es una teoría y técnica para el análisis y tratamiento de las estructuras geométricas, basada en la teoría de conjuntos, nació en 1964 como producto de la colaboración de Georges Matheron y Jean Serra en la *École des mines de Paris*, Francia.

Comúnmente es utilizada ampliamente en el campo de imágenes digitales, pero puede ser empleada también en gráficos, mallas poligonales, sólidos y muchas otras estructuras espaciales.

Las operaciones morfológicas simplifican imágenes y conservan las principales características de forma de los objetos. Un sistema de operadores de este tipo y su composición, permite que las formas subyacentes sean identificadas y reconstruidas de forma óptima a partir de sus formas distorsionadas y ruidosas, por lo que se puede usar, entre otros, con los siguientes objetivos:

- Preprocesamiento de imágenes (eliminación de ruido, simplificación de formas, crecimiento de regiones, ..., etc.)
- Destacar la estructura de los objetos (extraer el esqueleto, detección de objetos, envolventes convexas, ampliación, reducción, ..., etc.)
- Descripción de objetos (área, perímetro, ..., etc.)

Originalmente la morfología matemática se desarrollo para imágenes binarias y se extendió más tarde a funciones e imágenes en escala de grises. En imágenes binarias, las operaciones morfológicas como dilatación, erosión, transformada Hit-or-Miss, apertura y cierre, pueden ser utilizadas en aplicaciones como: extracción de fronteras y componentes conexas, llenado de regiones, adelgazamiento y engrosamiento, esqueletización.

2.7.1. Dilatación

La dilatación, es una operación de morfología muy utilizada y se describe como un crecimiento de píxeles, es decir, se marca con 1 la parte del fondo de la imagen que toque un pixel que forma parte de la región. Esto permite que aumente un pixel alrededor de la circunferencia de cada región y así poder incrementar dimensiones.

Ésta operación resulta muy útil para rellenar hoyos en una imagen o cuando es necesario unir regiones próximas que se han podido separar por una deficiente binarización.

El efecto de la dilatación dependerá del elemento estructural empleado. Dada una imagen A , y un elemento estructural B , (ambos imágenes binarias con fondo blanco), la dilatación de A por B se define como,

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (2.11)$$

donde \hat{B} es la reflexión de un conjunto B definido como,

$$\hat{B} = \{w | w = -b, \text{ para } b \in B\} \quad (2.12)$$

La traslación de un conjunto A por un punto $z = (z_1, z_2)$, denotado como $(A)_z$, se define como,

$$(A)_z = \{c | c = a + z, \text{ para } a \in A\} \quad (2.13)$$

Tengamos en cuenta que, para la intersección sólo consideramos los píxeles negros de A y B . El primer elemento de la dilatación, A , está asociado con la imagen que se está procesando y el segundo recibe el nombre de elemento estructural, es decir la forma que actúa sobre A en la dilatación para producir $A \oplus B$. En la Figura 2.19, se muestra el efecto que produce la operación morfológica de dilatación.

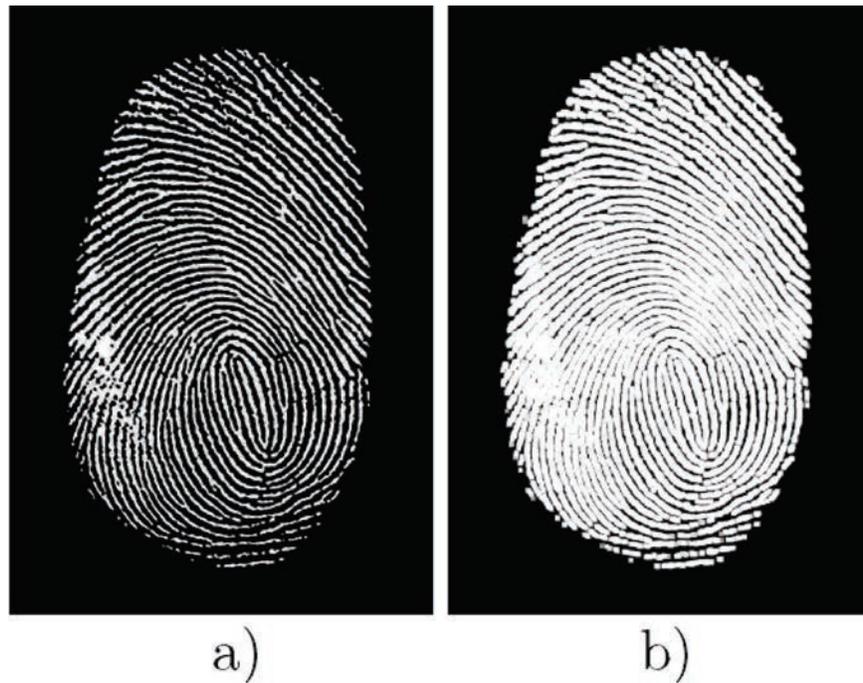


Figura 2.19: Ejemplo de dilatación. (a) Imagen original y (b) imagen dilatada con un elemento estructural de tamaño 3×3 .

2.7.2. Adelgazamiento

Es una operación morfológica usada para eliminar pixeles seleccionados en imágenes binarias. Puede ser utilizada para varias aplicaciones, pero es particularmente útil para la esquelización. En este modo, es comúnmente utilizada para ordenar la salida de detectores de bordes mediante la reducción del espesor de pixeles de todas las líneas que componen la imagen. Normalmente sólo se aplica a imágenes binarias produciendo como salida otra imagen binaria.

El adelgazamiento de un conjunto A por un elemento estructural B , denotado como $A \otimes B$, puede ser definido en términos de la transformada Hit-or-Miss,

$$A \otimes B = A - (A \circledast B) = A \cap (A \circledast B)^c \quad (2.14)$$

En términos generales, la operación de adelgazamiento es calculada trasladando el origen del elemento estructural a cada posible posición del pixel en la imagen, y en cada posición compararlo con los pixeles subyacentes de la imagen. Si los pixeles del objeto y el fondo del elemento estructural se relacionan exactamente con el objeto y el fondo de la imagen.

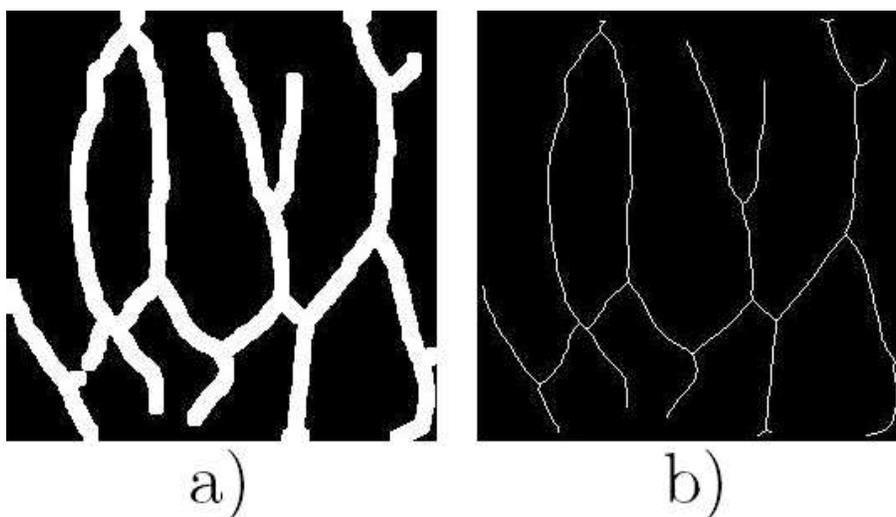


Figura 2.20: Ejemplo de dilatación. (a) Imagen original y (b) imagen adelgazada con un elemento estructural de tamaño 3×3 .

2.8. Conclusiones

En este capítulo se definieron las técnicas clásicas de procesamiento digital de imágenes para el mejoramiento del contraste de una imagen digital, el filtraje espacial pasa bajas para la eliminación de ruido e información innecesaria, así como la segmentación del área o región de interés a realzar, mejorar y/o describir.

Dentro de las técnicas de mejoramiento de contraste se encuentran las que están basadas en transformaciones de los niveles de gris de una imagen y las relacionadas con el uso y manipulación directa de su histograma.

Como se observó, los mejores resultados de mejoramiento y realce de información se obtienen a partir del uso de la técnica de Ecuilización Adaptativa del Histograma (AHE), el Contraste Limitado de la Ecuilización Adaptativa del Histograma (CLAHE) y de la Máscara de Desenfoque (Unsharp Masking) como se aprecia en las Figuras 2.7, 2.9 y 2.10.

En la imagen de la Figura 2.7, se puede apreciar en la parte del fondo que existe un aumento de ruido generado por el contraste de la imagen. Para el caso de la Figura 2.9, la limitación del contraste hace que este ruido desaparezca o que no aumente a medida que el contraste crece. Para el caso de la Máscara de Desenfoque, mientras mayor es el nivel del efecto de nitidez se genera en algunas ocasiones ruido parecido al que agrega el método de AHE.

La elección de los valores de control para estas tres funciones depende de la aplicación y de los elementos a realzar así como del preprocesamiento anterior que pudiera llegar a tener la imagen a mejorar.

Por otro lado, el método de binarización adaptativa, nos arroja mejores resultados que la binarización global y local debido a que no depende de una iluminación uniforme de la escena.

La extracción de la Región de Interés (ROI) se logró con el cálculo del centroide de una imagen. A partir del centro, se trazó un cuadro delimitador del área a analizar. Lo anterior, permite la reducción considerable del tiempo de preprocesado.

A partir de éstas imágenes segmentadas y binarizadas es posible extraer un conjunto de descriptores de forma. En el siguiente capítulo, se definen las técnicas implementadas para la extracción de características y su posterior clasificación.

Bibliografía

- [1] Ruben Medina, Gianfranco Passariello (Editor), "Imágenes médicas. Adquisición, Análisis, Procesamiento e Interpretación" *Equinoccio, Ediciones de la Universidad Simón Bolívar*, ISBN: 9802370924, 1995.
- [2] Rafael C. Gonzalez and Richard E. Woods. "Digital Image Processing", *Prentice Hall*, 2da Edición, (2002).
- [3] Jesús Silva y Wuilian Torres, Gianfranco Passariello (Editor), "Imágenes médicas. Adquisición, Análisis, Procesamiento e Interpretación" *Equinoccio, Ediciones de la Universidad Simón Bolívar*, ISBN: 9802370924, 1995.
- [4] Bashwati Das and Deepak Sharma, "Performance Enhancement of Content Based Image Retrieval System Using Contrast Limited Adaptive Histogram Equalization," *International Journal Of Engineering And Science*, vol. 3, pp. 29-36, 2013.
- [5] Shradha P. Dakhare and Manoj B. Chandak, "Performance Analysis of Image Filtering Technique and Combined Approach for Image Segmentation," *International Journal of Advanced Engineering Research and Technology (IJAERT)*, vol. 2, 2014.
- [6] D. P. Sharma, Dr. N. Chandra Sekhar Reddy, Prof. A. Praveen and Prof. Uma Shankar, "Problems of Intensity Transformation Methods and Selecting Appropriate Histogram Technique," *International Journal of Engineering and Technical Research*, vol. 2, 2014.
- [7] S. M. Pizer, E. P. Amburn, J. D. Austin, Robert Cromartie, ARI Geselowitz, Trey Greer, Bart Ter Haar Romeny, John B. Zimmerman and Karel Zuiderveld, "Adaptive Histogram Equalization and Its Variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, pp. 355-368, 1987.
- [8] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," *Graphics Gems IV, Academic Press*, ISBN: 0-12-336155-9, 1994.
- [9] Anupama Bhan, Anita Thakur and Garima Vyas, "Analysis of Histogram Based Compound Contrast Enhancement with noise reduction method for Endodontic therapy," *5th*

-
- International Conference- Confluence The Next Generation Information Technology Summit (Confluence)*, 2014.
- [10] D. P. Sharma, "Intensity Transformation using Contrast Limited Adaptive Histogram Equalization," *International Journal of Engineering Research*, vol. 2, pp, 282-285, Agosto 2013.
- [11] Jayaraman, "Digital Image Processing," *Tata McGraw-Hill Education*, ISBN: 9780070144798, 2011.
- [12] Joseph N. Wilson and Gerhard X. Ritter, "Handbook of Computer Vision Algorithms in Image Algebra," *CRC Press*, ISBN: 9781420042382, 2000.
- [13] www.scantips.com, "A few scanning tips, Sharpening - Unsharp Masking," <http://www.scantips.com/simple6.html>.
- [14] Hind Rostom Mohamed Shaban and Ameer Mohammad Husain, "Edge Detection for Objects in the Upper Triangle Gray Scale Image by Cholesky Decomposition and Unsharp Masking (USM)," *International Journal of Computer Applications*, vol. 85, January 2014.
- [15] Carlos A. Cattaneo, Ledda I. Larcher, Ana I. Ruggeri, Andrea C. Herrera y Enrique M. Biasoni, "Métodos de Umbralización de Imágenes Digitales Basados en Entropía de Shannon y Otros," *Asociación de Mecánica Computacional Argentina, Mecánica Computacional*, vol. XXX, pp. 2785-2805, 2011.
- [16] Javier Hernández Afonso, "Sistema de Detección y Almacenamiento de Variables Visualizadas por un Dispositivo Mediante Pantalla o Display a Través de un Sistema de Visión de Bajo Costo," *Universidad de la Laguna, Escuela Técnica Superior de Ingeniería Civil e Industrial*, 2014.
- [17] E. de la Fuente y F.M. Trespademe, "Visión Artificial Industrial. Procesamiento de imágenes para visión automática y robótica," *Universidad de Valladolid*, 2013.
- [18] Daniel Nieto Bonilla, "Capítulo 2. Técnicas de procesamiento digital de imágenes y reconocimiento de patrones," http://catarina.udlap.mx/u_dl_a/tales/documentos/lis/nieto_b_d/capitulo2.pdf, *Universidad de las Américas Puebla*.
- [19] Fernando Palero Molina, "Librerías Wrapper para el acceso a cámaras Kinect y herramientas para el calibrado de las cámaras RGB y profundidad," <https://riunet.upv.es/bitstream/handle/10251/15250/memoria.pdf?sequence=1>, *Universidad Politécnica de Valencia*, 2012.

- [20] Lingyu Wang, Graham Leedham and David Siu-Yeung Cho, "Minutiae feature analysis for infrared hand vein pattern biometrics," *Pattern Recognition*, vol. 41, pp. 920-929, 2008.
- [21] Ron Brinkmann, "The Art and Science of Digital Compositing," *Morgan Kaufmann, Editor*, ISBN: 9780121339609, 1999.

Capítulo 3

Técnicas de extracción de características y métodos de clasificación

3.1. Introducción

El reconocimiento y la clasificación de objetos en imágenes digitales es un área de investigación de mayor crecimiento en el campo de ingeniería y de ciencias de la imagen. La razón es evidente: la capacidad de reemplazar las habilidades visuales humanas con una máquina es una tarea muy importante y hay diversas aplicaciones. La idea principal es inspeccionar la escena de una imagen, mediante el procesamiento de datos obtenidos a partir de sensores.

Tales máquinas, pueden reducir substancialmente la carga de trabajo y mejorar la precisión en la toma de decisiones por parte de los operadores humanos en diversos campos, incluyendo: la defensa, sistemas de ingeniería biomédica, monitoreo de la salud, cirugías, sistemas de transporte inteligente, manufactura, robótica, entretenimiento y sistemas de seguridad.

El reconocimiento y clasificación es un campo multidisciplinario. Este requiere contribuciones de diversas tecnologías y del conocimiento en sensores, sistemas formadores de imágenes, algoritmos de procesamiento de imágenes / señales, hardware, software, entre otros.

Si bien ha habido avances significativos en el reconocimiento de imágenes y las tecnologías de la clasificación, existen grandes problemas técnicos y desafíos que enfrenta este campo. Estos incluyen grandes variaciones en las características del objeto inspeccionado debido a las condiciones del ambiente, variaciones geométricas, degradación y comportamiento del sensor, por ejemplo, las variaciones de una imagen térmica IR, los ángulos de reflexión, etc. Además, en muchas aplicaciones el objeto de interés es una pequeña parte de una escena muy compleja bajo inspección; es decir, la forma distorsionada del objeto está incrustada en un fondo ruidoso, tal como degradación ambiental, ruido del sensor, oclusión y/o camuflaje.

Algunas veces, los algoritmos son desarrollados con una disponibilidad de conjunto de datos de entrenamiento limitado, el cual puede no representar exactamente las variaciones actuales de los objetos o de la representación de la escena actual. Bajo estas condiciones adversas, un sistema fiable debe realizar el reconocimiento y clasificación en tiempo real y con alta probabilidad de detección y baja tasa de falsas alarmas [1].

Actualmente, no hay una sola técnica ganadora que pueda resolver todas las clases de problemas de reconocimiento y clasificación. En muchos casos, la solución parece ser dependiente de la aplicación y puede combinar un número de éstos métodos para obtener los resultados deseados.

Dada la necesidad de describir el contenido de manera automática y objetiva de la información de los objetos presentes en una imagen, aparecen las técnicas y métodos de extracción de descriptores o características. Para el caso de las imágenes se puede hacer referencia a los descriptores de imagen.

De manera general, los métodos de extracción de características deberían poseer las siguientes propiedades:

- Robustez: La extracción de características debe ser insensible al ruido de captura e iluminación.
- Simples: Representar las características extraídas de la imagen de forma clara y sencilla permitiendo un fácil entendimiento de su contenido.
- Discriminante: Las características extraídas deben servir para distinguir objetos o elementos de clases distintas y al mismo tiempo establecer una relación entre características similares.
- Repetible: El descriptor de características obtenido a partir del análisis de un objeto debe ser independiente del momento en que se obtenga.
- Invariancia: Deben de poseer determinadas invarianzas dependientes de la aplicación para poder relacionar las características aún bajo diferentes tipos de transformaciones como: Traslación, Rotación, Escala, Transformaciones no lineales de deformación, etc.
- Eficiente: Los recursos utilizados para obtener el descriptor deben ser aceptables para ser implementados en aplicaciones con restricciones de espacio y/o tiempo.

Existen diversos niveles de profundidad en cuanto a la interpretación del contenido hecha por los descriptores de características, dependiendo del nivel de abstracción que se trate. En el nivel más bajo se hallan los descriptores visuales, los cuales describen características como la forma, color, textura, movimiento, entre otros. Por otro lado, en un nivel más alto, se localizan

los descriptores específicos que contribuyen información acerca de los objetos de una escena. Este tipo de descriptores se apoyan en los descriptores de bajo nivel, esto con el fin de lograr la difícil tarea de efectuar una descripción semántica de las imágenes. Además, es posible clasificar a estos descriptores en dos grandes grupos:

- De información general: abarcan los descriptores conocidos como de bajo nivel, estos proporcionan una descripción del color, forma, textura y movimiento en la imagen.
- De información de dominio específico: llamados también descriptores semánticos, los cuales proveen información acerca de los objetos y eventos que constituyen la escena.

Al mismo tiempo, dentro de la categoría de los descriptores de información general, se pueden clasificar según el nivel de aplicación sobre el que actúan, es decir, sobre que áreas de una imagen se aplican las operaciones para obtener los resultados que componen el descriptor.

- Globales: Resumen el contenido de la imagen inspeccionada en un vector o matriz de características único.
- Locales: Actúan sobre regiones de interés identificadas previamente construyendo el vector de características para esa región tomando en cuenta tanto la información propia de la región inspeccionada como del vecindario en el que se encuentra dicha región.

En este capítulo se analizan diferentes técnicas de extracción de características basadas en momentos geométricos, invariantes de Hu y descriptores de venas. Los métodos fueron utilizados para el análisis y descripción de imágenes binarias. Además se describen los métodos de clasificación k vecinos próximos (k -nn) y Redes Neuronales Artificiales (RNA).

3.2. Patrones y Patrón de clases

Un patrón es un arreglo de descriptores. El nombre "característica" es usado a menudo en la literatura del reconocimiento de patrones denotando un descriptor. Un patrón de clases, es una familia de patrones que comparten algunas propiedades en común. El patrón de clases es denotado por $\omega_1, \omega_2, \dots, \omega_W$, donde W es el número de clases. El reconocimiento de patrones por máquina consiste en las técnicas de asignación de los patrones a sus respectivas clases automáticamente y con la menor intervención posible de personas. Los vectores de patrones son representados como,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (3.1)$$

donde cada componente, x_i representa el i -ésimo descriptor y n es el número total de tales descriptores asociados con el patrón. El vector de patrones es representado como columnas (es decir, matrices de tamaño $n \times 1$) o en su forma equivalente $x = (x_1, x_2, \dots, x_n)^T$, donde T indica la transpuesta [4].

A lo largo de este capítulo se describen dos técnicas para la extracción de características basadas en: funciones momento y descriptores de venas. Y como clasificadores se analizó el funcionamiento de dos métodos, el primero está basado en Redes Neuronales y el segundo en K -vecinos próximos.

3.3. Momentos geométricos

Los momentos son descriptores numéricos que pueden ser obtenidos a partir de una imagen dada. Con la ventaja que no sólo usan la silueta de un objeto, sino que toman en cuenta a todos los pixeles de la misma imagen. En la actualidad, son utilizados para reconocer la forma de los objetos presentes en una imagen. La teoría de los momentos provee una alternativa útil e interesante para la representación de la forma de un objeto.

Para una función 2-D continua $f(x, y)$, el momento de orden $(p + q)$ es definido como,

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, 3, \dots, \quad (3.2)$$

con p, q como índices de orden, (x, y) son coordenadas cartesianas, f es una función imagen continua no negativa. Para imágenes digitales de pixeles $M \times N$, la doble integral en la Ecuación 3.2 puede ser aproximada mediante sumatorias. Para calcular los momentos geométricos en el caso discreto se puede utilizar la siguiente expresión,

$$m_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y) \quad p, q = 0, 1, 2, 3, \dots, \quad (3.3)$$

La aplicación más útil de estos momentos, son los valores que pueden ser extraídos a partir de ellos tales como el área, la varianza, el centro de masa, etc. Adicionalmente, existen los momentos centrales los cuales son usados para el reconocimiento de un objeto independientemente de su posición en el eje cartesiano.

Los momentos centrales para una función 2-D continua $\mu_{p,q}$ están definidos como,

$$\mu_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) d(x - \bar{x}) d(y - \bar{y}), \quad (3.4)$$

por lo tanto, para calcular los momentos centrales en el caso discreto, se tiene la siguiente expresión,

$$\mu_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad p, q = 0, 1, 2, 3, \dots, \quad (3.5)$$

donde \bar{x} y \bar{y} son las coordenadas del centroide del objeto presente en la imagen, el cual esta dado por,

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (3.6)$$

Los momentos normalizados son usados para reconocimiento de objetos sin importar el tamaño o posición en el plano imagen. De modo que, los momentos normalizados son definidos por,

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma} \quad p, q = 0, 1, 2, 3, \dots, \quad (3.7)$$

donde $\gamma = \frac{p+q}{2} + 1$ para $p + q = 2, 3, \dots$

Con los momentos geométricos, el área $a = m_{00}$ de una figura puede ser calculada independientemente de su orientación y con los momentos centrales es posible distinguir un objeto a pesar de sí se trasladó. Además, los momentos normalizados pueden ser usados para reconocer objetos sin importar el tamaño dentro de la imagen.

A partir de los momentos normalizados se pueden obtener un conjunto de siete momentos invariantes estandarizados. Este conjunto de momentos es conocido como momentos de Hu. Estos momentos son usados para calcular los momentos de un objeto sin importar la rotación, traslación y cambios en la escala del objeto.

Este conjunto de momentos está definido de la siguiente manera [4],

$$\begin{aligned} \phi_1 &= \eta_{2,0} + \eta_{0,2} \\ \phi_2 &= (\eta_{2,0} - \eta_{0,2})^2 + 4(\eta_{1,1})^2 \\ \phi_3 &= (\eta_{3,0} - 3\eta_{1,2})^2 + (3\eta_{2,1} - \eta_{0,3})^2 \\ \phi_4 &= (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2 \\ \phi_5 &= (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3})[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] \\ \phi_6 &= (\eta_{2,0} - \eta_{0,2})[(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{2,1} + \eta_{0,3}) \\ \phi_7 &= (3\eta_{2,1} - \eta_{0,3})(\eta_{3,0} + \eta_{1,2})[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{1,2} - \eta_{3,0})(\eta_{2,1} + \eta_{0,3})[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] \end{aligned}$$

En la Figura 3.1 se muestra un ejemplo de invariancia de los descriptores a transformaciones en escala, rotación y traslación de los objetos u elementos en el plano imagen. El descriptor usado fue el ϕ_1 . En la gráfica (c) de la Figura 3.1, se observa una invariancia de los descriptores calculados a partir de las imágenes correspondientes a la persona 1 (a) y a la persona 2 (b).

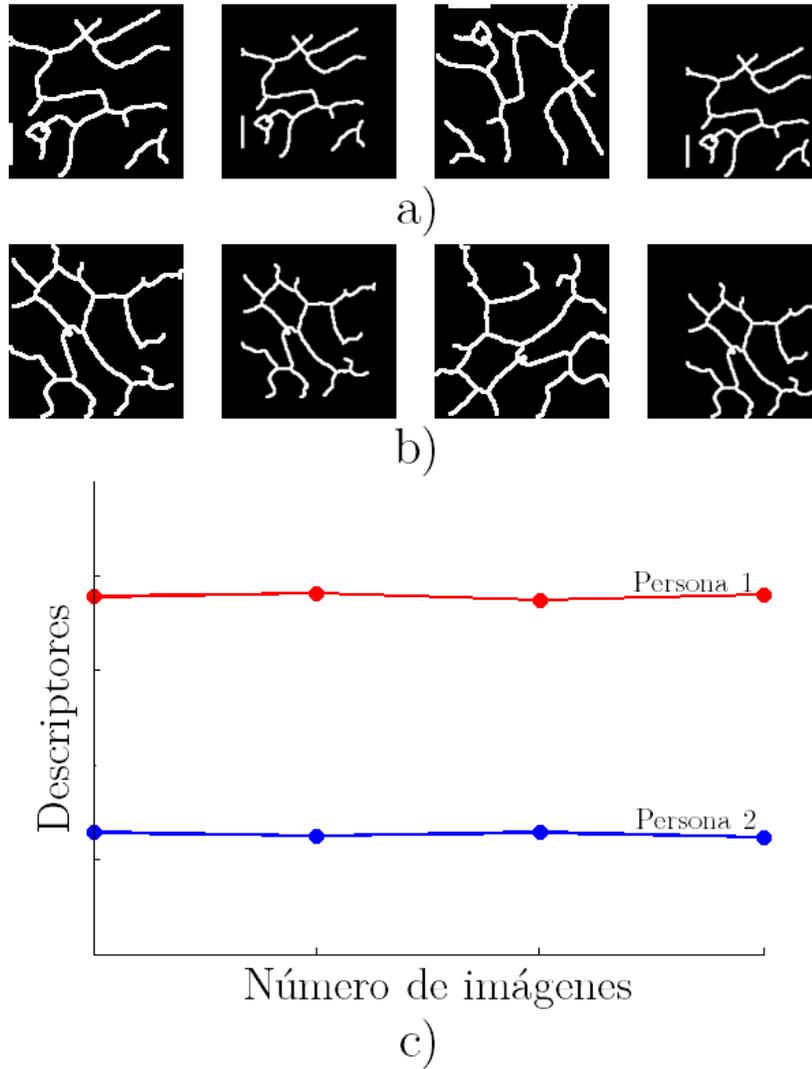


Figura 3.1: De izquierda a derecha para (a) Persona 1 y (b) Persona 2: imagen original, escalada, rotada y desplazada y (c) gráfica de los descriptores invariantes de Hu de 8 imágenes para 2 personas diferentes.

3.4. Descriptores de venas

Con este método se busca obtener un vector característico de cada imagen pero con una menor cantidad de datos en comparación a la imagen original y manteniendo información discriminante del patrón de venas. Para extraer las características que proporcionen información relevante de la estructura del patrón de venas, se realiza un conteo de pixeles presentes en una región mediante un sub-muestreo en la imagen del patrón de venas.

El sub-muestreo consiste en subdividir en pequeños cuadros o regiones de igual tamaño la imagen a la que se desea obtener los descriptores. La expresión para obtener los descriptores de las venas esta dada por,

$$V_{i,j} = \begin{cases} 1 & \text{if } Np_{i,j} \geq \delta \\ 0 & \text{en otro caso} \end{cases}, \quad (3.8)$$

donde Np es el número de pixeles en las sub-regiones, δ es un umbral de corte y V son los descriptores de venas vectorizados. Teniendo en cuenta que para el buen funcionamiento del mismo se deben definir las sub-divisiones que pueden ser en números pares o en potencias de 2, esto debido a que depende del tamaño de la imagen a sub-muestrear.

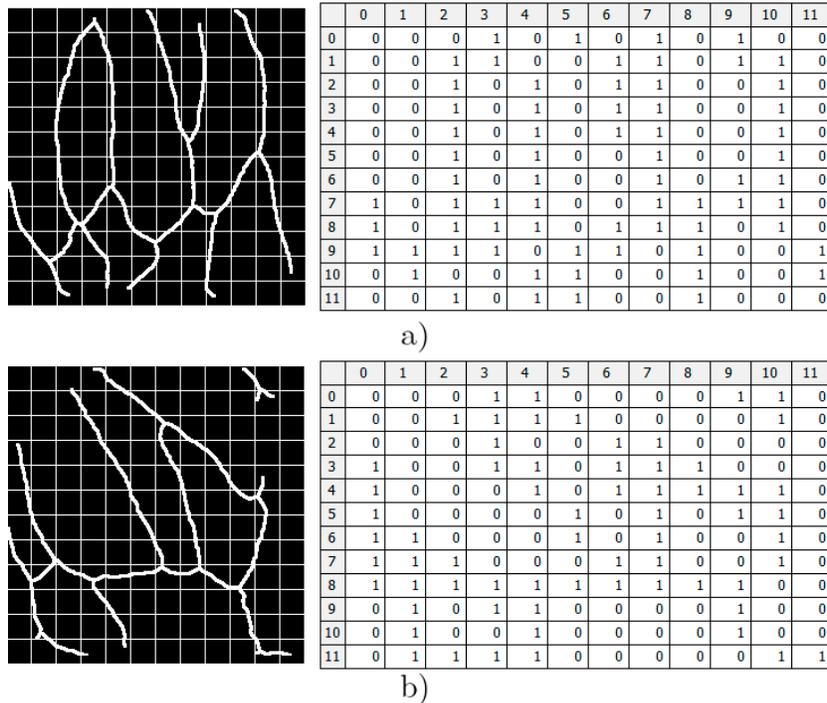


Figura 3.2: Extracción de descriptores del patrón de venas de la (a) Persona 1, (b) Persona 2.

En la Figura 3.2 se muestra un ejemplo del uso de esta técnica de extracción de características en la que se utilizó un umbral de corte $\delta=57$ con 12 sub-regiones o sub-muestras en una imagen de tamaño 264×264 píxeles.

3.5. Clasificadores

El término “clasificador” es empleado cuando se hace referencia al algoritmo que es utilizado para asignar un elemento u objeto sin etiqueta a una categoría o clase conocida. Este algoritmo, permite ordenar por clases dichos elementos a partir de la información característica de estos. En algunas ocasiones, es utilizado para referirse a la función matemática que implementa el algoritmo de clasificación.

Una forma de implementar un clasificador, es seleccionar un conjunto de elementos etiquetados y definir una regla que permita asignar una etiqueta o clase a cualquier otro elemento de entrada.

Para la implementación de un clasificador es necesario tener en cuenta una serie de características concretas. La selección de éstas, sin embargo, no es una tarea sencilla. La adición de muchos parámetros irrelevantes hace más difícil la clasificación para todos los métodos. Además, a medida que vamos añadiendo más información se van incrementando las dimensiones del espacio, hecho que supone que la optimización sea cada vez más difícil.

La finalidad del entrenamiento es construir un conjunto/grupo para que los objetos que se han clasificado como correctos, sean conocidos/identificados. Los parámetros característicos que describen a los objetos deben ser discriminatorios para la clasificación.

El conjunto de entrenamiento debe contener una lista de objetos con tipos conocidos. Idealmente este conjunto de entrenamiento debería contener muchos ejemplos, de este modo, se incluirían objetos comunes y no comunes.

Para crear el conjunto de entrenamiento se requiere una fuente de objetos clasificados de forma correcta. Una vez construido el clasificador se debe medir la precisión. Este paso es necesario tanto en la aplicación del clasificador como también para poder compararlo con otros clasificadores diferentes.

La precisión se puede determinar aplicando al clasificador un entrenamiento independiente de un conjunto de objetos de los que se conoce la clasificación. Se usan conjuntos conocidos ya que a veces no se tienen las fuentes necesarias para construir un nuevo modelo que se utilice únicamente para su prueba. Se debe evitar el entrenamiento y prueba con el mismo conjunto.

Los clasificadores tienen aplicaciones muy amplias, su uso se extiende desde la medicina, telefonía, robótica, juegos, procesos de fabricación, entre otros. En el caso de sistemas de visión por computadora, se utilizan para el reconocimiento facial, de voz, seguimiento de objetos, minería de datos, etc.

3.6. Redes Neuronales Artificiales (RNA)

Una red neuronal es un conjunto de elementos de procesamiento interconectados, unidades o nodos cuya funcionalidad está basada en una neurona biológica. La capacidad de procesamiento de la red es almacenada entre las unidades de conexión de puntos fuertes o pesos, obtenidos por un proceso de adaptación o aprendizaje mediante un conjunto de patrones [8]. Pero para entender un poco más acerca de las Redes Neuronales, hay que comprender el funcionamiento básico de una neurona biológica.

3.6.1. La neurona biológica

La teoría y modelado de las redes neuronales artificiales ha estado inspirada en la estructura y funcionamiento de los sistemas nerviosos, donde la neurona es el elemento fundamental. Una neurona, es una célula viva y, al ser una célula, contiene los mismos elementos que forman parte de todas las células biológicas. Además, estas contienen algunos elementos característicos que las diferencian de otras células. Uno de estos, es su capacidad de comunicación.

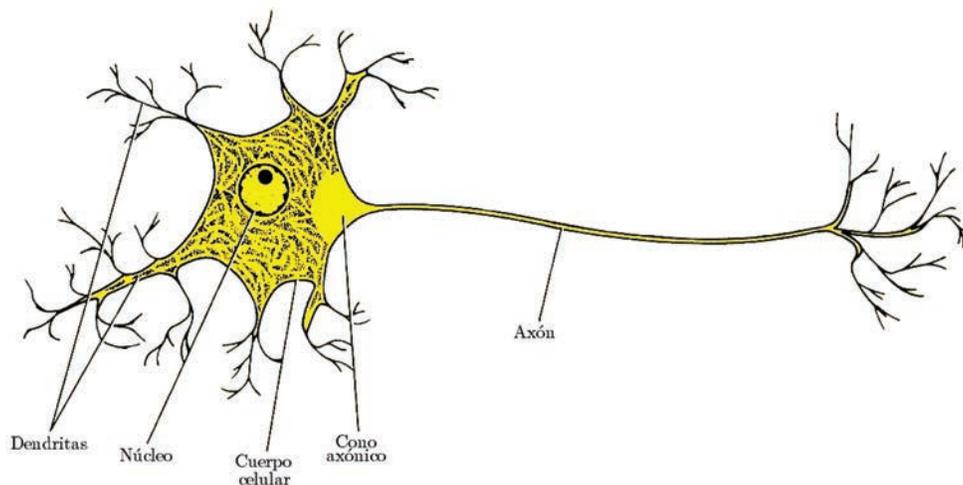


Figura 3.3: Elementos de una neurona biológica.

De manera general, una neurona consta de un cuerpo celular más o menos esférico de aproximadamente 5 a 10 micras de diámetro, de este cuerpo celular salen una rama principal, el axón, y las dendritas. De igual forma, el axón puede llegar a producir ramas en torno a su punto de arranque, y con frecuencia se ramifica extensamente cerca de su extremo [9]. Estos elementos se pueden visualizar en la Figura 3.3.

El cerebro humano se compone de un estimado de 10^{11} (100 billones) de células nerviosas o neuronas. Las neuronas se comunican a través de señales eléctricas que son impulsos de corta duración o "picos" en la tensión de la pared celular o membrana que recubre a la neurona.

Las conexiones son hechas por uniones electroquímicas llamadas sinapsis, que se producen en las ramas de la célula a las que se refiere como dendritas. Cada neurona recibe típicamente miles de conexiones de otras neuronas y por lo tanto está recibiendo constantemente una cantidad enorme de señales entrantes, que finalmente alcanza el cuerpo de la célula.

Es aquí donde se integran todas estas señales entrantes y, en términos generales, si la señal resultante de esta integración supera cierto umbral, entonces la neurona va a "disparar" o generar un impulso como respuesta. Esta respuesta se transmite entonces a otras neuronas por medio del axón y es conocida como sinapsis [9].

3.6.2. Neurona artificial

Las redes neuronales artificiales son modelos que intentan reproducir el comportamiento del cerebro. Como tal modelo, realiza una simplificación, averiguando cuáles son los elementos relevantes del sistema, bien porque la cantidad de información disponible puede ser excesiva o bien porque es redundante.

Cualquier modelo de red neuronal consta de dispositivos elementales de proceso: las neuronas. A partir de ellas, se pueden generar representaciones específicas, de tal forma que un conjunto de ellas puede significar una letra, un número o un objeto.

La neurona artificial pretende mimetizar (imitar) las características más importantes de las neuronas biológicas. Cada neurona i -ésima está caracterizada en cualquier instante por un valor numérico denominado valor o estado de activación $a_i(t)$; asociado a cada unidad, existe una función de salida f_i , que transforma el estado actual de activación en una señal de salida, y_i . Las señales moduladas que han llegado a la j -ésima unidad se combinan entre ellas, generando así la entrada total, Net_j como,

$$Net_j = \sum_i y_i w_{i,j} \quad (3.9)$$

donde $w_{i,j}$ son los pesos. Una función de activación, F , determina el nuevo estado de activación $a_j(t+1)$ de la neurona, teniendo en cuenta la entrada total calculada y el anterior estado de activación $a_j(t)$ [9].

3.6.3. El perceptrón

Fue el primer modelo de red neuronal artificial desarrollado por Rosenblatt en 1958. Despertó un enorme interés en los años 60, debido a su capacidad para “aprender” a “reconocer” patrones sencillos. Un perceptrón está formado por varias neuronas lineales que reciben las entradas a la red y una neurona de salida, que es capaz de decidir cuándo una entrada presentada a la red pertenece a una de las dos clases que es capaz de “reconocer”. En la Figura 3.4 se muestra la representación gráfica de un perceptrón.

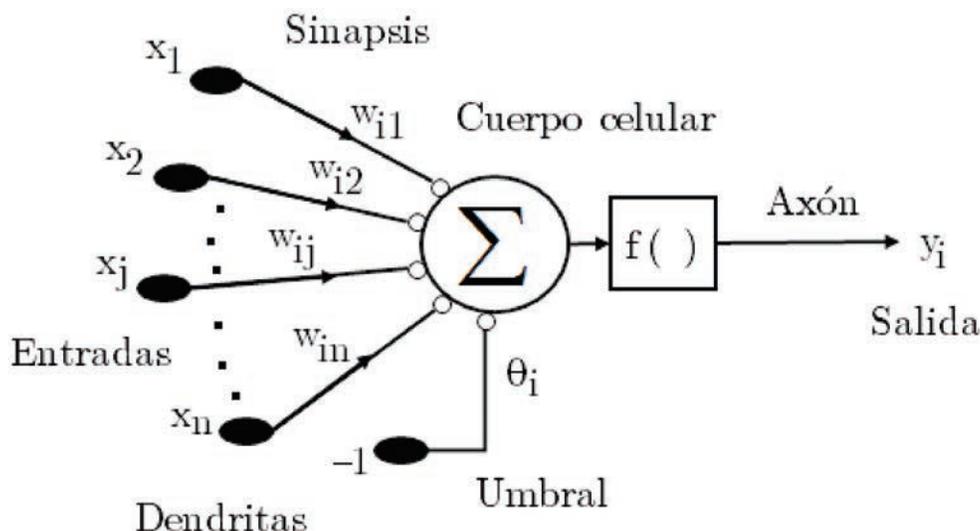


Figura 3.4: Representación gráfica del perceptrón.

Una técnica utilizada para analizar el comportamiento de redes como el perceptrón es representar en un mapa las regiones de decisión creadas en el espacio multidimensional de entradas a la red. En estas regiones se visualiza qué patrones pertenecen a una clase y cuáles a otra.

El perceptrón separa regiones por un hiperplano cuya ecuación queda determinada por los pesos de las conexiones y el valor umbral de la función de activación de la neurona. En este caso, los valores de los pesos pueden fijarse o adaptarse utilizando diferentes algoritmos de entrenamiento de la red.

Sin embargo, el perceptrón, al constar sólo de una capa de entrada y otra de salida con una única neurona, tiene una capacidad de representación bastante limitada. Este modelo sólo es capaz de discriminar patrones muy sencillos, linealmente separables. El caso más conocido es la imposibilidad del perceptrón de representar la función OR-EXCLUSIVA.

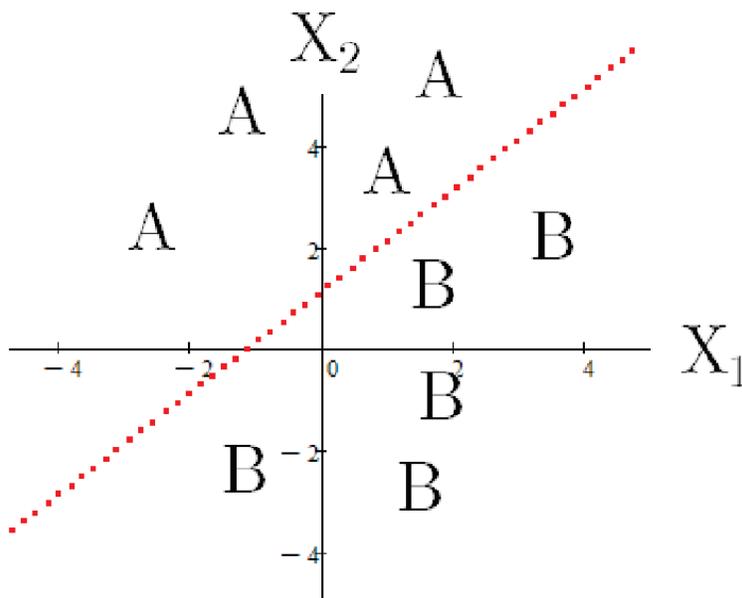


Figura 3.5: Clases linealmente separables.

La separabilidad lineal limita a las redes con sólo dos capas a la resolución de problemas en los cuales el conjunto de puntos sean separables geoméricamente. En el caso de dos entradas, la separación se lleva a cabo mediante una línea recta como se muestra en la Figura 3.5.

Para tres entradas, la separación se realiza mediante un plano en el espacio tridimensional, y así sucesivamente hasta el caso de N entradas, en el cual el espacio N -dimensional es dividido en un hiperplano [9]. Lo anterior se explica en la siguiente sección con ayuda de la Figura 3.6.

3.6.4. Tipos de redes neuronales artificiales

La topología o arquitectura de las redes neuronales consiste en la organización y disposición de las neuronas en la red formando capas o agrupaciones de neuronas más o menos alejadas de la entrada y salida de la red. En este sentido, los parámetros fundamentales de la red son:

- El número de capas.

- El número de neuronas por capa.
- El grado de conectividad.
- El tipo de conexiones entre neuronas.

Cuando se realiza una clasificación de las redes en términos topológicos como en la Figura 3.6, se suele distinguir entre las redes con una sola capa o nivel de neuronas y las redes con múltiples capas (2, 3, etc.).

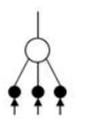
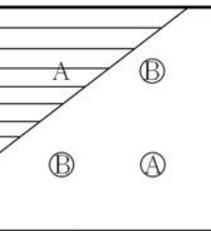
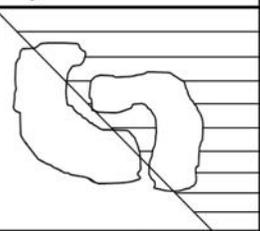
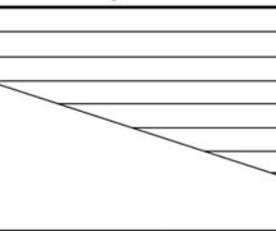
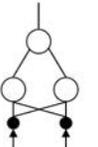
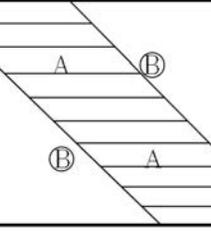
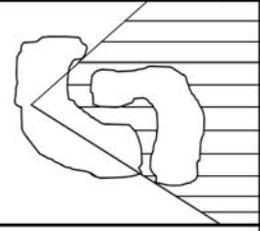
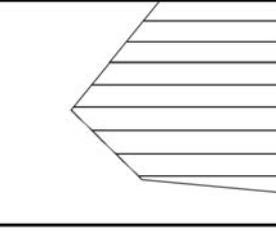
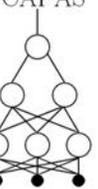
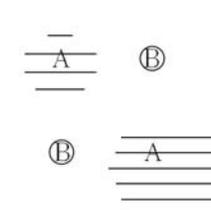
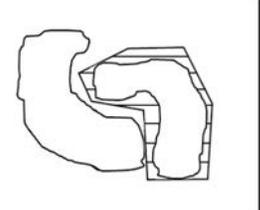
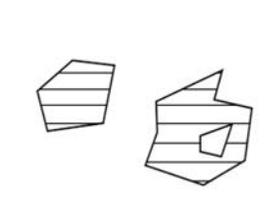
Estructura	Regiones de decisión	Problema de la XOR	Clases con regiones mezcladas	Formas de regiones más generales
<p>2 CAPAS</p> 	<p>MEDIO PLANO LIMITADO POR UN HIPERPLANO</p>			
<p>3 CAPAS</p> 	<p>REGIONES CERRADAS O CONVEXAS</p>			
<p>4 CAPAS</p> 	<p>ARBITRARIA COMPLEJIDAD LIMITADA POR EL NÚMERO DE NEURONAS</p>			

Figura 3.6: Clasificación de las Redes Neuronales Artificiales [9].

3.6.5. Función de activación

En un perceptrón el envío de señales o impulsos inhibitorios o excitatorios como lo haría una neurona biológica está regulada o modelada por la función de activación, siendo la que se encarga de enviar o no un impulso de salida. Una función de activación cumple generalmente con las siguientes propiedades:

- La decisión inhibitoria o excitatoria se encuentra centrada en el origen.

- Los valores de salida están frecuentemente acotados en el intervalo de $[-1,1]$ o $[0,1]$, de tal manera que cuando el valor absoluto de la integración sea muy grande, el valor de salida o activación permanece pequeño o constante.

Existen cuatro funciones de activación clásicas que determinan distintos tipos de neuronas, las cuales son [9]:

- Función escalón
- Función sigmoideal
- Función Gaussiana
- Función lineal y mixta

En la Figura 3.7 se muestra el perfil de la función de activación Sigmoideal la cual es utilizada en este trabajo de Tesis.

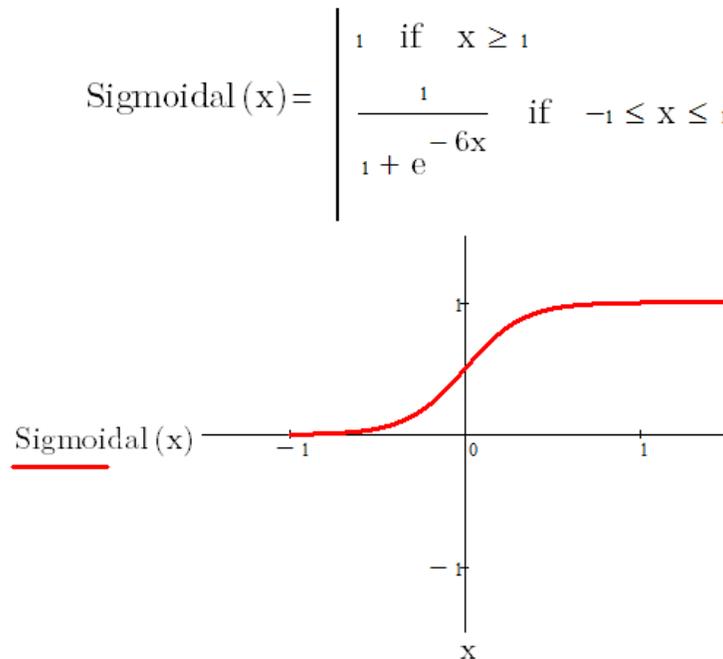


Figura 3.7: **Función Sigmoideal.**

3.6.6. Regla de aprendizaje del perceptrón

El algoritmo que se utiliza para el "aprendizaje" del perceptrón es de tipo supervisado, es decir que por cada ciclo de entrenamiento, al modelo se le presenta un patrón de entrada y uno deseado de salida. Para que el modelo pueda asociar los patrones de entrada con los de salida se necesita una regla de aprendizaje cuyo funcionamiento consiste en modificar los pesos del perceptrón para ajustar la frontera de decisión con la separación de clases, es decir, que si los patrones son linealmente separables, el perceptrón puede encontrar la solución.

Es muy común que se pueda llegar a encontrar la siguiente regla de aprendizaje:

$$\begin{aligned}w_i(t+1) &= w_i(t) + \alpha[e(t)x_i(t)] \\e(t) &= d(t) - y(t) \\(0 \leq i \leq N)\end{aligned}$$

donde t es el número de iteración, $d(t)$ es el valor deseado de salida, $y(t)$ es el valor de salida de la función de activación, $e(t)$ es el error, $x_i(t)$ es el elemento i del patrón de entrada, $w_i(t)$ es el peso correspondiente al elemento i del patrón de entrada y α es el factor de aprendizaje, este es ajustado experimentalmente en un intervalo sugerido de 0.025.

Con esta regla de aprendizaje el perceptrón es capaz de ubicar la frontera de decisión en cualquier punto del espacio de soluciones, la frontera de decisión en un ciclo completo de entrenamiento se va a modificar t veces, siendo cada t una iteración. Si en una iteración el error es 0, resulta evidente que el peso no va a ser modificado.

El factor de aprendizaje α es un parámetro de velocidad de aprendizaje, si este valor es muy pequeño es posible que el entrenamiento necesite un gran número de iteraciones, por otro lado, si es muy grande aumentan las probabilidades de que no exista convergencia [9].

En la Figura 3.8 se muestra un ejemplo de entrenamiento de un perceptrón para la función AND y OR en el que se puede apreciar que el perceptrón encuentra la solución para la separabilidad de dos clases. En el caso de la función XOR se pretende que para los valores de entrada (0,0) y (1,1) se devuelva a la clase 0, y para los patrones (0,1) y (1,0), la clase 1. Como puede comprobarse en la Figura 3.9, el problema radica en que no existe ninguna recta que separe los patrones de una clase de los de la otra.

La solución podría darse si se descompusiera el espacio en tres regiones: una región pertenecería a una de las clases de salida y las otras dos pertenecerían a la segunda clase. Si en lugar de utilizar únicamente una neurona de salida se utilizaran dos, se obtendrían dos rectas, delimitando tres zonas.

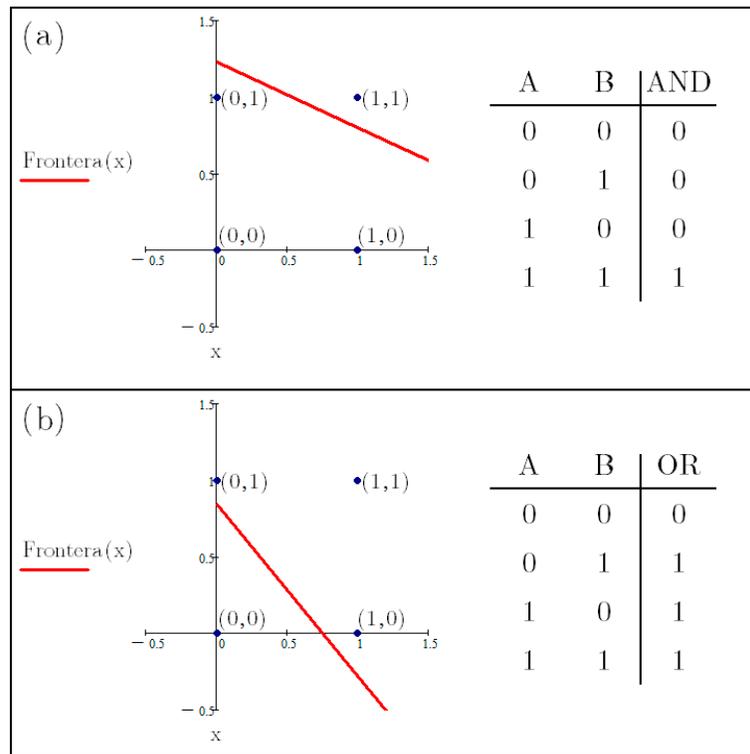


Figura 3.8: Mapas de decisión para las operaciones (a) AND y (b) OR

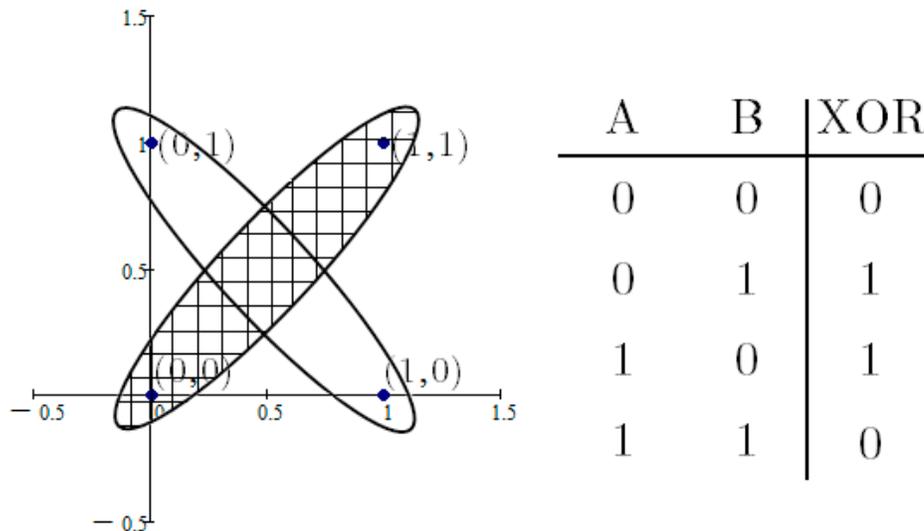


Figura 3.9: En la función XOR no es posible obtener una recta que separe las dos clases.

Para poder elegir entre una zona u otra de las tres, es necesario utilizar otra capa con una neurona cuyas entradas serán las salidas de las neuronas anteriores. Las dos zonas o regiones que contienen a los puntos $(0,0)$ y $(1,1)$ se asocian a una salida 0 de la red, y la zona central se asocia a la salida con valor 1 [9]. De esta manera, es posible encontrar una solución al problema de la función XOR, esta solución es mostrada en la Figura 3.10.

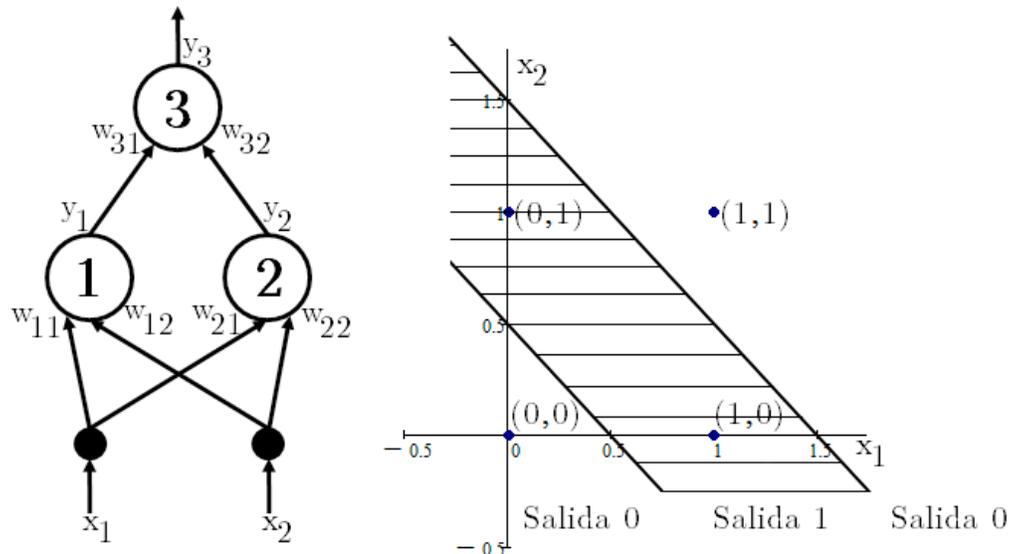


Figura 3.10: Solución del problema de la función XOR.

3.7. La red backpropagation

Este método, conocido en general como *backpropagation* o *propagación del error hacia atrás*, está basado en la generalización de la regla delta y, a pesar de sus propias limitaciones, ha ampliado de forma considerable el rango de aplicaciones de las redes neuronales.

El algoritmo de propagación hacia atrás, o retropropagación, es una regla de aprendizaje que se puede aplicar en modelos de redes con más de dos capas de neuronas. Una característica importante de este algoritmo es la representación interna del conocimiento que es capaz de organizar en la capa intermedia de las neuronas para conseguir cualquier correspondencia entre la entrada y la salida de la red.

De forma simplificada, el funcionamiento de una red *backpropagation* consiste en un aprendizaje de un conjunto predefinido de pares de entradas - salidas dados como ejemplo, emplean-

do un ciclo de propagación - adaptación de dos fases: primero se aplica un patrón de entrada como estímulo para la primera capa de las neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida, se compara el resultado obtenido en las neuronas de salida con la salida deseada y se calcula un valor del error para cada neurona de salida.

A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original.

Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se presente el mismo patrón, la salida esté más cercana a la deseada; es decir, el error disminuya.

La importancia de la red *backpropagation* consiste en su capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones dados como ejemplo y sus salidas correspondientes. Para poder aplicar esa misma relación, después del entrenamiento, a nuevos vectores de entrada con ruido o incompletas, dando una salida activa si la nueva entrada es parecida a las presentadas durante el aprendizaje.

Esta característica importante, que exige a los sistemas de aprendizaje, es la capacidad de generalización, entendida como la facilidad de dar salidas satisfactorias a entradas que el sistema no ha visto nunca en su fase de entrenamiento. La red debe encontrar una representación interna que le permita generar las salidas deseadas cuando se le dan las entradas de entrenamiento, y que pueda aplicar, además, a entradas no presentadas durante la etapa de aprendizaje para clasificarlas según las características que compartan con los ejemplos de entrenamiento [9].

3.7.1. Regla delta generalizada

La regla propuesta por Widrow en 1960 (*regla delta*) ha sido extendida a redes con capas intermedias (*regla delta generalizada*) con conexiones hacia adelante (*feedforward*) y cuyas neuronas tienen funciones de activación continuas (*lineales o sigmoidales*), dando lugar al algoritmo de retropropagación (*backpropagation*). Estas funciones continuas son no decrecientes y derivables. La función sigmoideal pertenece a este tipo de funciones, a diferencia de la función escalón, ya que esta última no es derivable en el punto en el que se encuentra la discontinuidad.

Este algoritmo utiliza también una función o superficie de error asociada a la red, buscando el estado estable de mínima energía o de mínimo error a través del camino descendente de la superficie del error. Por ello, realimenta el error del sistema para realiza la modificación de los pesos en un valor proporcional al gradiente decreciente de dicha función de error.

El método que sigue la regla delta generalizada para el ajuste de los pesos es que, los pesos se actualicen de forma proporcional a la delta, o diferencia entre la salida deseada y la obtenida ($\delta = \text{salida deseada} - \text{salida obtenida}$). Dada una neurona (unidad U_j) y la salida que produce, y_i , el cambio que se produce en el peso de la conexión que une la salida de dicha neurona con la unidad $U_j(w_{ji})$ para un patrón de aprendizaje p determinado es:

$$\Delta w_{ji}(t+1) = \alpha \delta_{pj} y_{pi} \quad (3.10)$$

En donde el subíndice p se refiere al patrón de aprendizaje concreto y α es la constante o tasa de aprendizaje.

El punto en el que difieren la regla delta generalizada de la regla delta es en que el valor concreto de δ_{pj} . Por otro lado, en las redes multinivel, a diferencia de las redes sin neuronas ocultas, en principio no se puede conocer la salida deseada de las neuronas de las capas ocultas para poder determinar los pesos en función del error cometido. Sin embargo, inicialmente sí podemos conocer la salida deseada de las neuronas de salida. Según esto, si consideramos la unidad U_j de salida, entonces definimos:

$$\delta_{pj} = (d_{pj} - y_{pj}) \cdot f'(net_j) \quad (3.11)$$

donde d_{pj} es la salida deseada de la neurona j para el patrón p y net_j es la entrada neta que recibe la neurona j .

Esta fórmula es como la de la regla delta, excepto en lo que se refiere a la derivada de la función de transferencia. Este término representa la modificación que hay que realizar en la entrada que recibe la neurona j . En el caso de que dicha neurona no sea de salida, el error que se produce estará en función del error que se cometa en la neuronas que reciban como entrada la salida de dicha neurona. Esto es lo que se denomina procedimiento de propagación del error hacia atrás.

Según esto, en el caso de que U_j no sea una neurona de salida, el error que se produce está en función del error que se comete en las neuronas que reciben como entrada la salida de U_j ,

$$\delta_{pj} = \left(\sum_k \delta_{pk} w_{kj} \right) \cdot f'(net_j) \quad (3.12)$$

donde el rango de k cubre todas aquellas neuronas a las que está conectada la salida de U_j . De esta forma, el error que se produce en una neurona oculta es la suma de los errores que se

producen en las neuronas a las que está conectada la salida de ésta, multiplicando cada uno de ellos por el peso de la conexión.

Adición de un momento en la regla delta generalizada

el método de retropropagación del error, también conocido como del gradiente descendente, requiere un importante número de cálculos para lograr el ajuste de los pesos de la red. En la implementación del algoritmo, se toma una amplitud de paso que viene dada por la tasa de aprendizaje α . A mayor tasa de aprendizaje, mayor es la modificación de los pesos en cada iteración, con lo que el aprendizaje será más rápido, pero, por otro lado, puede dar lugar a oscilaciones.

Para filtrar estas oscilaciones, se añade en la expresión del incremento de los pesos un término (momento), β , de manera que dicha expresión quede,

$$\begin{aligned} w_{ji}(t+1) &= w_{ji}(t) + \alpha \delta_{pj} y_{pi} + \beta (w_{ji}(t) - w_{ji}(t-1)) = \\ \Delta w_{ji}(t+1) &= \alpha \delta_{pj} y_{pi} + \beta \Delta w_{ji}(t) \end{aligned} \quad (3.13)$$

donde β es una constante (momento) que determina el efecto en $t+1$ del cambio de los pesos en el instante t .

Con este momento se consigue la convergencia de la red en menor número de iteraciones, ya que si en t el incremento de un peso era positivo y en $t+1$ también, entonces el descenso por la superficie de error en $t+1$ es mayor. Sin embargo, si en t el incremento era positivo y en $t+1$ es negativo, el paso que se da en $t+1$ es más pequeño, lo cual es adecuado, ya que eso significa que se ha pasado por un mínimo y que los pasos deben ser menores para poder alcanzarlo.

3.8. K-Vecinos Próximos (KNN)

El método KNN (K nearest neighbors) es un método de clasificación supervisada que sirve para estimar la función de densidad $F(\frac{x}{C_j})$ de los elementos x por cada clase C_j . 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 muestras.

En el proceso de aprendizaje no se hace ninguna suposición acerca de la distribución de las variables x . La salida depende de si k -nn se utiliza para clasificación o regresión [5]:

- En la clasificación k -nn, la salida es una pertenencia de una clase. Un objeto se clasifica por mayoría de votos de sus k vecinos, el objetivo es asignar a la clase más común entre sus k vecinos más cercanos (k es un número entero positivo, por lo general pequeño). Si $k = 1$, entonces el objeto simplemente se asigna a la clase de los que solo es vecino más cercano.
- En la regresión k -nn, la salida es el valor de la propiedad para el objeto. Este valor es el promedio de los valores de sus k vecinos más cercanos.

En el reconocimiento de patrones, el algoritmo k -nn es usado como método de clasificación de objetos (elementos) basado en un entrenamiento mediante ejemplos cercanos en el espacio de los elementos. K -nn es un tipo de "Aprendizaje Lento", donde la función se aproxima sólo localmente y todo el cómputo es diferido a la clasificación. El algoritmo k -nn está entre el más simple de todos los algoritmos de aprendizaje automáticos.

Tanto para la clasificación y regresión, puede ser útil ponderar las contribuciones de los vecinos, de modo que los vecinos más cercanos contribuyen más a la media que los distantes. Por ejemplo, un esquema de ponderación común consiste en dar a cada vecino un peso de $1/d$, donde d es la distancia del vecino.

Los vecinos se toman de un conjunto de elementos x para los cuales la clase (para clasificación k -nn) o el valor de propiedad del objeto (para regresión k -nn) es conocido. Esto puede estar pensado como el conjunto de entrenamiento para el algoritmo, aunque no se requiera ninguna etapa de formación explícita.

3.8.1. Algoritmo

Los ejemplos de entrenamiento son vectores en un espacio característico multidimensional, cada ejemplo está descrito en términos de p atributos considerando q clases para la clasificación. Los valores de los atributos del i -ésimo ejemplo (donde $1 \leq i \leq n$) se representan por el vector p -dimensional,

$$x_i = (x_{1i}, x_{2i}, \dots, x_{pi}) \in X \quad (3.14)$$

El espacio es particionado en regiones por localizaciones y etiquetas de los ejemplos de entrenamiento. Un punto en el espacio es asignado a la clase C si esta es la clase más frecuente entre los k ejemplos de entrenamiento. Generalmente se usa la distancia euclidiana [6].

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^p (x_{ir} - x_{jr})^2} \quad (3.15)$$

La fase de entrenamiento del algoritmo consiste en almacenar los vectores característicos y las etiquetas de las clases de los ejemplos de entrenamiento. En la fase de clasificación, la evaluación del ejemplo (clase desconocida) es representada por un vector en el espacio característico. Se calcula la distancia entre los vectores almacenados y el nuevo vector, y se seleccionan los k ejemplos más cercanos. El nuevo ejemplo es clasificado con la clase que más se repite en los vectores seleccionados.

Este método supone que los vecinos más cercanos nos dan la mejor clasificación y esto se hace utilizando todos los atributos; el problema de dicha suposición es que es posible que se tengan muchos atributos irrelevantes que dominen sobre la clasificación: dos atributos relevantes perderían peso entre otros veinte irrelevantes.

Para corregir el posible sesgo se puede asignar un peso a las distancias de cada atributo, dándole así mayor importancia a los atributos más relevantes. Otra posibilidad consiste en tratar de determinar o ajustar los pesos con ejemplos conocidos de entrenamiento. Finalmente, antes de asignar pesos es recomendable identificar y eliminar los atributos que se consideran no necesarios. En síntesis, el método k -nn se resume en dos algoritmos:

Algoritmo de entrenamiento

Para cada ejemplo $\langle x, f(x) \rangle$, donde $x \in X$, agregar el ejemplo a la estructura representando los ejemplos de aprendizaje.

Algoritmo de clasificación

Dado un ejemplar x_q que debe ser clasificado, sean x_1, \dots, x_k los k vecinos más cercanos a x_q en los ejemplos de aprendizaje, regresar

$$\hat{f}(x) \leftarrow \operatorname{argmax}_{v \in V} \sum_{i=1}^k \delta(v, f(x_i))$$

donde $\delta(a, b) = 1$ si $a = b$; y 0 en cualquier otro caso.

el valor $\hat{f}(x_q)$ devuelto por el algoritmo como un estimador de $f(x_q)$ es solo el valor más común de f entre los k vecinos más cercanos a x_q . Si elegimos $k = 1$; entonces el vecino más cercano a x_i determina su valor.

3.8.2. Elección del k

La mejor elección del k depende fundamentalmente de los datos; generalmente, valores grandes de k reducen el efecto de ruido sobre la clasificación, pero crean límites entre clases parecidas. Un buen k puede ser seleccionado mediante una optimización de uso. El caso especial en que la clase es predicha para ser la clase más cercana al ejemplo de entrenamiento (cuando $k = 1$) es llamado **Nearest Neighbor Algorithm**, Algoritmo del Vecino más Cercano.

La exactitud de este algoritmo puede ser severamente degradada por la presencia de ruido o características irrelevantes, o si las escalas de características no son consistentes con lo que uno considera importante. Muchas investigaciones y esfuerzos fueron puestos en la selección y crecimiento de características para mejorar las clasificaciones. Particularmente una aproximación en el uso de algoritmos que evolucionan para optimizar características de escalabilidad.

En problemas binarios (de dos clases) de clasificación, resulta útil elegir un k impar para evitar empates. Una forma popular para la elección del k óptimo de forma empírica es mediante el método de arranque. A menudo se utiliza una matriz de confusión como herramienta para validar la exactitud de la clasificación [7].

En la Figura 3.11 se muestra un ejemplo en el cual se desea clasificar el círculo verde. Entonces, para un $k=3$, éste elemento será clasificado como la clase triángulo ya que solo hay un cuadrado y dos triángulos dentro del círculo que los contiene. Si, por otro lado se tiene un $k=5$, éste sería clasificado con la clase cuadrado, ya que hay dos triángulos y tres cuadrados dentro del círculo exterior.

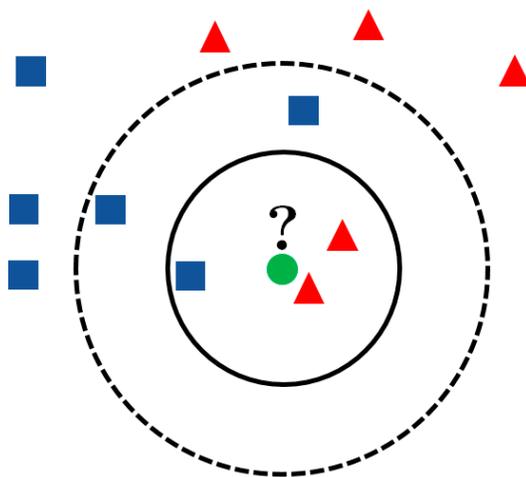


Figura 3.11: Algoritmo K -nn para la clasificación de las clases círculo, cuadrado y triángulo.

3.9. Conclusiones

Se analizaron diferentes técnicas de extracción de características basadas en: a) funciones momento de una imagen digital y b) submuestreo por regiones.

Los descriptores o momentos de una imagen, nos dieron información tanto global como local de la distribución de intensidades de una imagen. Como se puede apreciar en la Figura 3.1, los descriptores son invariantes al cambio en la posición, orientación y escala de una imagen.

Los descriptores de venas son vectores binarios característicos generados a partir de la cantidad de información relevante que hay en una región dada.

Los descriptores obtenidos alimentan una red neuronal (perceptrón multicapa) la cual usa un algoritmo de aprendizaje supervisado para entrenamiento conocido como propagación hacia atrás (Back Propagation).

Además se realiza un estudio comparativo usando un clasificador lineal llamado k vecinos próximos (k nearest neighbors). Como se observa en la Figura 3.11 los resultados de clasificación dependerán del número de k vecinos a considerar.

Bibliografía

- [1] Bahram Javidi, "Image Recognition and Classification: Algorithms, Systems, and Applications," *CRC Press*, ISBN: 9780203910962, 2002.
- [2] Rafael C. Gonzalez and Richard E. Woods. "Digital Image Processing", *Prentice Hall*, 2da Edición, (2002).
- [3] Kevin Gurney, "An introduction to neural networks," *CRC Press*, ISBN: 9780203451519, 2003.
- [4] José Ramón Hilerá González y Víctor José Martínez Hernando, "Redes neuronales artificiales: fundamentos, modelos y aplicaciones," *RA-MA*, 1994.
- [5] N. S. Altman, "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression ," *The American Statistician*, vol. 46, pp. 175-185, 1992.
- [6] Pablo A. Jaskowiak and Ricardo J. G. B. Campello, "Comparing Correlation Coefficients as Dissimilarity Measures for Cancer Classification in Gene Expression Data," *Brazilian Symposium on Bioinformatics (BSB 2011)*, pp. 1-8, 2014.
- [7] Peter Hall, Byeong U. Park and Richards J. Samworth, "Choice of Neighbor Order in Nearest-Neighbor Classification," *The Annals of Statistics*, vol. 36, pp. 2135-2152, 2008.

Capítulo 4

Sistema de adquisición de imágenes

4.1. Introducción

La visión por computadora es un campo de investigación que se ha desarrollado ampliamente en los últimos años. Su objetivo principal es el análisis y descripción de una escena a partir de una imagen digital.

El rango de aplicaciones en las que la visión por computadora está, son las áreas: 1) Militares, como en la detección y seguimiento de objetivos y el análisis de terreno; 2) Robótica, como el guiado y navegación de robots móviles e industriales; 3) Agricultura, en el análisis de los cultivos, como el crecimiento y detección de enfermedades; 4) Identificación, en el reconocimiento automático a través de las huellas, caras o iris; 5) Control de tráfico, como en la identificación de las matrículas de los vehículos o el control de tráfico vial; 6) Control de calidad, en la verificación e inspección de etiquetas, conectores, motores, circuitos, madera, tela, cristales o comida; 7) Biomedicina, en el análisis de imágenes tomadas por ultrasonidos, rayos X, sangre, ADN y 8) Seguridad, en la vigilancia de edificios, detección de explosivos por rayos X, etc.

Los elementos necesarios para un sistema de visión incluyen al hardware para la adquisición y almacenamiento de las imágenes capturadas y el software para procesar e interpretar las imágenes. El avance de la tecnología y la reducción de costos en los componentes, permite tener aplicaciones específicas y de bajo costo, dando solución a diversos problemas. Por lo tanto, para la visualización y adquisición de imágenes de la distribución del patrón de venas de la palma y dorso de la mano, se emplean sistemas de visión por computadora.

Los sistemas de identificación de personas basados en imágenes de la mano utilizan la distribución de venas de la palma o dorso de la mano como información útil para la identificación personal. El patrón de venas, es un rasgo fisiológico útil como indicador biométrico debido a que satisface las propiedades de: Universalidad, Unicidad, Permanencia y Medibilidad.

De todo el espectro electromagnético, el ojo de un ser humano únicamente puede detectar un pequeño segmento de éste, comprendido en el rango de los 400 *nm* a los 700 *nm* aproximadamente. Este rango es conocido comúnmente como rango del espectro visible, sin embargo, existen otros rangos que generalmente contienen mucha más información útil [1].



Figura 4.1: Imágenes de las características que afectan la visibilidad de las venas en las manos: (a) Pliegues de la piel de la palma, (b) Vello del dorso, (c) Cicatrices, (d) Lesiones, lunares o verrugas, (e) Pigmentación de la piel y (f) Falta de visibilidad a simple vista.

A manera de ejemplo, en el rango visible, la estructura de las venas del dorso o palma, es poco perceptible a simple vista, esta visibilidad dependerá de factores como la edad, la temperatura, los niveles de grasa subcutánea, la actividad física, la posición de la mano, el

medio ambiente, etc. Algunas otras características que afectan la visibilidad de las venas son los lunares, las cicatrices, los vellos del dorso, la pigmentación de la piel, algún tipo de verruga y los pliegues de la piel, algunas de estas características pueden oscurecer la imagen o añadir ruido [1]. En la Figura 4.1, se muestran imágenes de la palma y dorso de la mano adquiridas en el visible.

Con el uso de imágenes adquiridas en el infrarrojo se puede ver un contraste entre los vasos sanguíneos y el entorno, además de eliminar ciertos rasgos indeseables de la superficie de la piel. El uso de luz en el infrarrojo cercano para la detección del patrón de la distribución de venas, se basa en el principio de absorción de la desoxihemoglobina como respuesta a la exposición infrarroja [2]. En la Figura 4.2, se muestran imágenes de la palma, dorso y muñeca de la mano adquiridas en el infrarrojo cercano.

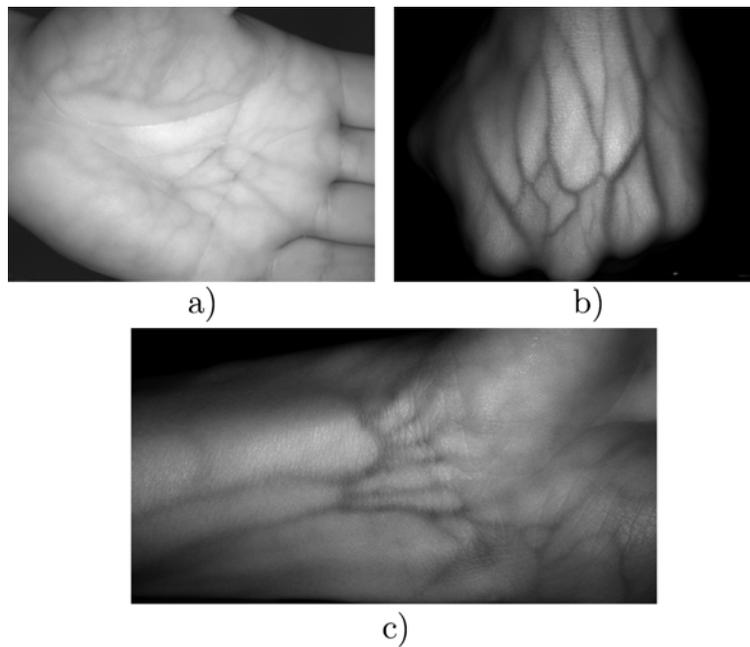


Figura 4.2: Imágenes en el infrarrojo cercano de: (a) Palma, (b) Dorso y (c) Muñeca de la mano.

En el sistema circulatorio humano, las venas son los vasos sanguíneos que transportan la sangre desoxigenada (pobre en oxígeno) de los tejidos hacia el corazón [3]. Los principales absorbentes del infrarrojo en la sangre son la oxihemoglobina, la desoxihemoglobina y el agua. Las dos primeras sustancias presentan casi las mismas características de absorción hasta los 600 nm . Para el rango de entre los 700 nm a los 950 nm , la oxihemoglobina, la desoxihemoglobina y el agua, presentan un mayor nivel de absorción. Esto se aprecia en la Figura 4.3. A partir de los 950 nm , la curva de la desoxihemoglobina cae rápidamente [2].

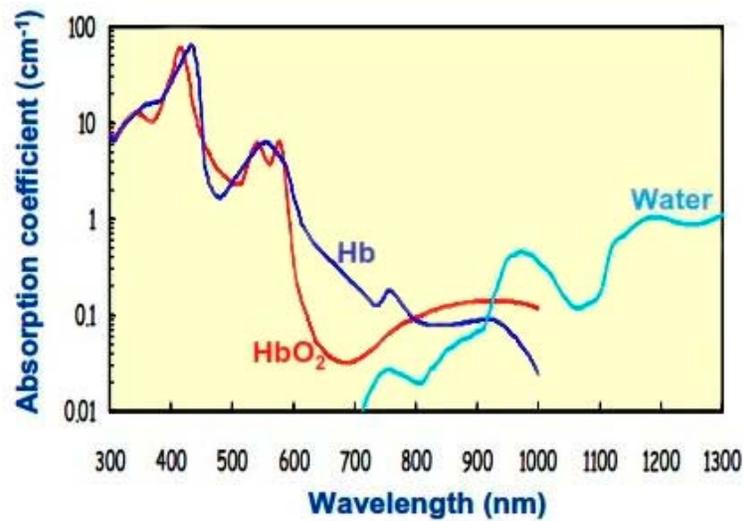


Figura 4.3: Niveles de absorción de la oxihemoglobina, desoxihemoglobina y el agua.

Por otro lado, la luz llega a diferentes profundidades de los tejidos según la longitud de onda que se esté utilizando como se observa en la Figura 4.4. Entre los 300 *nm* y 400 *nm* la luz sólo llega a la zona superficial de la epidermis y dermis de la piel en donde no hay venas. A partir de los 700 *nm* a los 900 *nm*, la luz penetra profundamente en los tejidos llegando incluso a la sangre de los vasos sanguíneos ubicados en el tejido subcutáneo [2].

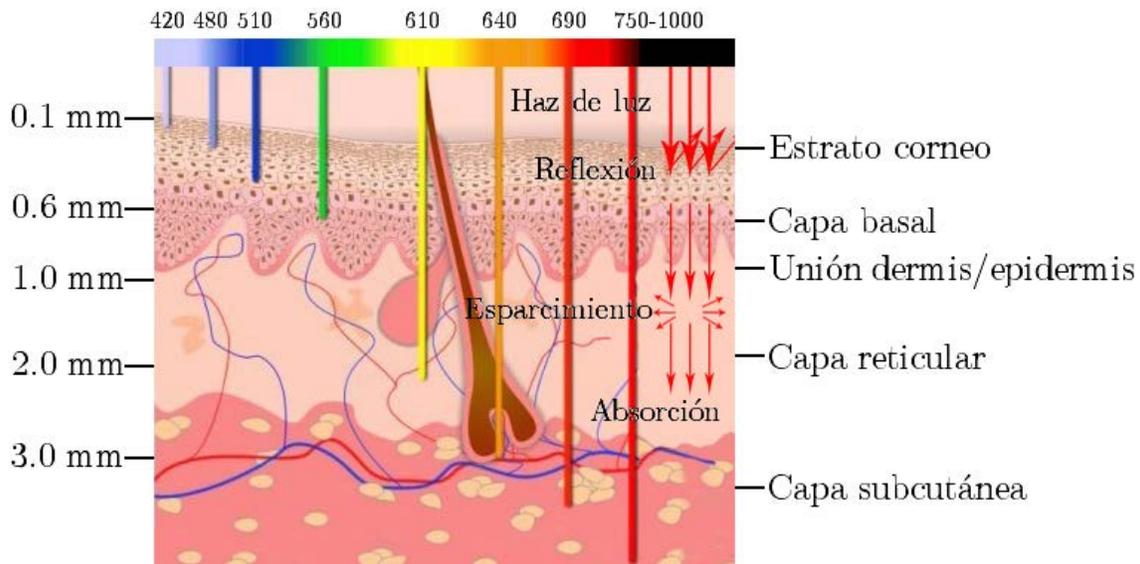


Figura 4.4: Profundidad de penetración de la luz en la piel a diferentes longitudes de onda.

El resultado de utilizar luz infrarroja incidente será absorbida en mayor cantidad por las venas y reflejada por el tejido circundante, por lo tanto, el resultado es una imagen en donde el patrón de venas aparece más oscuro que el resto del tejido [2].

En la literatura, existen diversas configuraciones para un sistema de adquisición de imágenes del patrón vascular tanto de la palma como del dorso de la mano, éstas adquieren imágenes del patrón de venas tanto por reflexión como por transmisión de la luz incidente del infrarrojo.

Óscar Motato et al. [1] proponen un sistema compuesto por una cámara de video convencional con una respuesta espectral extendida hasta los 1200 nm y un filtro óptico infrarrojo en el que su porcentaje de transmisión es significativo a partir de los 700 nm . El sistema está basado en la configuración tanto por transmisión como por reflexión. En la Figura 4.5 se muestra el prototipo implementado y la imagen adquirida con este sistema.

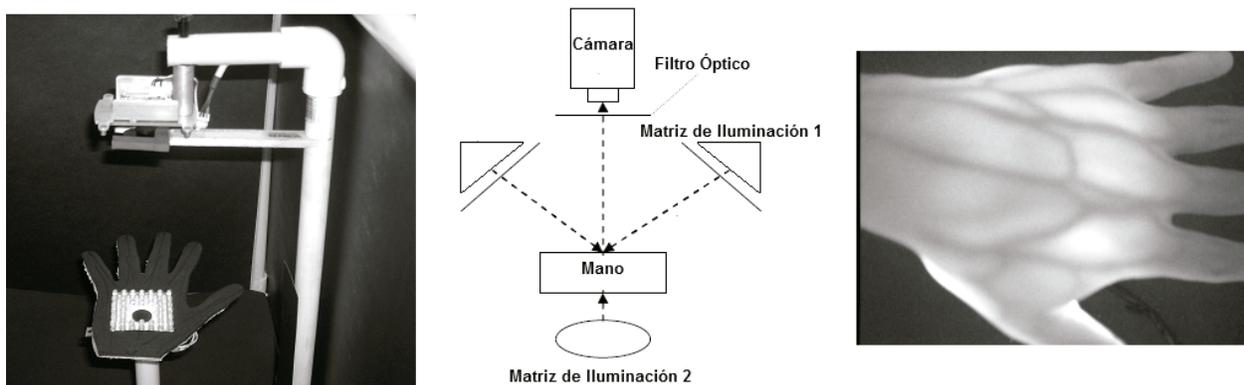


Figura 4.5: Sistema de adquisición de imágenes del dorso de la mano.

La imagen adquirida con este sistema presenta una iluminación uniforme debido a que se cuenta con 3 fuentes de iluminación distribuidas de la siguiente manera: una fuente de iluminación colocada bajo la base donde se coloca la mano y 2 fuentes laterales ubicadas en la parte superior de la base de la mano.

Aída Marcotti et al. [2] proponen un sistema basado en una configuración por reflexión, compuesto por una cámara web Eurocase Eucc 816 Infralight y un anillo de LED's infrarrojo con rangos de emisión de 850 y 940 nanómetros. El sistema propuesto es mostrado en la Figura 4.6. Como se muestra en la imagen, la interfaz adquiere, delimita y segmenta el patrón de venas.

En este capítulo se presenta la descripción e implementación de un sistema de adquisición de imágenes en el infrarrojo cercano del dorso y palma de la mano con una configuración basada en reflexión como la mostrada en la Figura 4.6.

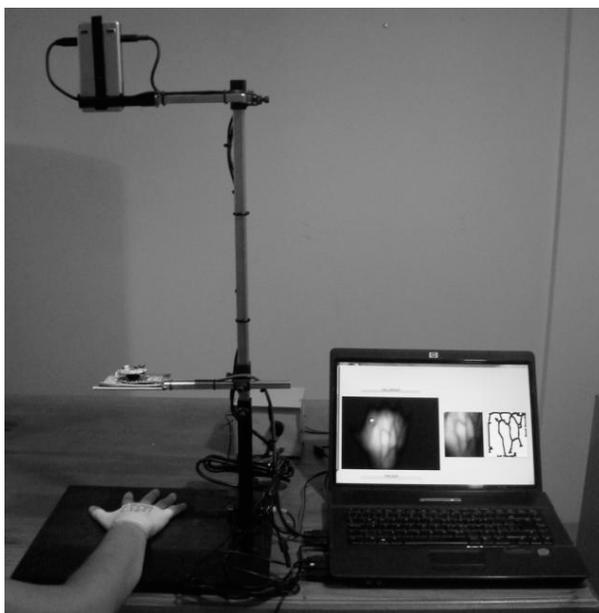


Figura 4.6: Sistema de adquisición de imágenes del dorso de la mano.

Uno de los principales problemas para la creación de nuevos métodos de segmentación y clasificación es la falta de una evaluación comparativa de datos que permita obtener resultados óptimos.

Con el propósito de probar el rendimiento y crear una plataforma común para la comparación del desempeño de los algoritmos propuestos, se procesaron tres bases de datos públicas con fines de investigación llamadas:

- **PolyU Multispectral Palmprint Database** perteneciente a la Universidad Politécnica de Hong Kong.
- **Vein Pattern Database** perteneciente al Instituto de Control e Ingeniería de la Información de la Universidad Tecnológica de Polonia.
- **CASIA Multi-Spectral Palmprint Image Database** perteneciente al Instituto de Automatización de la Academia de Ciencias China.

Los sistemas de adquisición y las características de las imágenes de éstas tres bases de datos son descritas a lo largo del capítulo.

4.2. Sistema para la captura de imágenes en el NIR

Con el objetivo de adquirir una base de datos propia de la Universidad Politécnica de Tulancingo, se implementó un sistema de adquisición de imágenes compuesto por una cámara 2CCD multiespectral de escaneo progresivo, una fuente de luz en el infrarrojo cercano (NIR), una base para colocar la mano y una tarjeta de adquisición conectada a una computadora para la adquisición y procesado de los datos.

Tomando como ventaja el efecto de la absorción de la radiación infrarroja por las venas de la mano, se utiliza una fuente de iluminación en el infrarrojo cercano. Las especificaciones técnicas de esta fuente de luz son:

- Pico de intensidad de radiación emitida de aproximadamente 880 *nm*.
- Distancia óptima de trabajo de 10.16 cm.
- 60 diodos emisores de luz (LED's) distribuidos en dos círculos concéntricos. El círculo externo cuenta con 36 LED's y el interno con 26 LED's.

Una cámara 2CCD Multiespectral de la marca jAi como la que se muestra en la Figura 4.7, es utilizada para la adquisición de las imágenes. Ésta cámara puede capturar simultáneamente los canales visible e infrarrojo cercano a través del mismo camino óptico, produciendo así, una imagen con un tamaño de 1024×768 pixeles.

Cada pixel es un cuadro de tamaño 4.65μ , por lo que el tamaño del sensor es de $4,76mm \times 3,57mm$ tiene una resolución de 800K pixeles así como una conexión GigaE de alta velocidad, esto permite una comunicación rápida con la computadora [4].



Figura 4.7: Cámara 2CCD Multiespectral jAi de escaneo progresivo.

Ésta cámara está basada en un prisma dicróico permitiendo la separación precisa del canal visible (color) y del infrarrojo cercano. El rango de longitud de onda para el canal visible (canal 1) va desde los 400 nm a los 650 nm, y para el infrarrojo cercano (canal 2) desde los 760 nm a 1000 nm, esto se puede apreciar en la Figura 4.8. Ambos canales pueden ser configurados para operar de manera separada o sincronizada [4].

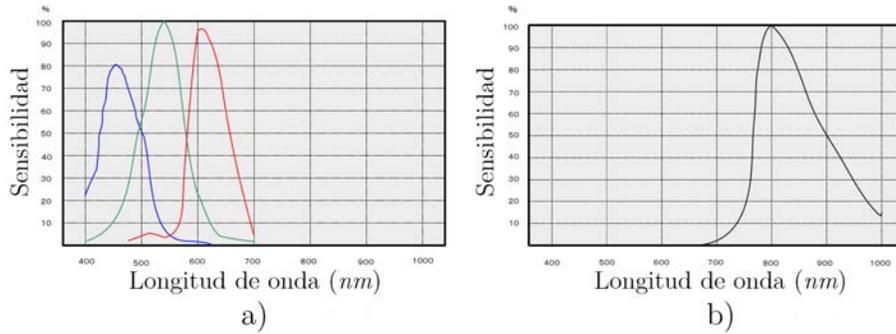


Figura 4.8: Respuesta espectral total incluyendo prisma y sensor para los canales (a) color y (b) infrarrojo.

La palabra dicróico proviene del griego “dikhroos” que significa “dos colores”. Esto hace referencia a los dispositivos ópticos capaces de dividir un haz de luz en dos o más haces con diferentes longitudes de onda.

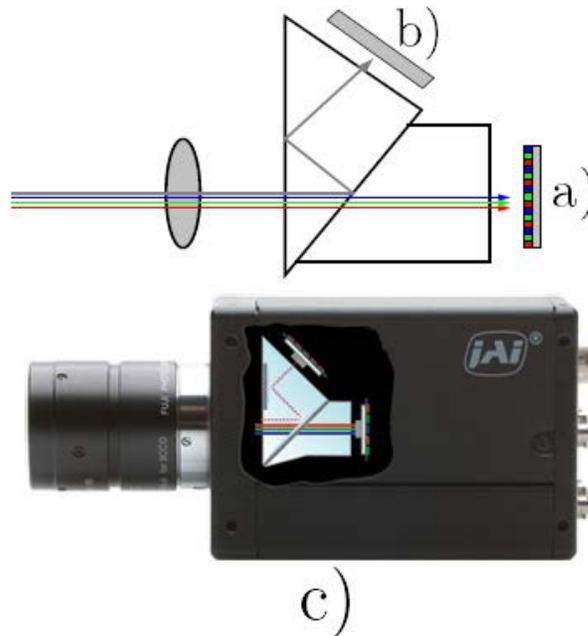


Figura 4.9: Diagrama conceptual del prisma dicróico que separa dos canales (a) Visible, (b) Infrarrojo y (c) Ubicación del prisma en el interior de la cámara.

Un ejemplo de este tipo de dispositivos es el prisma dicróico, mencionado anteriormente y representado de forma conceptual en la Figura 4.9. Las cámaras que incorporan 2CCD y 3CCD con prisma óptico requieren de unas ópticas diseñadas especialmente para corregir la aberración cromática [5]. Las características que posee este tipo de óptica son:

- Distancia focal: 25mm.
- Número f: 1.8 a 16.
- Distancia mínima: 200mm.
- Diámetro máximo: 30mm.
- Longitud: 47mm
- Campo de Visión: 16.36° .

En las Figuras 4.10 y 4.11, se observan los sistemas de adquisición de imágenes de la palma y dorso de la mano, respectivamente, implementados en el Laboratorio de Óptica y Sistemas de Visión de la Universidad Politécnica de Tulancingo.

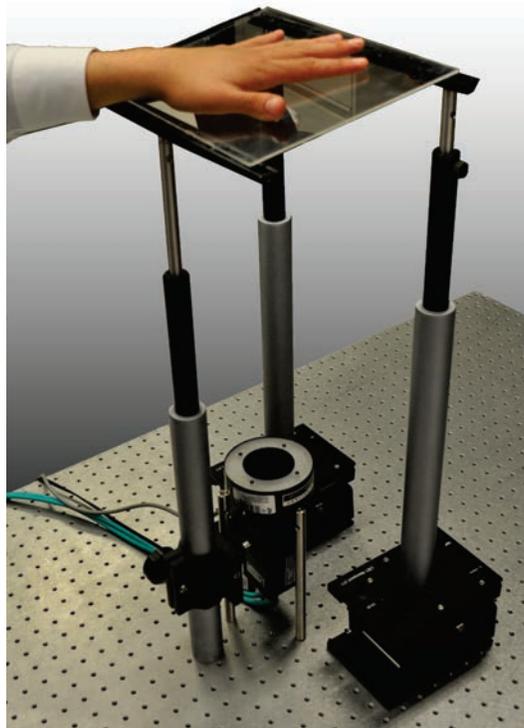


Figura 4.10: Sistema de adquisición de imágenes de la palma de la mano.



Figura 4.11: Sistema de adquisición de imágenes del dorso de la mano.

4.2.1. Interfaz para la captura de imágenes

Se programó una interfaz con el objetivo de establecer una comunicación dinámica entre la computadora y la cámara 2CCD Multispectral jAi para la adquisición y procesamiento de imágenes adquiridas en el infrarrojo cercano.

MATLAB, conocido también como *MATrix LABoratory* o "Laboratorio de Matrices", es una herramienta de software matemático la cual que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio [6], sobre el cual se diseñó y desarrolló esta interfaz la cual es capaz de capturar y procesar las imágenes obtenidas a partir de la cámara.

En la Figura 4.12 se muestra el inicio de la interfaz en la que se observan 5 botones de control, los cuales son: "Iniciar", "Capturar", "Preprocesar", "Detener" y "Salir". Cada uno de estos botones de control lleva a cabo funciones específicas las cuales se describirán a continuación en las siguientes imágenes.

Al momento de presionar el botón "Iniciar", se envía una señal a la cámara para que

inicie la previsualización del área para la adquisición del patrón de venas de la mano de una persona. Esto puede verse en la Figura 4.13, en la cual se tiene la mano de una persona lista para ser capturada.

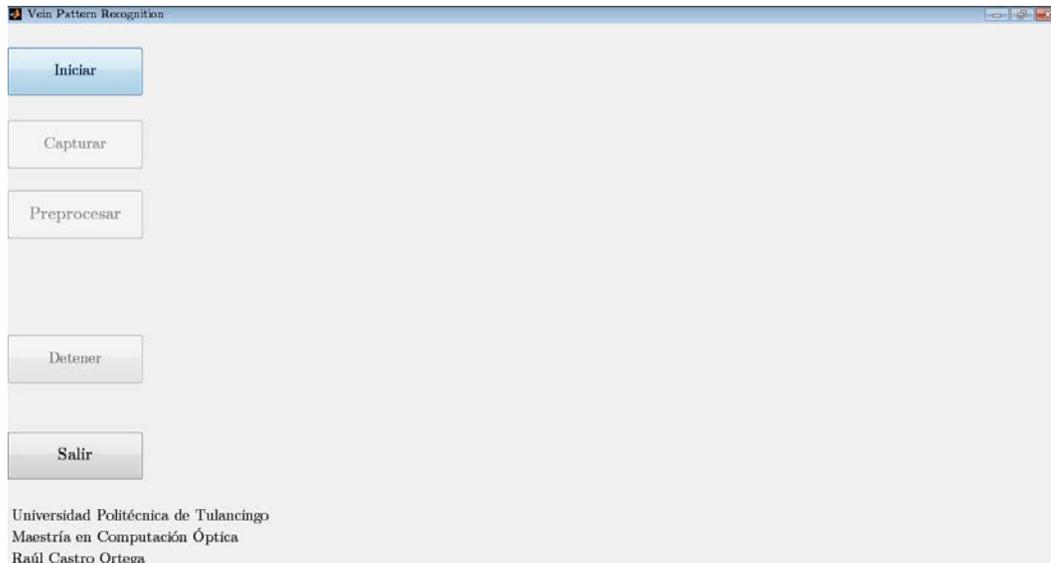


Figura 4.12: Inicio de la Interfaz para la adquisición y procesamiento de las imágenes.



Figura 4.13: Previsualización del área para adquisición del patrón de venas.

Una vez colocada la mano en la posición adecuada para realizar la captura, se presiona el botón **"Capturar"** con el propósito de adquirir o tomar una fotografía de esta, con esta

acción, en la Figura 4.14 se observa que la imagen ha sido capturada y guardada en la memoria de la computadora en una ruta específica.

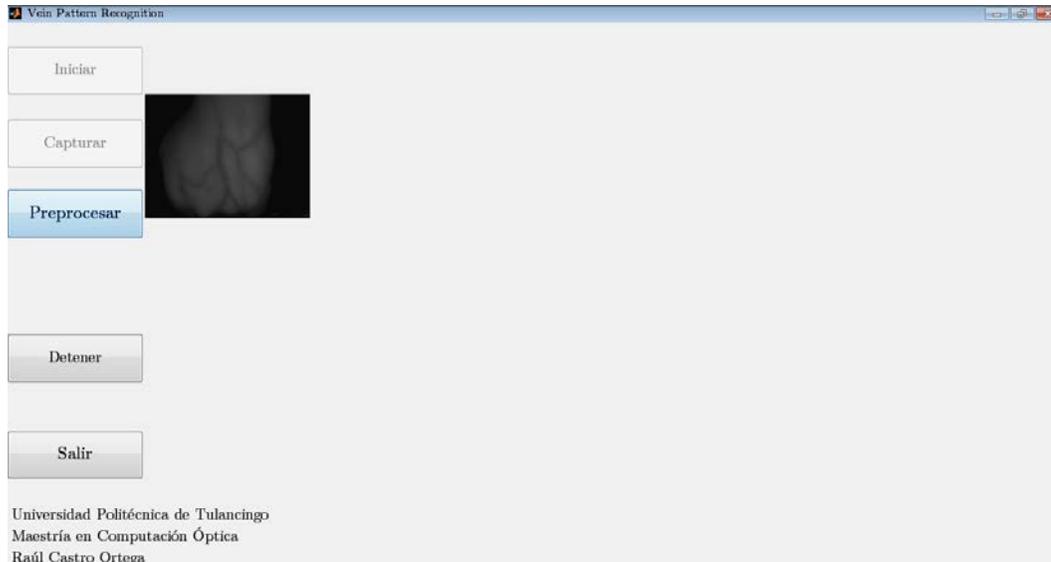


Figura 4.14: Imagen capturada de la distribución del patrón de venas de una persona.

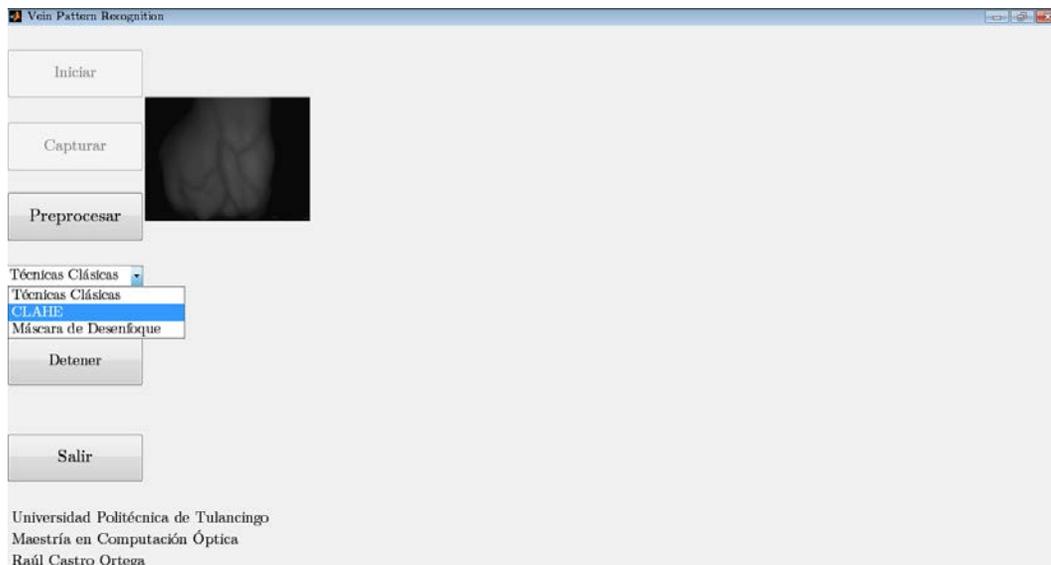


Figura 4.15: Visualización de las técnicas disponibles de preprocesamiento para la imagen capturada.

Acto seguido, al presionar el botón “**Preprocesar**”, se muestran los métodos de preprocesamiento digital de imágenes que pueden ser aplicados a la imagen previamente capturada.

Los métodos programados son las Técnicas clásicas de mejoramiento de imágenes, el Contraste Limitado de la Ecuación Adaptativa del Histograma (CLAHE) y la Máscara de desenfoque (Unsharp Masking). Lo anterior se visualiza en la Figura 4.15.

Al seleccionar el ítem “*Técnicas Clásicas*”, se aplica el preprocesamiento a la imagen adquirida con el objetivo de observar el proceso de *Delimitación* de la Region de Interés, *Mejoramiento* y *Segmentación* del patrón de venas como lo muestra la Figura 4.16.

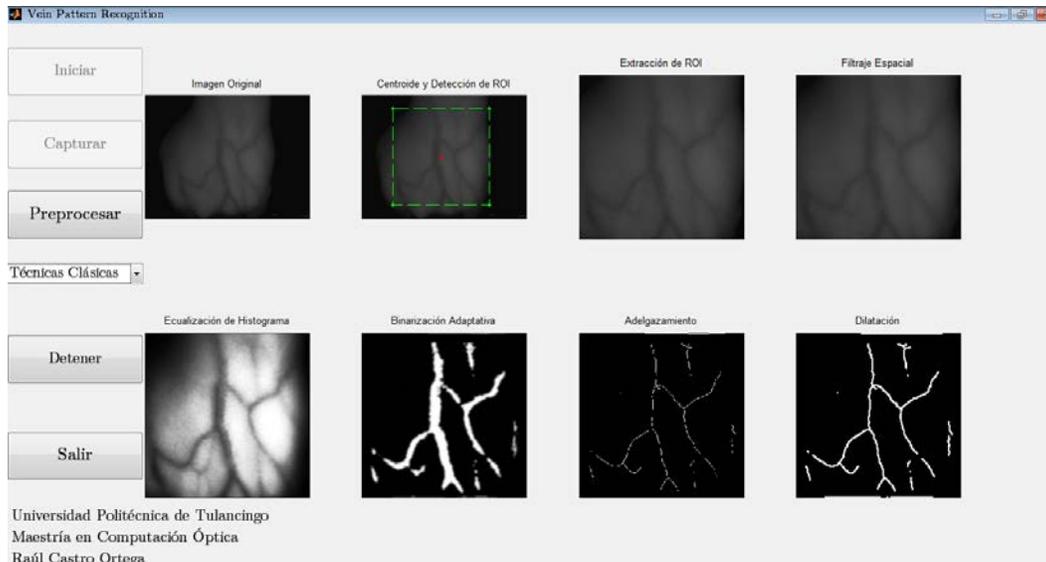


Figura 4.16: **Proceso de Delimitación, Mejoramiento y Segmentación de una imagen utilizando las Técnicas Clásicas.**

De igual manera, al seleccionar el ítem “*CLAHE*”, se aplica el preprocesamiento a la imagen adquirida con el objetivo de observar nuevamente el proceso de *Delimitación* de la Region de Interés, *Mejoramiento* y *Segmentación* del patrón de venas como se puede apreciar en la Figura 4.17.

De la misma forma, al seleccionar el ítem “*Máscara de Desenfoque*”, se aplica el preprocesamiento a dicha imagen adquirida con el propósito de observar el proceso de *Delimitación* de la Region de Interés, *Mejoramiento* y *Segmentación* del patrón de venas como se puede ver en la Figura 4.18.

Las acciones de los botones “**Detener**” y “**Salir**” son básicamente, limpiar el contenido de la ventana así como de los valores de las variables utilizadas regresando al estado de inicial de la aplicación como en la Figura 4.12 y terminar la aplicación, respectivamente. Como se puede observar en las Figuras 4.16, 4.17 y 4.18, la última imagen arroja el patrón de venas segmentado a partir de la imagen original.

A partir de la imagen del patrón de venas, es posible proyectarlo sobre la piel de la mano de una persona con propósitos de venopunción. O bien, compararlo con patrones de venas de una base de datos para la identificación o autenticación de la identidad de un individuo.

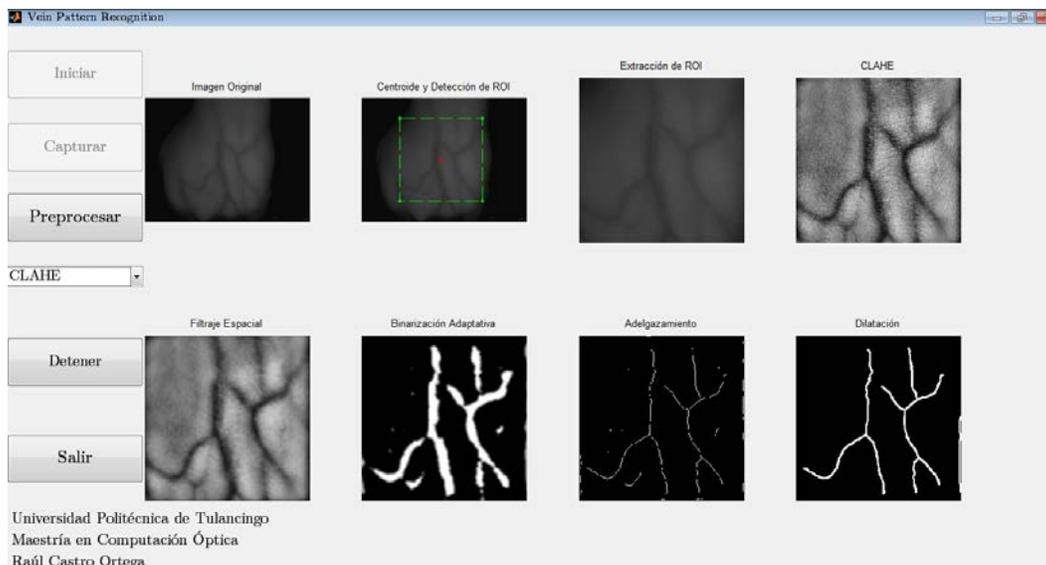


Figura 4.17: Proceso de Delimitación, Mejoramiento y Segmentación de una imagen utilizando la técnica de Contraste Limitado de la Ecuación Adaptativa del Histograma (CLAHE).

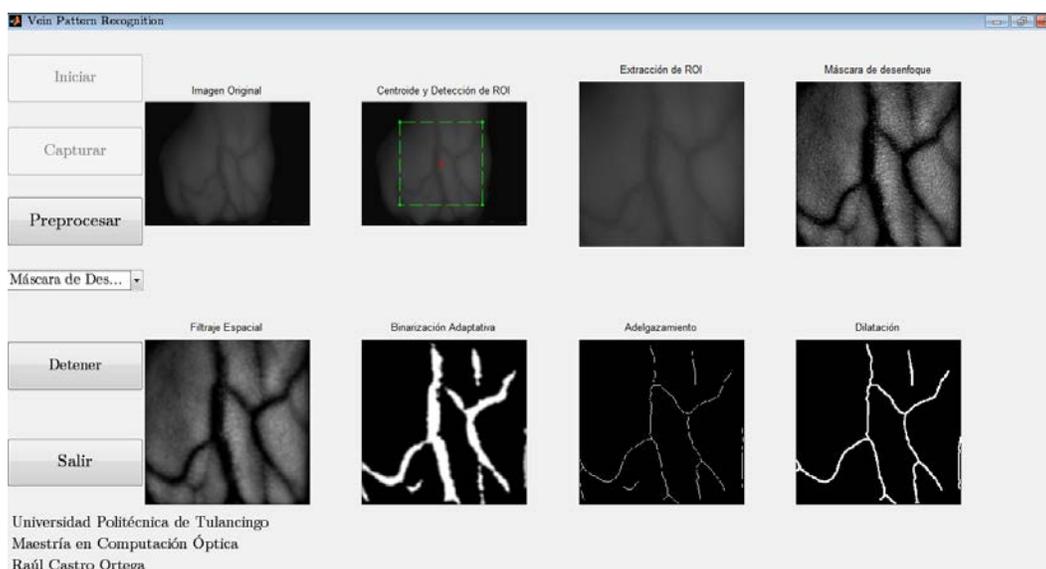


Figura 4.18: Proceso de Delimitación, Mejoramiento y Segmentación de una imagen utilizando la técnica de Máscara de Desenfoque.

4.3. Imágenes adquiridas con el sistema de la UPT

En las Figuras 4.19 y 4.20 se muestran algunas imágenes adquiridas con los sistemas propuestos en las Figuras 4.10 y 4.11, respectivamente, de la palma y dorso de la mano de 4 personas distintas denotadas por (a), (b), (c) y (d). Las imágenes fueron obtenidas a partir de 38 personas, siendo 20 mujeres y 18 hombres. La edad de los participantes es de 20 a 25 años. A cada persona se le tomaron 12 imágenes de ambas manos tanto de la palma como del dorso, por lo tanto, el total de imágenes para cada persona es de 48 y, para la base de datos es de 1824 imágenes.

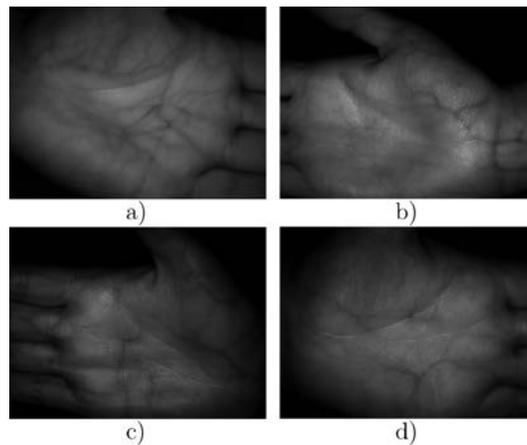


Figura 4.19: Imágenes de la palma de la mano adquiridas con el sistema de la Figura 4.10.

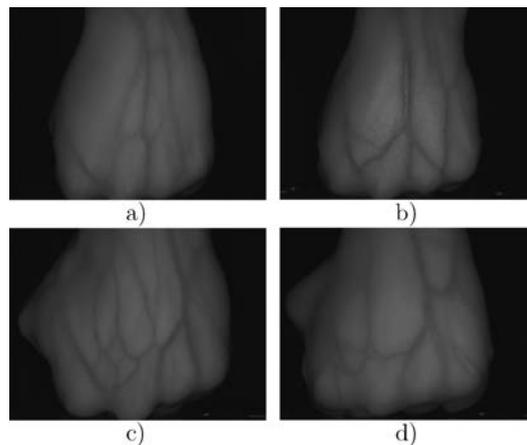


Figura 4.20: Imágenes del dorso de la mano adquiridas con el sistema de la Figura 4.11.

Como se puede apreciar las imágenes adquiridas muestran el patrón de venas de la mano, por lo que se puede ver que, la distribución de venas es diferente para cada persona. En la

siguiente sección se describe el proceso de delimitación de la región de interés, mejoramiento y segmentación de la distribución del patrón vascular.

4.3.1. Base de datos UPT preprocesada

Con el objetivo de delimitar una zona o región de interés (ROI), mejorar y segmentar las imágenes adquiridas con el sistema implementado y, a su vez, hacer una comparación entre éstas técnicas, se lleva a cabo un análisis para determinar qué método es el adecuado para realzar y separar la distribución del patrón de venas del resto de la imagen.

Los métodos de mejoramiento de imágenes implementados son: a) Ecuación del histograma, b) Contraste Limitado de la Ecuación del Histograma (CLAHE) y c) Máscara de desenfoque. Los resultados de aplicar las técnicas clásicas de procesamiento digital de imágenes se muestran en la Figura 4.21, las basadas en CLAHE se ilustran en la Figura 4.22 y las basadas en la Máscara de Desenfoque o Unsharp Masking, en la Figura 4.23.

El procedimiento seguido para la extracción del patrón de venas es el siguiente:

- 1 a,b,c** Dado que, el tamaño de la imagen adquirida es de 1024×768 píxeles, si se preprocesara la imagen completa el costo computacional sería elevado. Por lo tanto, a partir de la imagen original (Figura 4.21 (a)) es posible trazar una zona delimitadora (Figura 4.21 (b)) comúnmente conocida como **región de interés** o **ROI** (Figura 4.21 (c)).
- 2 d** A partir de la imagen delimitada (Figura 4.21 (c)), son aplicadas las técnicas clásicas de preprocesamiento. Se aplica un **filtraje pasa bajas** basado en convolución, con el propósito de eliminar información innecesaria como los vellos o pliegues del dorso y palma de la mano respectivamente. El resultado se muestra en la Figura 4.21 (d). Si ésta información innecesaria no fuese removida, se estaría añadiendo ruido, y por lo tanto, el patrón de venas comenzaría a deformarse. La máscara utilizada para este filtro, fue de tamaño de 5×5 píxeles.
- 3 e** Con el objetivo de redistribuir de manera uniforme los valores de intensidad de los niveles de gris presentes en la imagen se aplicó la **ecuación del histograma**. El resultado, se muestra en la Figura 4.21 (e), la cual es una imagen con el patrón de venas contrastado.
- 4 f** Debido a que los valores de intensidad de los niveles de gris varían en cada imagen y a través de estas, el uso de técnicas de binarización global son deficientes en casos donde la iluminación no es uniforme como se observa en la parte superior izquierda de la Figura 4.21 (e).

Por esta razón, se implementó la técnica de **binarización adaptativa**, la cual es un algoritmo de umbralización que combina el umbral adaptativo global y local, es decir

umbralización local con reducción global. Esta técnica utiliza diferentes valores de umbral, los cuales son asignados a cada pixel de la imagen, esto basado en el análisis de los niveles de gris de sus vecinos.

Los umbrales corresponden al valor promedio de los niveles de gris de una vecindad de $m \times n$ alrededor de cada pixel restado por el umbral de corte global. El resultado de aplicar este algoritmo es ilustrado en la Figura 4.21 (f).

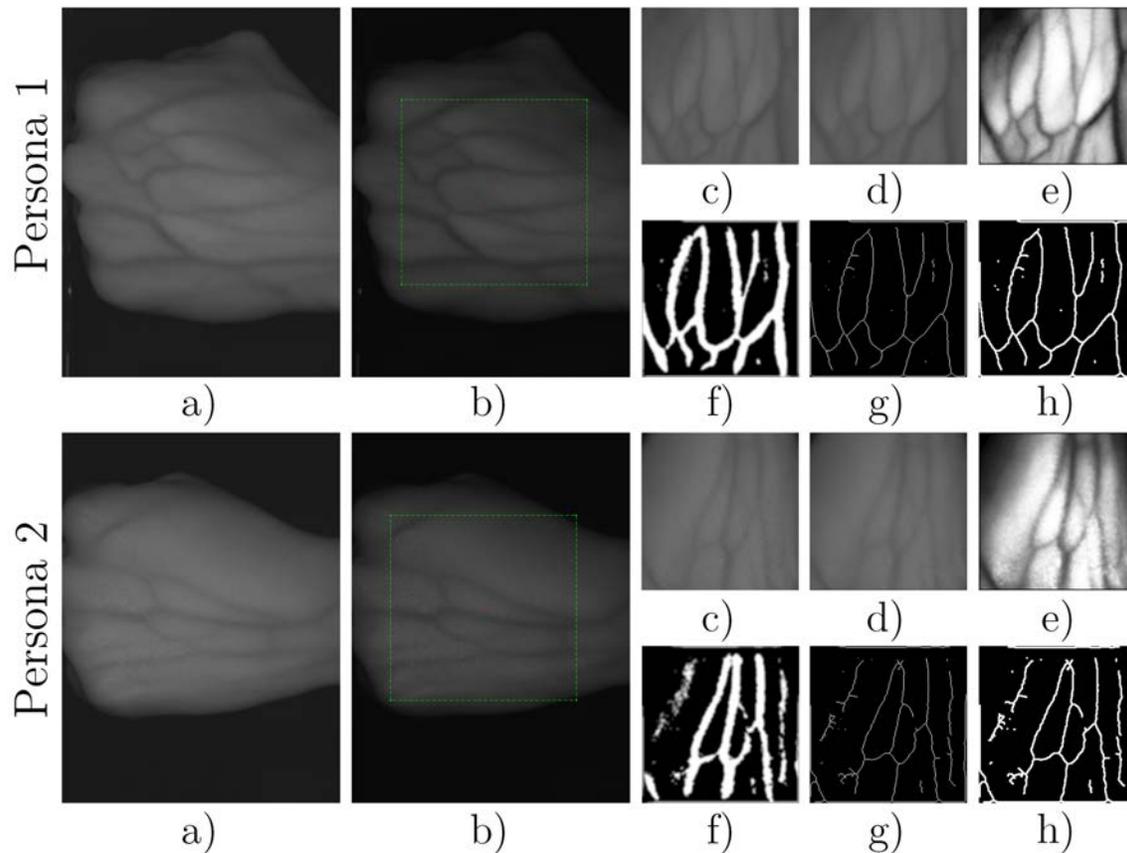


Figura 4.21: (a) Imagen original, (b) Centroide y detección de la región de interés (ROI), (c) Extracción de ROI, (d) Filtrado espacial, (e) Ecuilización del histograma, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación e imagen final con el patrón de venas segmentado.

5 g Una vez realizado esto, fue aplicado un algoritmo de **adelgazamiento** con el fin de obtener únicamente el esqueleto (Figura 4.21 (g)) de la imagen binarizada.

6 h Acto seguido, a la imagen del patrón adelgazado, se le aplicó un algoritmo de **dilatación** (Figura 4.21 (h)). Con este procedimiento se observa que el patrón de venas fue extraído casi por completo.

Como se puede observar, en la Figura 4.21 (h), fue posible extraer el patrón de venas de las personas 1 y 2. Sin embargo, el inconveniente que se tiene al aplicar este procedimiento en el orden que se presentó, es que a partir de la imagen filtrada, se observa un marco negro alrededor de la imagen que continúa presente hasta la dilatación generando ruido y a su vez falsa ubicación de las venas.

Un segundo algoritmo para el mejoramiento de imágenes se muestra en la Figura 4.22. El procedimiento anteriormente descrito fue aplicado en el mismo orden y con los mismos parámetros excepto en 2 (d), 3 (e) y 6 (h).

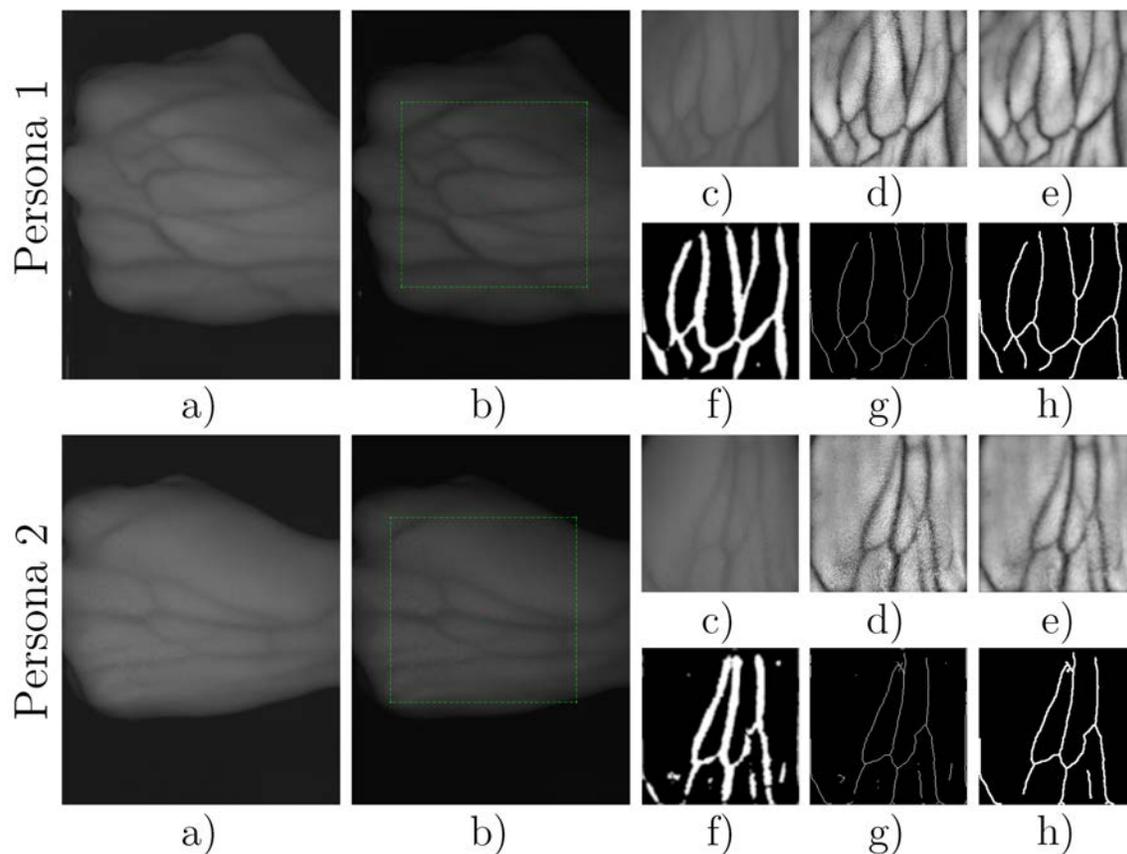


Figura 4.22: (a) Imagen original, (b) Centroide y detección de ROI, (c) Extracción de ROI, (d) Contraste limitado de la ecualización adaptativa del histograma, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

En la Figura 4.22 (d), fue aplicado el Contraste Limitado de la Ecuación Adaptativa del Histograma (CLAHE) en el que se observa una mejora visible en el contraste de la imagen. Además, la presencia de sombra en la parte superior izquierda a sido aclarada.

Este efecto de aclarado es debido a que el algoritmo utiliza distintos histogramas de la imagen por regiones, por lo tanto, la imagen se verá más contrastada y sus valores de intensidad de niveles de gris serán más uniformes.

Una vez aplicado CLAHE, se realizó un filtrado espacial con el objetivo de reducir la presencia de los pliegues y vellos de la cara dorsal de la mano (Figura 4.22 (e)), pero manteniendo la información relevante de las venas. Posteriormente después de aplicar binarización adaptativa (Figura 4.22 (f)), Adelgazamiento (Figura 4.22 (g)) y dilatación (Figura 4.22 (h)), en esta última, se eliminó el ruido presente que pudiera llegar a tener la imagen mediante un conteo de píxeles.

Los resultados de mejoramiento de imágenes basados en Máscara de desenfoque se muestran en la Figura 4.23. El procedimiento se aplicó en el mismo orden y con los mismos parámetros excepto en 2 (d), 3 (e) y 6 (h).

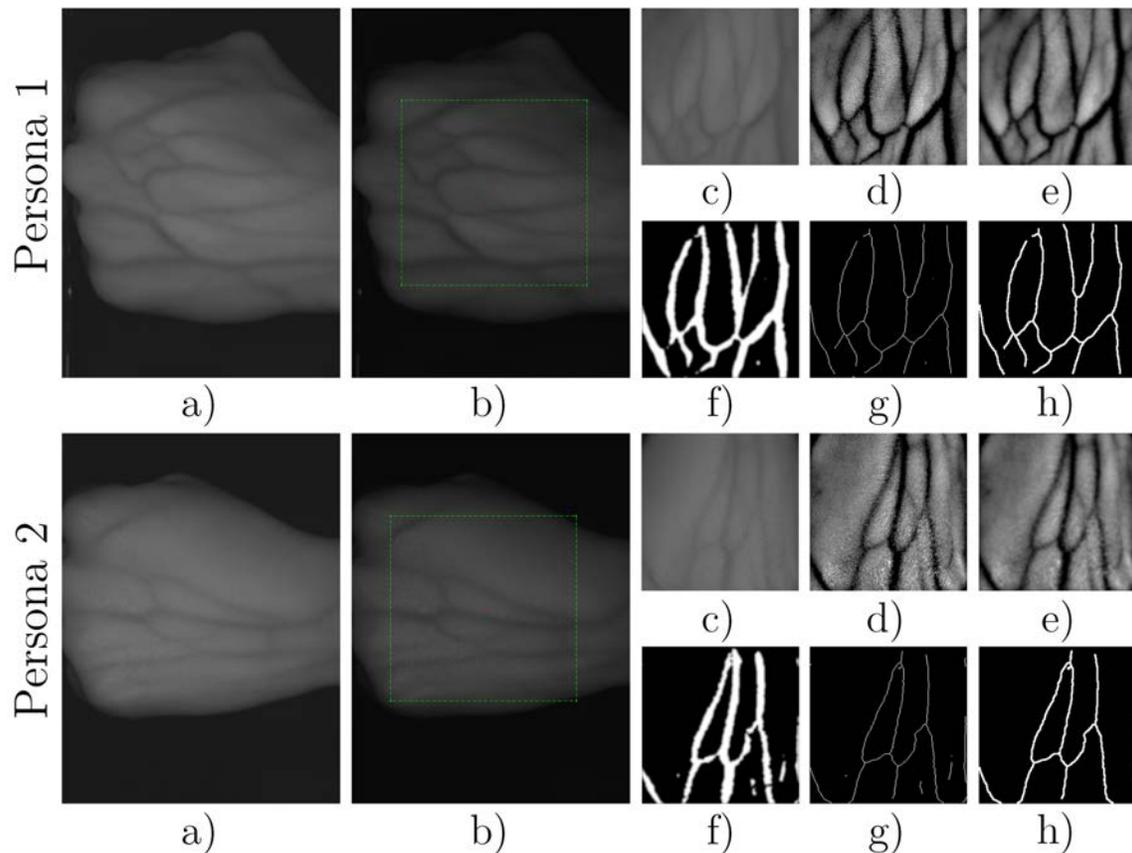


Figura 4.23: (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Máscara de desenfoque, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

En la Figura 4.23 (d), fue aplicado el algoritmo de Máscara de Desenfoque o "Unsharp Masking". Aplicando esta técnica es posible observar un aumento considerable en el contraste de las regiones en donde hay presencia de venas. De forma similar al CLAHE, la presencia de sombra en la parte superior izquierda a sido aclarada y sus valores de intensidad de niveles de gris serán más uniformes.

Después de aplicar la Máscara de Desenfoque, se realizó un filtrado espacial con el objetivo de reducir la presencia de los pliegues y vellos del dorso de la mano (Figura 4.23 (e)), pero manteniendo la información relevante de las venas. Posteriormente después de aplicar Binarización Adaptativa (Figura 4.23 (f)), Adelgazamiento (Figura 4.23 (g)) y dilatación (Figura 4.23 (h)), en esta última, se eliminó el ruido presente que pudiera llegar a tener la imagen mediante un conteo de pixeles.

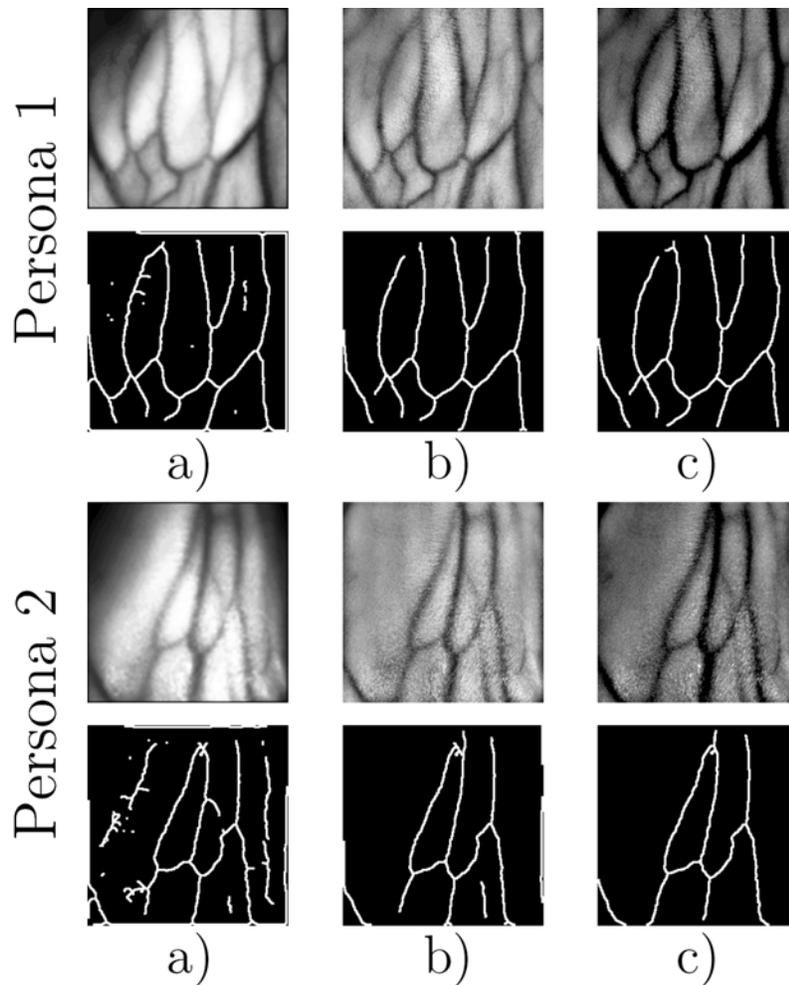


Figura 4.24: Imágenes mejoradas usando: (a) Ecuilización del Histograma, (b) Contraste Limitado de la Ecuilización Adaptativa del Histograma y (c) Máscara de desenfoque.

En la Figura 4.24 se hace una comparación entre los resultados de mejoramiento usando las técnicas basadas en Ecualización del Histograma, Contraste Limitado de la Ecualización del Histograma y Máscara de desenfoque. En la Figura 4.24 (a), a pesar de que se realiza la información del patrón de venas ecualizando el histograma de la imagen, existe la presencia de sombras en la parte superior izquierda para ambos casos, esto puede llegar a distorsionar y afectar el resultado final.

En el caso de CLAHE (Figura 4.24 (b)) y la Máscara de desenfoque (Figura 4.24 (c)), se observa un aumento del contraste así como la eliminación de sombras en la parte superior izquierda en ambas imágenes. Por lo tanto, éstas últimas dos técnicas de mejoramiento de imágenes ayudan a redistribuir los valores de intensidad de los niveles de gris de manera más uniforme haciendo que las imágenes sean más contrastadas.

En ambos casos, para CLAHE (Figura 4.24 (b)) y Máscara de desenfoque (Figura 4.24 (c)) no hay algún tipo de bordes en el marco de la imagen final como en la Figura 4.24 (a) donde se aprecian dichos bordes, lo cual se traduce como ruido o falsa información sobre la presencia de venas.

De lo anterior, es posible concluir que las técnicas: Contraste Limitado de la Ecualización del Histograma y Máscara de desenfoque, son adecuadas para el mejoramiento del contraste y posterior extracción del patrón de venas con las técnicas de segmentación anteriormente descritas. Con el propósito de evaluar la eficiencia de los algoritmos propuestos, se analizaron bases de datos públicas, las cuales serán descritas en la siguiente sección.

4.4. Imágenes adquiridas con el sistema de la UPHK

La base de datos PolyU Multispectral Palmprint Database pertenece al Centro de Investigación Biométrica (UGC/CRC) de la Universidad Politécnica de Hong Kong. En este centro se ha desarrollado un dispositivo como el que se puede observar en la Figura 4.25 para la adquisición de imágenes multiespectral de la palma de la mano en tiempo real.

El sistema puede capturar imágenes bajo fuentes de luz tales como: Azul, Verde, Rojo e Infrarrojo Cercano (NIR), construyéndose así, una base de datos multiespectral de la palma de la mano a gran escala. Las imágenes multiespectrales de la palma de la mano fueron obtenidas a partir de 250 voluntarios, en el que se incluyen a 195 hombres y 55 mujeres. La distribución de la edad de los participantes es de 20 a 60 años [7].

Dichas imágenes fueron adquiridas en dos sesiones por separado. En cada sesión, se le pidió a cada persona proporcionar 6 imágenes para cada palma. Por lo tanto, 24 imágenes en Visible e Infrarrojo de las 2 palmas fueron adquiridas para cada persona. En total, la base de datos

contiene 6000 imágenes de 500 palmas diferentes por cada tipo de iluminación como se puede observar en la Figura 4.26. El intervalo de tiempo promedio entre la primera y la segunda sesión de captura fue de alrededor de 9 días [7].

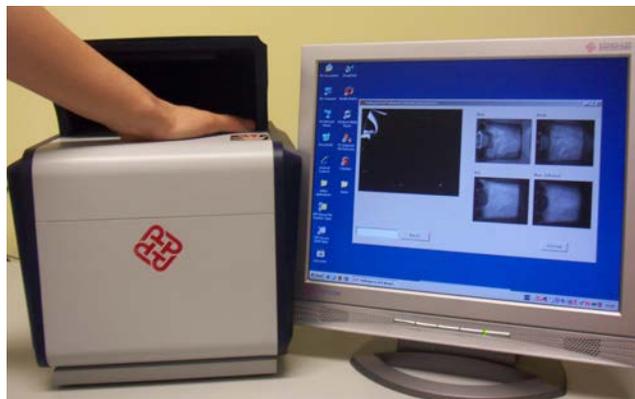


Figura 4.25: Sistema de adquisición de imágenes multispectrales de la palma de la mano [7].

Como se aprecia en la Figura 4.26, las imágenes adquiridas en el canal Rojo (a) muestran información de la distribución de patrón de venas así como de los pliegues de la mano. Tanto el canal Verde (b) y Azul (c), solo muestran la información de los pliegues de la piel, por lo que los canales R, G y B no muestran suficiente información de la distribución de venas. En el caso de las imágenes adquiridas en el infrarrojo, la distribución de venas es visible y los pliegues de la mano son casi nulos. Por ésta razón, estas imágenes fueron utilizadas para este trabajo.

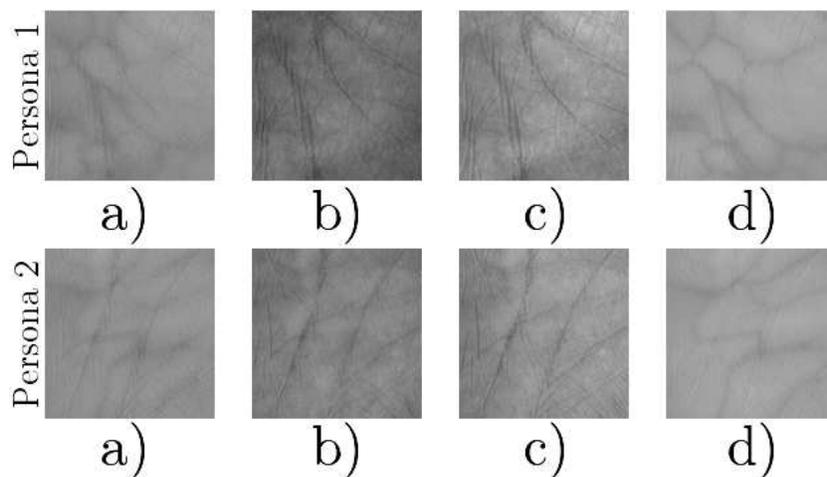


Figura 4.26: Imágenes de la palma de la mano de 2 personas diferentes adquiridas en los canales: (a) Rojo (R), (b) Verde (G), (c) Azul (B) y (d) Infrarrojo cercano (NIR) [7].

4.4.1. Base de datos PolyU Multispectral Palmprint Database pre-procesada

Las imágenes de la base de datos PolyU Multispectral Palmprint cuentan con una delimitación previa de la region de interes (ROI), por lo que ya no fue necesario obtenerla. Los resultados de aplicar las técnicas de mejoramiento de imágenes se muestran en las Figuras 4.27, 4.28 y 4.29. Cabe hacer mención de que el proceso aplicado a estas imágenes es similar al de la sección 4.3.1. Donde los métodos de mejoramiento de imágenes implementados son: a) Ecualización del histograma, b) Contraste Limitado de la Ecualización Adaptativa del Histograma (CLAHE) y c) Máscara de desenfoco.

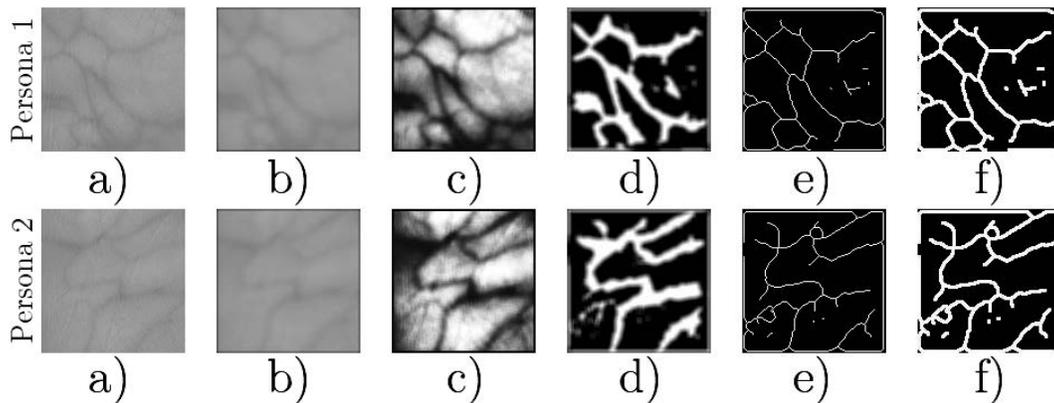


Figura 4.27: (a) Imagen original, (b) Filtraje espacial, (c) Ecualización del histograma, (d) Binarización adaptativa, (e) Adelgazamiento y (f) Dilatación e imagen final con el patrón de venas segmentado.

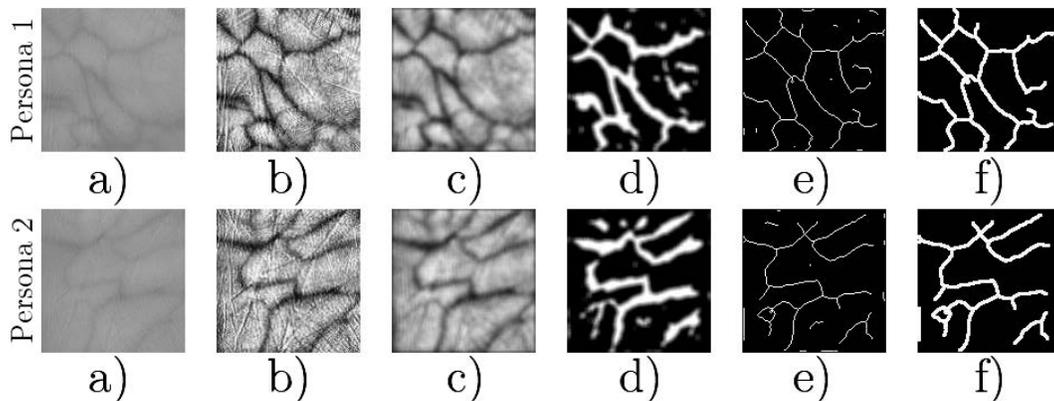


Figura 4.28: (a) Imagen original, (b) Contraste limitado de la ecualización adaptativa del histograma, (c) Filtraje espacial, (d) Binarización adaptativa, (e) Adelgazamiento y (f) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

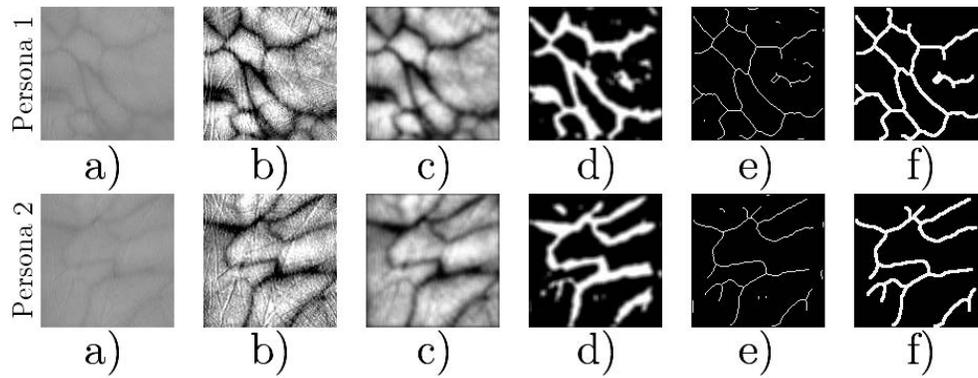


Figura 4.29: (a) Imagen original, (b) Máscara de Desenfoque, (c) Filtrado espacial, (d) Binarización adaptativa, (e) Adelgazamiento y (f) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

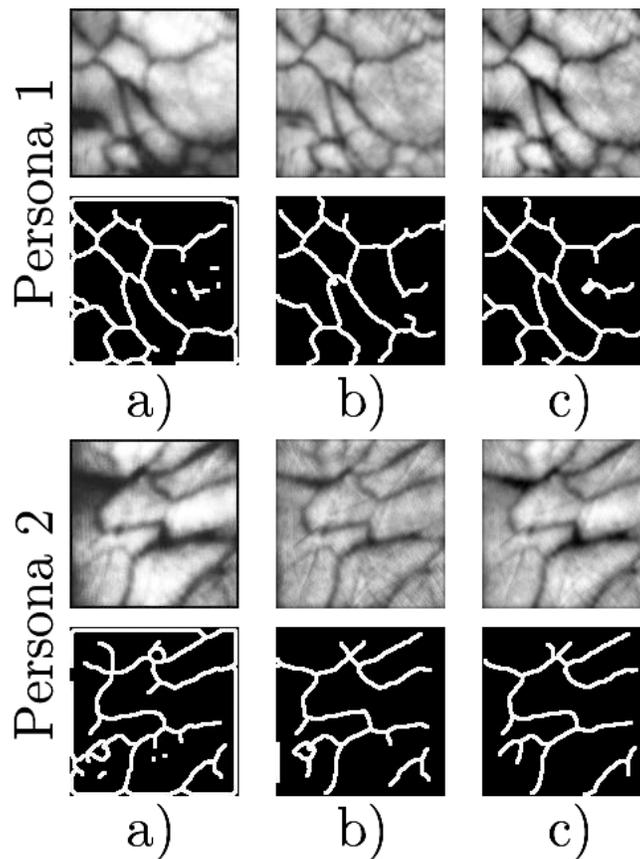


Figura 4.30: Imágenes mejoradas usando: (a) Ecuación del Histograma, (b) Contraste Limitado de la Ecuación Adaptativa del Histograma y (c) Máscara de desenfoque.

En la Figura 4.30 se hace una comparación entre los métodos de mejoramiento de imágenes basados en la Ecualización del Histograma, CLAHE y la Máscara de desenfoque. Como se puede observar, nuevamente las técnicas de mejoramiento basadas en el Contraste Limitado de la Ecualización Adaptativa del Histograma y Máscara de Desenfoque, realzan adecuadamente el contraste de las imágenes IR a pesar de la iluminación no uniforme del sistema de adquisición.

4.5. Imágenes adquiridas con el sistema del CIE de la UTP

Las imágenes adquiridas de la base de datos Vein Pattern Database pertenecen al Instituto de Control e Ingeniería de la Información (CIE) de la Universidad Tecnológica de Polonia, tienen una resolución de 1280×960 píxeles guardadas en formato de mapa de bits (.bmp) de 24 bits y fueron adquiridas en series de 3, con 4 imágenes por cada serie. El intervalo entre cada serie fue de al menos una semana. Para capturar la región de la palma de la mano, los voluntarios colocaron su mano en el dispositivo con el fin de cubrir la ventana de adquisición, de modo que la línea por debajo de sus dedos coincidieran con su borde. Para esta adquisición, no se requirieron sistemas de posicionamiento de la mano adicionales.



Figura 4.31: Dispositivo utilizado para adquirir imágenes de la palma y muñeca de la mano.

En el caso de la muñeca, se utilizó únicamente la construcción en el dispositivo que permite situar la palma y muñeca de forma cómoda al colocar la mano. En la Figura 4.31 se muestra el dispositivo con el que fueron adquiridas las imágenes de esta base de datos. El dispositivo, esta basado en una cámara USB de bajo costo, cuenta con una construcción modular para la adquisición de la palma y la muñeca mediante un sistema de posicionamiento cuasirígido.

Es una base de datos libre disponible con fines de investigación, consta de 2400 imágenes de la distribución de venas. La mitad de la base de datos cuenta con 1200 imágenes del patrón de venas de la palma y, la otra mitad, cuenta con 1200 imágenes del patrón de venas de la

muñeca. Las imágenes fueron adquiridas a partir de las dos manos de 50 estudiantes. Por lo tanto, se tiene un total de 100 patrones diferentes.

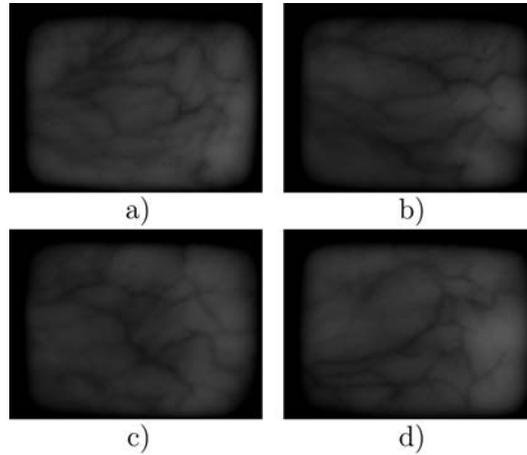


Figura 4.32: Imágenes de la palma de la mano de 4 personas diferentes [8].

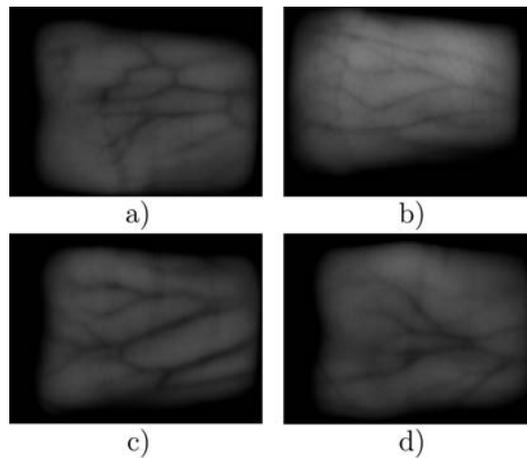


Figura 4.33: Imágenes de la muñeca de la mano de 4 personas diferentes [8].

Estas imágenes encuentran divididas en dos secciones principales: la palma de la mano y la muñeca, cada una cuenta con 50 personas. A su vez, cada carpeta tiene 12 imágenes de la palma derecha y 12 de la palma izquierda dando un total de 24 imágenes para cada persona. En el caso de la región de la muñeca, cada carpeta contiene 12 imágenes de la muñeca derecha y 12 de la muñeca izquierda con un total de 24 imágenes para cada persona [8]. A continuación en las Figuras 4.32 y 4.33 se muestran algunas imágenes adquiridas de la palma y muñeca de la mano de 4 personas distintas con el sistema de la Figura 4.31.

4.5.1. Base de datos Vein Pattern Database preprocesada

Los resultados de aplicar las técnicas de mejoramiento de imágenes se muestran en las Figuras 4.34, 4.35 y 4.36. Cabe mencionar de que el proceso aplicado a estas imágenes es similar al de la sección 4.3.1. Los métodos de mejoramiento de imágenes implementados son: a) Ecualización del histograma, b) Contraste Limitado de la Ecualización Adaptativa del Histograma (CLAHE) y c) Máscara de desenfoque.

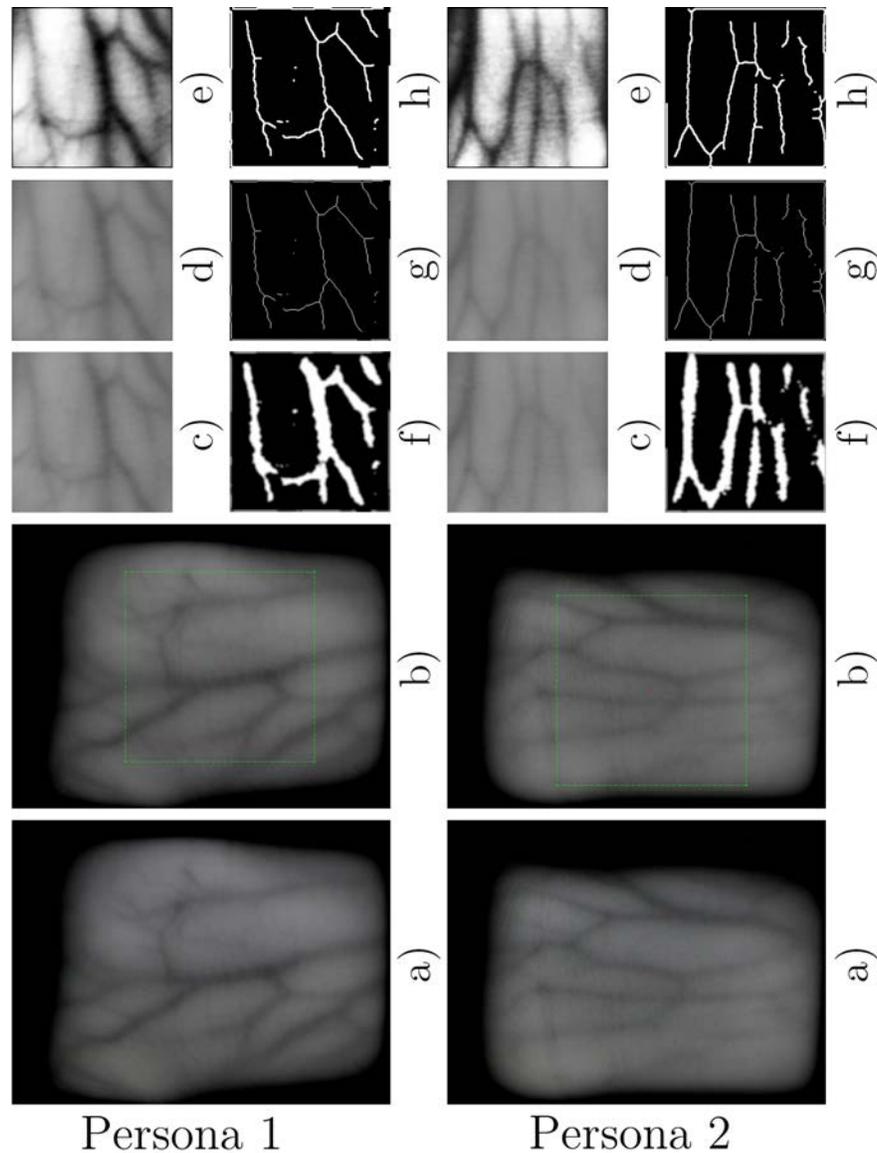


Figura 4.34: (a) Imagen original, (b) Centroide y detección de la región de interés (ROI), (c) Extracción de ROI, (d) Filtrado espacial, (e) Ecualización del histograma, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación e imagen final con el patrón de venas segmentado.

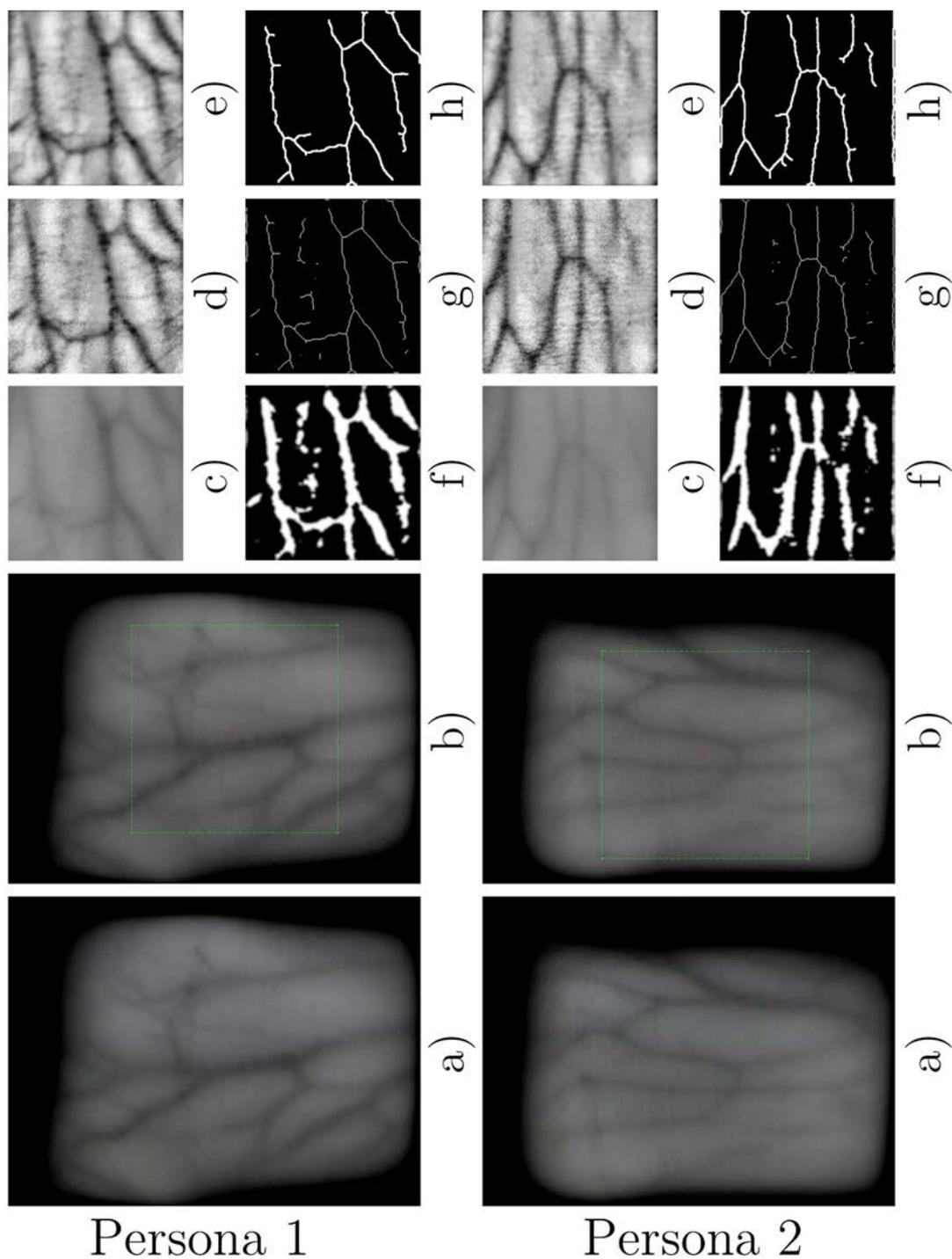


Figura 4.35: (a) Imagen original, (b) Centroide y detección de ROI, (c) Extracción de ROI, (d) Contraste limitado de la ecualización adaptativa del histograma, (e) Filtraje espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

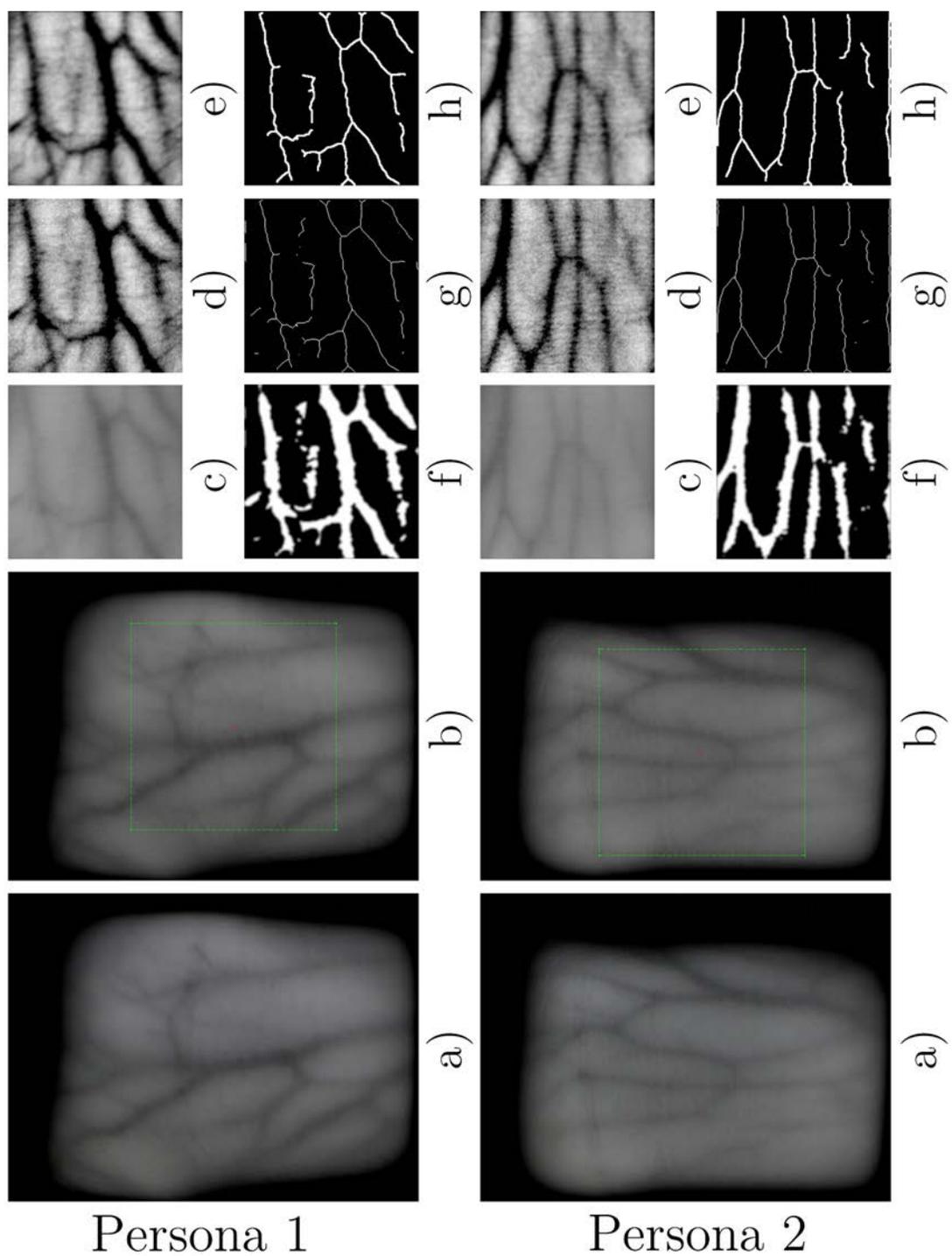


Figura 4.36: (a) Imagen original, (b) Centroides y detección de ROI, (c) Extracción de ROI, (d) Máscara de desenfoque, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

En la Figura 4.37 se hace una comparación entre los métodos de mejoramiento de imágenes basados en la Ecuación del Histograma, CLAHE y la Máscara de desenfoque.

El patrón de venas resultante de entre los tres métodos de extracción implementados, contiene más información cuando se usan las técnicas de Contraste Limitado de la Ecuación del Histograma y la Máscara de desenfoque. Además de que no contienen la falsa información del contorno.

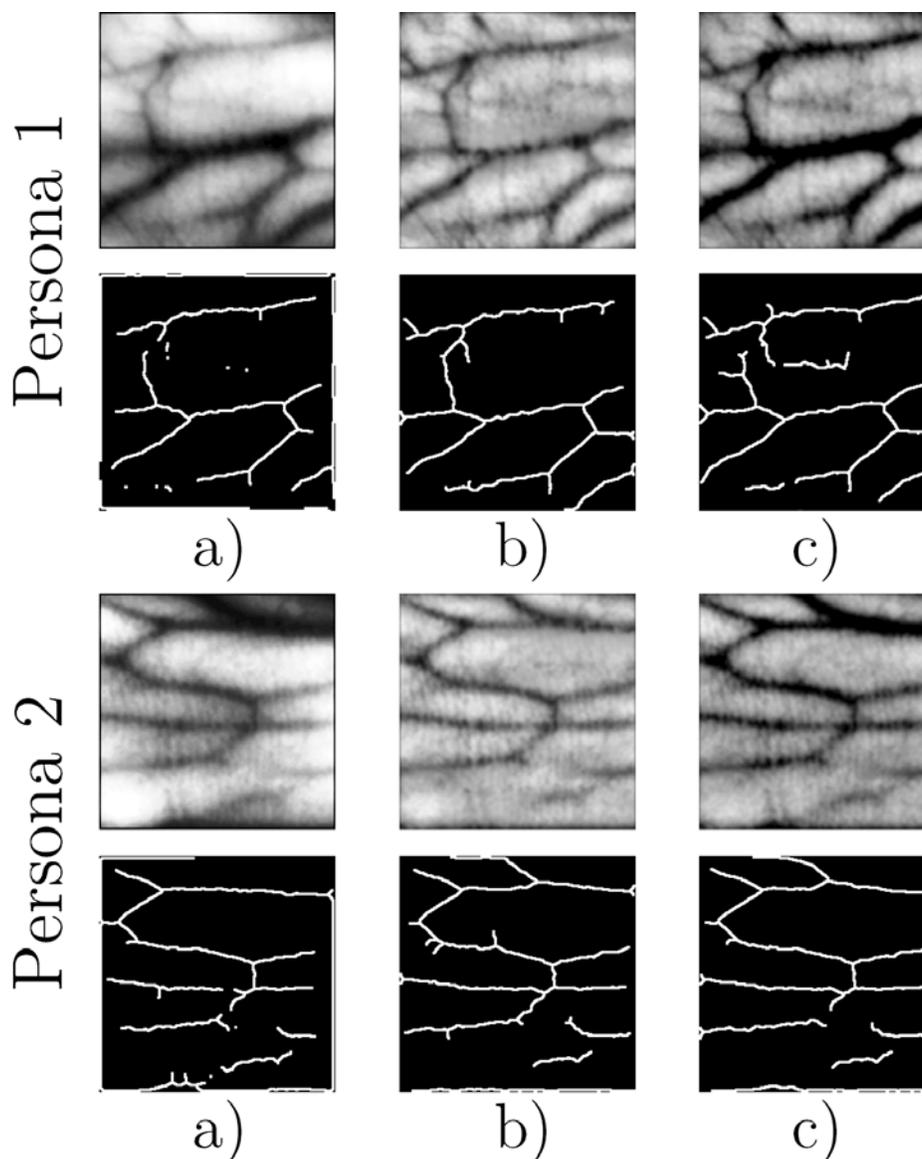


Figura 4.37: Imágenes mejoradas usando: (a) Ecuación del Histograma, (b) Contraste Limitado de la Ecuación Adaptativa del Histograma y (c) Máscara de desenfoque.

4.6. Imágenes adquiridas con el sistema CASIA

La base de datos pertenece al Instituto de Automatización de la Academia de Ciencias China (CASIA) V1.0, y consta de imágenes multispectrales de la palma de la mano, se encuentra publicada con el objetivo de promover la investigación y el progreso en múltiples imágenes espectrales de modalidades biométricas. La base de datos contiene 7200 imágenes de la palma capturadas de 100 personas diferentes usando un dispositivo de imágenes multispectral de diseño propio como se muestra en la Figura 4.38. Todas las imágenes de la palma son guardadas en 8 bits en niveles de gris con formato jpg.

Para cada mano, se capturaron dos sesiones de imágenes de la palma. El intervalo de tiempo entre las dos sesiones es mayor a un mes. En cada sesión se adquirieron 3 muestras. Cada muestra contiene 6 imágenes de la palma las cuales son capturadas al mismo tiempo con 6 diferentes espectros electromagnéticos [9].

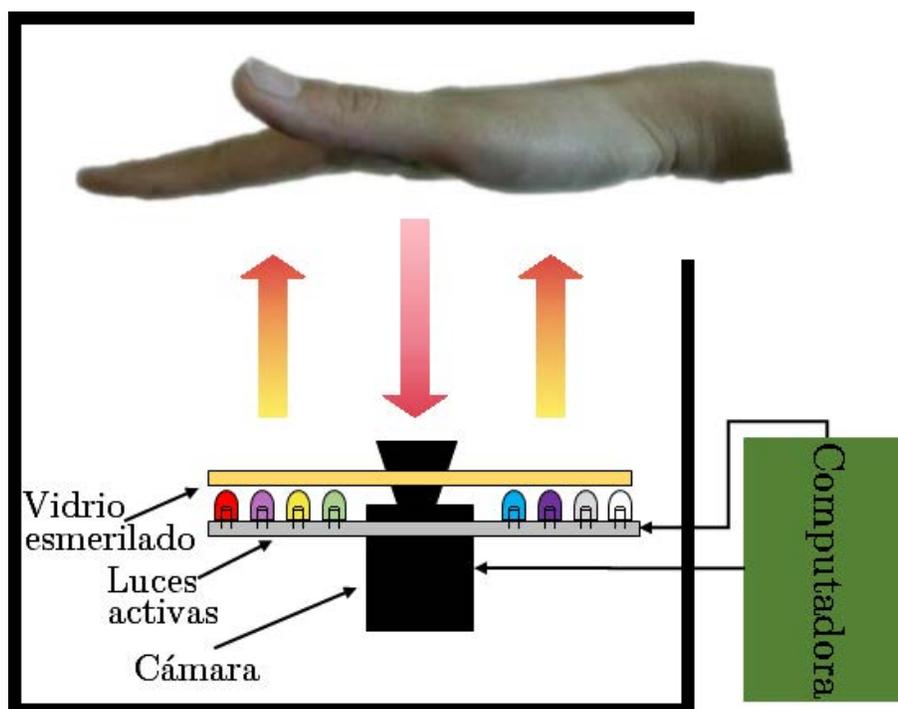


Figura 4.38: Diseño conceptual del dispositivo multispectral de adquisición de imágenes de la palma de la mano [9].

Las longitudes de onda del iluminador correspondientes a los 6 espectros son 460 nm , 630 nm , 700 nm , 850 nm , 940 nm y luz blanca respectivamente. Entre las dos muestras, **se permitió un cierto grado de variaciones en la posición de la mano**. Por eso, el objetivo es incrementar la diversidad de muestras entre clases y **simular el uso práctico**.

En el dispositivo, no hay delimitadores para restringir la postura y posición de las palmas. Se requirió que los sujetos colocaran su palma en el dispositivo y que la pusieran delante de un fondo de color uniforme.

El dispositivo provee una iluminación distribuida uniformemente y captura las imágenes de la palma usando una cámara CCD colocada en la parte baja del dispositivo. Se diseñó además un circuito de control para ajustar automáticamente los espectros [9]. En la Figura 4.39, se muestran 6 imágenes de la base de datos, típicas de la palma.

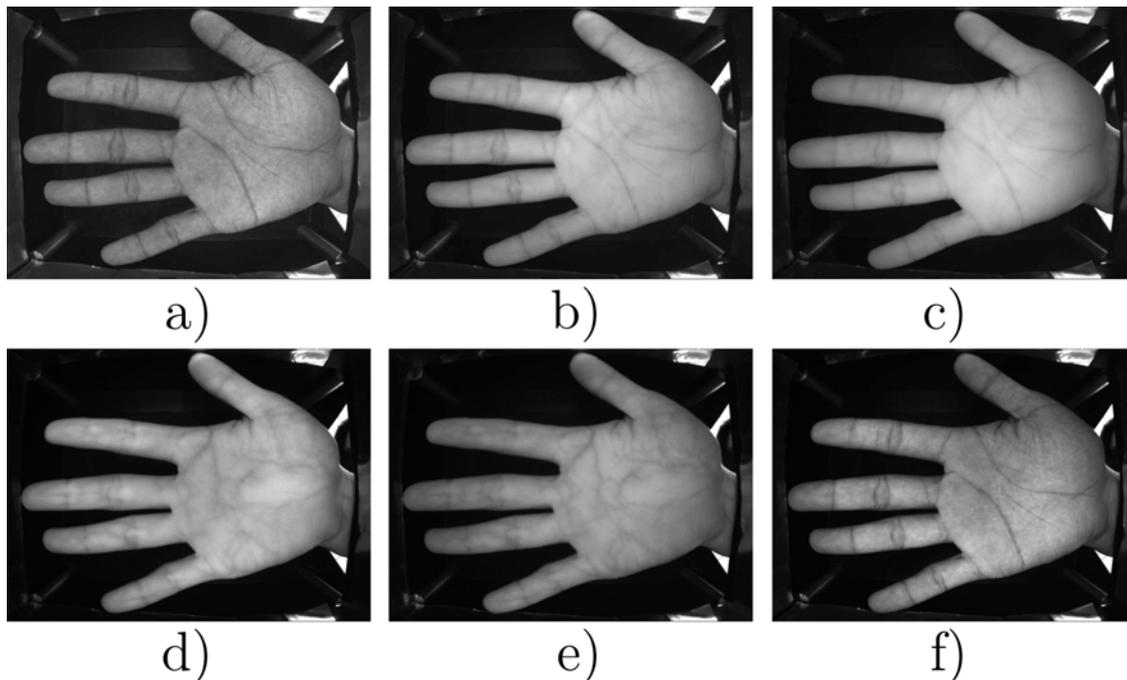


Figura 4.39: Imágenes de la palma de la mano adquiridas en longitudes de onda de: (a) 460 nm , (b) 630 nm , (c) 700 nm , (d) 850 nm , (e) 940 nm y (f) Luz blanca [9].

4.6.1. Base de datos CASIA Multi-Spectral Palmprint Image Database preprocesada

Se procesaron las imágenes que fueron adquiridas en los 850 nm y 940 nm . Los resultados de aplicar las técnicas de mejoramiento de imágenes se muestran en las Figuras 4.40, 4.41 y 4.42. Es importante mencionar que el proceso aplicado es similar al de la sección 4.3.1. Los métodos de mejoramiento de imágenes implementados son: a) Ecuilización del histograma, b) Contraste Limitado de la Ecuilización Adaptativa del Histograma (CLAHE) y c) Máscara de desenfoco.

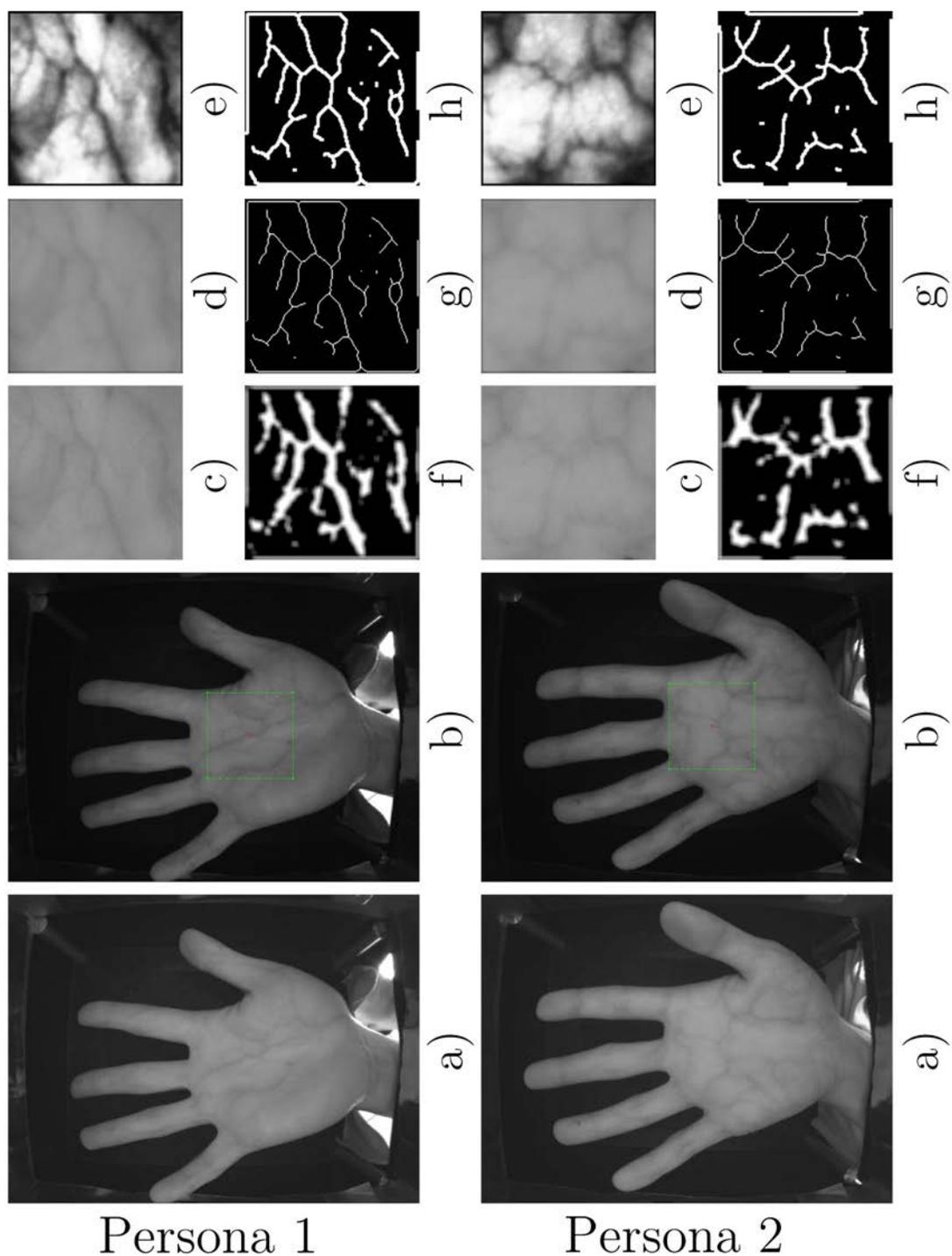


Figura 4.40: (a) Imagen original, (b) Centroide y detección de la región de interés (ROI), (c) Extracción de ROI, (d) Filtrado espacial, (e) Ecualización del histograma, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación e imagen final con el patrón de venas segmentado.

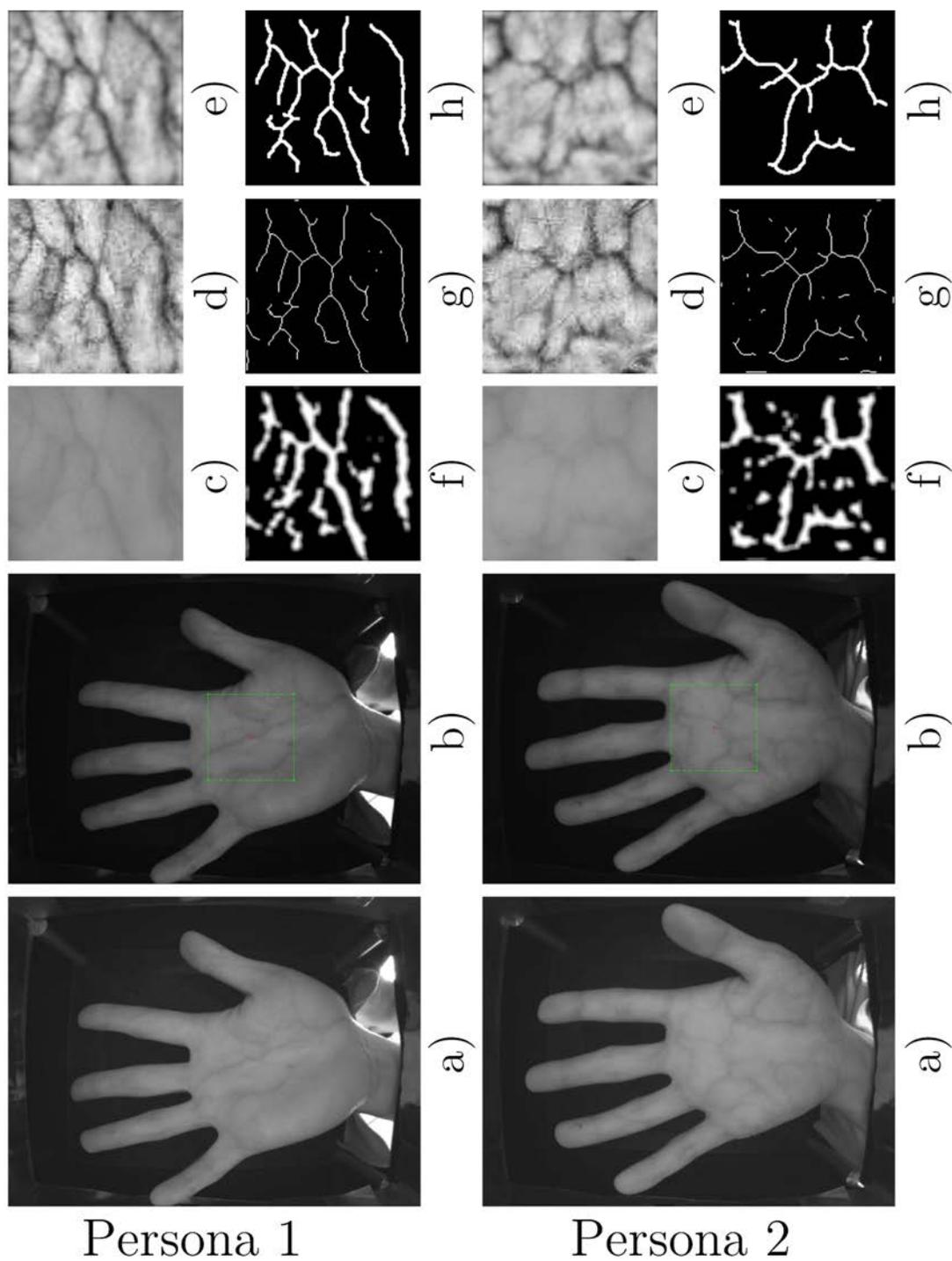


Figura 4.41: (a) Imagen original, (b) Centroide y detección de ROI, (c) Extracción de ROI, (d) Contraste limitado de la ecualización del histograma, (e) Filtrado espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

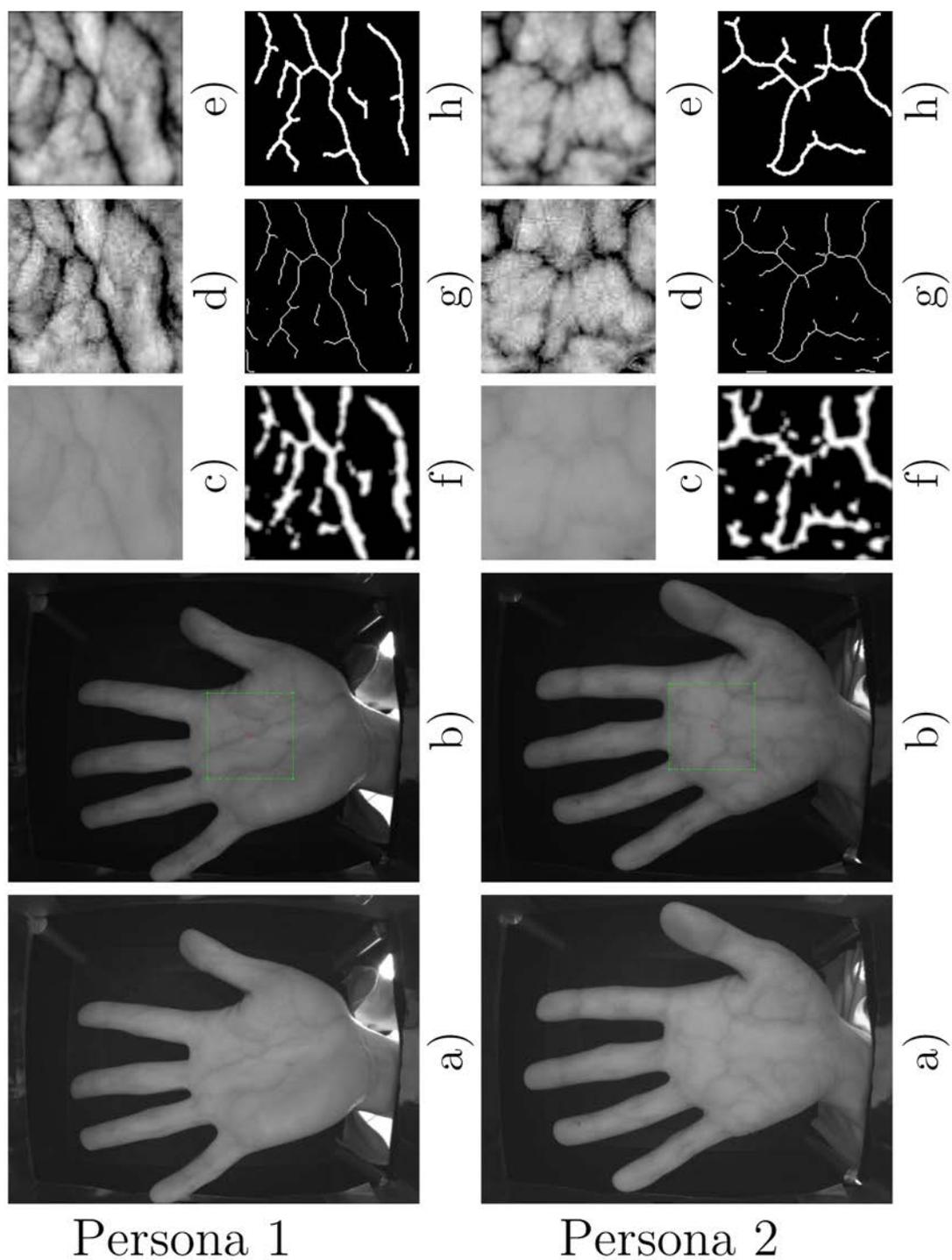


Figura 4.42: (a) Imagen original, (b) Centroide y detección de ROI, (c) Extracción de ROI, (d) Máscara de desenfoque, (e) Filtraje espacial, (f) Binarización adaptativa, (g) Adelgazamiento y (h) Dilatación, eliminación de ruido e imagen final con el patrón de venas segmentado.

En la Figura 4.43 se hace una comparación entre los métodos de mejoramiento de imágenes basados en la Ecuación del Histograma, CLAHE y la Máscara de desenfoque. Los resultados de procesar las imágenes IR de la palma de la mano, son más exactos y tienen menos ruido cuando se aplican los métodos de Contraste Limitado de la Ecuación Adaptativa del Histograma y la Máscara de desenfoque. Adicionalmente, son más robustos a la iluminación no uniforme durante la adquisición de imágenes.

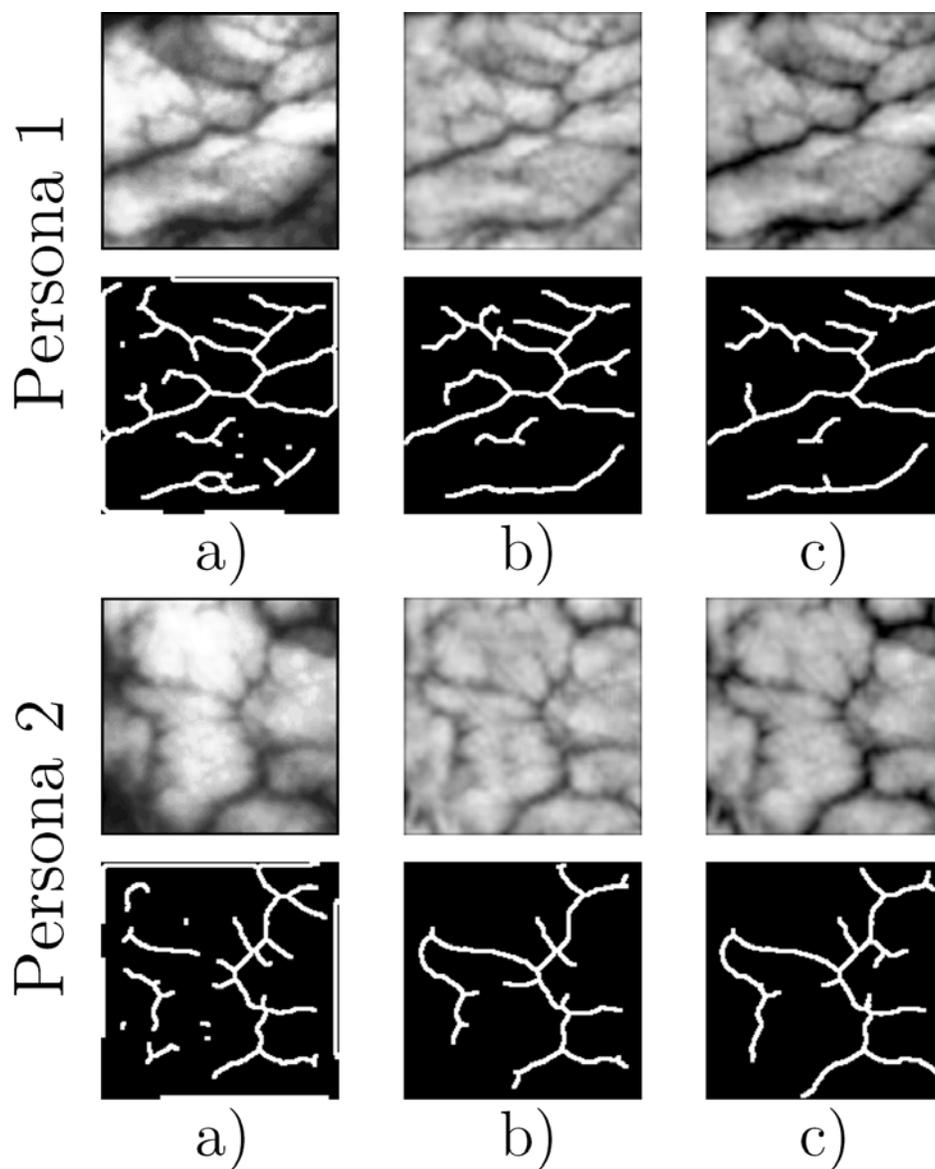


Figura 4.43: Imágenes mejoradas usando: (a) Ecuación del Histograma, (b) Contraste Limitado de la Ecuación Adaptativa del Histograma y (c) Máscara de desenfoque.

4.7. Conclusiones

En este capítulo se describió la configuración del sistema óptico implementado para la adquisición de imágenes de la palma y dorso de la mano. Debido a las características de absorción de los principales componentes del infrarrojo en la sangre, se determinó usar una fuente y sensor entre los 700 a los 950 nanómetros.

El sistema óptico está basado en una configuración por reflexión, que es capaz de detectar el patrón vascular del dorso y palma de la mano. A partir de éstas imágenes adquiridas, es llevado a cabo el proceso de mejoramiento, extracción y descripción de la distribución de venas.

Se desarrolló una interfaz en MATLAB para la captura y mejoramiento de imágenes basados en los métodos de: a) Ecuilización del Histograma, b) Contraste Limitado de la Ecuilización Adaptativa del Histograma (CLAHE) y c) Máscara de desenfoque.

Se llevó a cabo una evaluación comparativa de los algoritmos propuestos para el mejoramiento de la información, para esto, se analizaron las bases de datos:

- **Base de datos UPT (UPT).**
- **PolyU Multispectral Palmprint Database** perteneciente a la Universidad Politécnica de Hong Kong (PUMPD).
- **Vein Pattern Database** perteneciente al Instituto de Control e Ingeniería de la Información de la Universidad Tecnológica de Polonia (VPD).
- **CASIA Multi-Spectral Palmprint Image Database** perteneciente al Instituto de Automatización de la Academia de Ciencias China (CASIA).

En todos los casos, los algoritmos de mejoramiento basados en la Ecuilización del Histograma, realzan la información del patrón de venas, pero existe la presencia de sombras que puede llegar a distorsionar y afectar el resultado final.

De igual forma, en el caso del Contraste Limitado de la Ecuilización Adaptativa del Histograma y la Máscara de desenfoque, se observa tanto un aumento del contraste como la eliminación de sombras en las imágenes. Por lo tanto, estas últimas dos técnicas de mejoramiento de imágenes ayudan a redistribuir los valores de intensidad de los niveles de gris de manera más uniforme haciendo que las imágenes sean más contrastadas.

El análisis de los resultados de las comparaciones anteriores, se pueden observar en las Figuras 4.24, 4.30, 4.37 y 4.43

Las características más importantes de las bases de datos analizadas, se muestran en la Tabla 5.25.

Característica	UPT	PUMPD	VPD	CASIA
Region	DD, DI, PD, PI	PD y PI	DD, DI, PD, PI	PD y PI
Longitud de Onda (λ)	880 <i>nm</i>	***	***	460, 630, 700, 850 940 <i>nm</i> y Luz Blanca
Personas	38	250	50	100
Versiones	12	12	12	6
Total de imágenes	1824	6000	2400	7200
ROI	No	Si	No	No
Tiempo entre captura	1 Día	***	1 Semana	1 Mes
Edad	20 a 25	20 a 60	***	***
Tamaño de imagen	1024 × 768	128 × 128	1024 × 768	768 × 576

Tabla 4.1: Resumen de las características más importantes de las bases de datos analizadas. Simbología: (DD) Dorso Derecho, (DI) Dorso Izquierdo, (PD) Palma Derecha, (PI) Palma Izquierda, (***) Sin Datos.

Bibliografía

- [1] Óscar Fernando Motato Toro and Humberto Loaiza Correa, "Identificación biométrica utilizando imágenes infrarrojas de la red vascular de la cara dorsal de la mano," *Revista Ingeniería e Investigación*, vol. 29, pp. 90-100, 2009.
- [2] Aída Marcotti, Belén Hidalgo and Ladislao Mathé, "Método de detección no invasivo utilizando luz infrarroja," *Universidad Nacional de Córdoba, Argentina, Master Degree Thesis*, 2012.
- [3] Deepak Prasanna. R, Neelamegam. P, Sriram. S and Nagarajan Raju, "Enhancement of vein patterns in hand image for biometric and biomedical application using various image enhancement techniques," *Procedia Engineering*, vol. 38, pp. 1174-1185, 2012.
- [4] www.jai.com, "AD-080GE, Digital 2CCD Progressive Scan Multi-Spectral Camera," *User's Manual*, 2014.
- [5] www.infaimon.com, "Infaimon, Su asesor en visión artificial," *Catálogo de Productos*, 2011.
- [6] <http://www.mathworks.com/matlabcentral/>, "MATLAB, El lenguaje de cálculo técnico," *MATLAB Forums Programmers, MathWorks*.
- [7] <http://www4.comp.polyu.edu.hk/>, "The Hong Kong Polytechnic University (PolyU), Multispectral Palmprint Database," <http://www4.comp.polyu.edu.hk/biometrics/MultispectralPalmprint/MSP.htm>, 2012.
- [8] www.cie.put.poznan.pl, "Universidad Tecnológica de Polonia, The Institute of Control and Information Engineering," <http://biometrics.put.poznan.pl/vein-dataset/download-vpd-2/>, 2014.
- [9] CASIA-MS-PalmprintV1, "Chinese Academy of Sciences Institute of Automation (CASIA)," <http://biometrics.idealtest.org/>, 2012.

Capítulo 5

Análisis de resultados

5.1. Introducción

En este capítulo se analizan los resultados de clasificación de imágenes del patrón de venas de la mano obtenidas en el Infrarrojo Cercano (NIR). Las imágenes procesadas pertenecen a las bases de datos:

- **Base de datos UPT.**
- **PolyU Multispectral Palmprint Database.**
- **Vein Pattern Database.**
- **CASIA Multi-Spectral Palmprint Image Database.**

Los porcentajes de clasificación correcta dependerán de los siguientes factores:

- a) El sistema de adquisición de imágenes: las imágenes procesadas provienen de 4 bases de datos distintas.
- b) Las técnicas de preprocesamiento de imágenes: se implementaron tres métodos para el mejoramiento, realce y segmentación de la distribución del patrón de venas. las técnicas están basadas en Ecuilización del Histograma, Contraste Limitado de la Ecuilización Adaptativa del Histograma y Máscara de desenfoque.
- c) Las técnicas de extracción de características: los métodos clásicos para la descripción de objetos están basados en los Momentos Geométricos y los Momentos de Hu, además, se propuso un método denominado Descriptor de Venas.
- d) El método de clasificación. Se utilizaron dos algoritmos, uno basado en K Vecinos Próximos (K-nn) y el otro en Redes Neuronales Artificiales (RNA).

5.2. Resultados de clasificación para la base de datos UPT

A partir de las imágenes preprocesadas de la base de datos UPT generadas en la Sección 4.3.1, se extraen un conjunto de descriptores de la distribución del patrón de venas mostrado previamente en la Figura 4.24. Dichos descriptores están basados en Momentos Geométricos, Momentos de Hu y Descriptor de Venas.

La base de datos UPT consta de 152 patrones de la distribución de venas de la palma y dorso de ambas manos obtenidas a partir de 38 personas con 12 versiones cada una, dando un total de 1824 imágenes.

A continuación, en las Tablas 5.1 y 5.2, se muestran los porcentajes de clasificación utilizando las técnicas clásicas de preprocesamiento basadas en la **Ecuación del Histograma** así como de las técnicas descritas en la Figura 4.21.

En las Tablas 5.3 y 5.4, se muestran los porcentajes de clasificación usando el **Contraste Limitado de la Ecuación Adaptativa del Histograma** y las técnicas descritas en la Figura 4.22.

En las Tablas 5.5 y 5.6, se muestran los porcentajes de clasificación usando la **Máscara de desenfoque** y las técnicas descritas en la Figura 4.23. Los algoritmos de clasificación utilizados son K Vecinos Próximos y Redes Neuronales Artificiales.

Clasificación basada en las técnicas clásicas de mejoramiento de imágenes

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	42.93 %	23.90 %	76.54 %
Dorso Izquierdo	35.31 %	19.52 %	69.96 %
Palma Derecha	32.68 %	10.09 %	62.94 %
Palma Izquierda	27.85 %	9.87 %	72.15 %

Tabla 5.1: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

5.2. RESULTADOS DE CLASIFICACIÓN PARA LA BASE DE DATOS UPT

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	63.60 %	32.46 %	78.73 %
Dorso Izquierdo	50.00 %	24.56 %	70.39 %
Palma Derecha	52.63 %	17.32 %	70.18 %
Palma Izquierda	45.83 %	16.67 %	72.59 %

Tabla 5.2: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Contraste Limitado de la Ecuación Adaptativa del Histograma

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	29.17 %	14.04 %	55.70 %
Dorso Izquierdo	24.34 %	17.76 %	35.09 %
Palma Derecha	23.46 %	12.74 %	11.62 %
Palma Izquierda	24.78 %	10.96 %	23.03 %

Tabla 5.3: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	41.45 %	15.35 %	62.50 %
Dorso Izquierdo	37.72 %	19.74 %	42.11 %
Palma Derecha	10.96 %	12.50 %	24.78 %
Palma Izquierda	36.62 %	13.38 %	34.43 %

Tabla 5.4: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Máscara de desenfoque

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	40.13 %	21.71 %	59.43 %
Dorso Izquierdo	35.53 %	26.10 %	42.32 %
Palma Derecha	25.88 %	25.48 %	15.57 %
Palma Izquierda	34.43 %	23.72 %	17.98 %

Tabla 5.5: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	59.21 %	24.12 %	72.15 %
Dorso Izquierdo	54.82 %	30.26 %	50.22 %
Palma Derecha	13.38 %	17.11 %	25.66 %
Palma Izquierda	20.18 %	17.11 %	27.41 %

Tabla 5.6: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Los porcentajes de clasificación de imágenes IR de la palma y dorso de la mano que pertenecen a la base de datos UPT son mayores cuando se usan:

- A) Las técnicas de procesamiento Clásicas para el mejoramiento, realce y segmentación de imágenes.
- B) Los Descriptores de Venas para la extracción de características del patrón.
- C) Redes Neuronales Artificiales para la clasificación en 38 clases o personas.

5.3. Resultados de clasificación para la base de datos de la UPHK

A partir de las imágenes preprocesadas de la base de datos PolyU Multispectral Palmprint Database generadas en la Sección 4.4.1, se extraen un conjunto de descriptores de la distribución del patrón de venas mostrado previamente en la Figura 4.30.

La base de datos PolyU Multispectral Palmprint Database consta de 500 patrones de la distribución de venas de la palma de la mano obtenidas a partir de 250 personas con 12 versiones cada una, dando un total de 6000 imágenes.

En las Tablas 5.7 y 5.8, se muestran los porcentajes de clasificación se muestran los porcentajes de clasificación utilizando las técnicas clásicas de preprocesamiento basadas en la **Ecuación del Histograma** así como de las técnicas descritas en la Figura 4.27.

En las Tablas 5.9 y 5.10, se muestran los porcentajes de clasificación usando el **Contraste Limitado de la Ecuación Adaptativa del Histograma** y las técnicas descritas en la Figura 4.28.

En las Tablas 5.11 y 5.12, se muestran los porcentajes de clasificación usando la **Máscara de desenfoque** y las técnicas descritas en la Figura 4.29. Los algoritmos de clasificación utilizados son K Vecinos Próximos y Redes Neuronales Artificiales.

Clasificación basada en las técnicas clásicas de mejoramiento de imágenes

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma	11.32 %	0.65 %	96.17 %

Tabla 5.7: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma	54.23 %	26.41 %	97.50 %

Tabla 5.8: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Contraste Limitado de la Ecuación Adaptativa del Histograma

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma	6.20 %	0.42 %	65.33 %

Tabla 5.9: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma	46.22 %	18.73 %	85.19 %

Tabla 5.10: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Máscara de desenfoque

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma	11.73 %	0.55 %	92.23 %

Tabla 5.11: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma	49.38 %	16.33 %	82.86 %

Tabla 5.12: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Los porcentajes de clasificación de imágenes IR de la palma de la mano que pertenecen a la base de datos PolyU Multispectral Palmprint Database son mayores cuando se usan:

- A) Las técnicas de procesamiento Clásicas para el mejoramiento, realce y segmentación de imágenes.
- B) Los Descriptores de Venas para la extracción de características del patrón.
- C) Redes Neuronales Artificiales para la clasificación en 38 clases o personas.

5.4. Resultados de clasificación para la base de datos VPD

A partir de las imágenes preprocesadas de la base de datos Vein Pattern Database generadas en la Sección 4.5.1, se extraen un conjunto de descriptores de la distribución del patrón de venas como las mostradas en la Figura 4.37. Dichos descriptores están basados en Momentos Geométricos, Momentos de Hu y Descriptor de Venas.

La base de datos Vein Pattern Database consta de 100 patrones de la distribución de venas de la palma y muñeca de la mano obtenidas a partir de 50 personas con 12 versiones cada una, dando un total de 2400 imágenes.

En las Tablas 5.13 y 5.14, se muestran los porcentajes de clasificación se muestran los porcentajes de clasificación utilizando las técnicas clásicas de preprocesamiento basadas en la **Ecuación del Histograma** así como de las técnicas descritas en la Figura 4.34.

En las Tablas 5.15 y 5.16, se muestran los porcentajes de clasificación usando el **Contraste Limitado de la Ecuación Adaptativa del Histograma** y las técnicas descritas en la Figura 4.35.

En las Tablas 5.17 y 5.18, se muestran los porcentajes de clasificación usando la **Máscara de desenfoque** y las técnicas descritas en la Figura 4.36.

Clasificación basada en las técnicas clásicas de mejoramiento de imágenes

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	24.00 %	7.83 %	41.83 %
Dorso Izquierdo	18.83 %	10.67 %	42.00 %
Palma Derecha	20.83 %	10.53 %	40.00 %
Palma Izquierda	18.67 %	10.00 %	39.00 %

Tabla 5.13: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

5.4. RESULTADOS DE CLASIFICACIÓN PARA LA BASE DE DATOS VPD

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	25.33 %	6.33 %	43.00 %
Dorso Izquierdo	23.33 %	7.50 %	36.83 %
Palma Derecha	31.83 %	12.33 %	42.67 %
Palma Izquierda	27.17 %	9.67 %	35.17 %

Tabla 5.14: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Contraste Limitado de la Ecuación Adaptativa del Histograma

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	11.33 %	5.00 %	34.33 %
Dorso Izquierdo	11.00 %	6.00 %	28.33 %
Palma Derecha	11.83 %	6.00 %	30.33 %
Palma Izquierda	9.17 %	3.17 %	29.17 %

Tabla 5.15: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	17.67 %	9.00 %	29.83 %
Dorso Izquierdo	19.50 %	7.50 %	22.33 %
Palma Derecha	19.50 %	6.67 %	26.67 %
Palma Izquierda	18.83 %	6.67 %	23.00 %

Tabla 5.16: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Máscara de desenfoque

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	13.17 %	8.33 %	41.67 %
Dorso Izquierdo	12.83 %	6.83 %	41.33 %
Palma Derecha	19.67 %	4.50 %	39.17 %
Palma Izquierda	15.33 %	7.67 %	36.67 %

Tabla 5.17: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Dorso Derecho	24.83 %	11.50 %	34.33 %
Dorso Izquierdo	25.17 %	9.33 %	33.00 %
Palma Derecha	30.33 %	9.33 %	36.50 %
Palma Izquierda	28.33 %	9.00 %	35.17 %

Tabla 5.18: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Los porcentajes de clasificación son mayores para:

- A) Las técnicas de procesamiento Clásicas.
- B) Descriptores de Venas.
- C) Mediante el algoritmo de K vecinos próximos y Redes Neuronales Artificiales.

5.5. Resultados de clasificación para la base de datos CASIA

SIA

A partir de las imágenes preprocesadas de la base de datos CASIA Multi-Spectral Palmprint Image Database generadas en la Sección 4.6.1, se extraen un conjunto de descriptores de la distribución del patrón de venas como las mostradas en la Figura 4.43. Dichos descriptores están basados en Momentos Geométricos, Momentos de Hu y Descriptor de Venas.

La base de datos CASIA Multi-Spectral Palmprint Image Database consta de 7200 patrones de la distribución de venas de la palma de la mano adquiridas en 2 longitudes de onda diferentes (850 *nm* y 940 *nm*) obtenidas a partir de 100 personas con 6 versiones cada una, dando un total de 2400 imágenes.

En las Tablas 5.19 y 5.20, se muestran los porcentajes de clasificación se muestran los porcentajes de clasificación utilizando las técnicas clásicas de procesamiento basadas en la **Ecuación del Histograma** así como de las técnicas descritas en la Figura 4.40.

En las Tablas 5.21 y 5.22, se muestran los porcentajes de clasificación usando el **Contraste Limitado de la Ecuación Adaptativa del Histograma** y las técnicas descritas en la Figura 4.41.

En las Tablas 5.23 y 5.24, se muestran los porcentajes de clasificación usando la **Máscara de desenfoque** y las técnicas descritas en la Figura 4.42.

Clasificación basada en las técnicas clásicas de mejoramiento de imágenes

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma Derecha 850 <i>nm</i>	12.17 %	2.17 %	25.50 %
Palma Izquierda 850 <i>nm</i>	16.50 %	3.17 %	30.67 %
Palma Derecha 940 <i>nm</i>	15.33 %	2.17 %	26.50 %
Palma Izquierda 940 <i>nm</i>	16.17 %	1.83 %	26.33 %

Tabla 5.19: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma Derecha 850 <i>nm</i>	11.33 %	0.33 %	27.83 %
Palma Izquierda 850 <i>nm</i>	8.83 %	1.17 %	35.83 %
Palma Derecha 940 <i>nm</i>	11.33 %	0.87 %	29.83 %
Palma Izquierda 940 <i>nm</i>	8.50 %	0.83 %	33.83 %

Tabla 5.20: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Contraste Limitado de la Ecuación Adaptativa del Histograma

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma Derecha 850 <i>nm</i>	13.83 %	4.00 %	16.33 %
Palma Izquierda 850 <i>nm</i>	14.33 %	2.17 %	21.33 %
Palma Derecha 940 <i>nm</i>	13.33 %	4.67 %	18.33 %
Palma Izquierda 940 <i>nm</i>	14.33 %	5.17 %	21.00 %

Tabla 5.21: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma Derecha 850 <i>nm</i>	14.00 %	2.33 %	22.83 %
Palma Izquierda 850 <i>nm</i>	14.17 %	1.17 %	23.50 %
Palma Derecha 940 <i>nm</i>	16.00 %	3.33 %	23.17 %
Palma Izquierda 940 <i>nm</i>	15.83 %	1.83 %	24.17 %

Tabla 5.22: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Clasificación basada en la técnica de Máscara de desenfoque

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma Derecha 850 <i>nm</i>	12.17 %	2.17 %	25.50 %
Palma Izquierda 850 <i>nm</i>	16.50 %	3.17 %	30.67 %
Palma Derecha 940 <i>nm</i>	15.33 %	2.17 %	26.50 %
Palma Izquierda 940 <i>nm</i>	16.17 %	1.83 %	26.33 %

Tabla 5.23: Porcentajes de clasificación correcta usando el algoritmo de los K vecinos próximos (K-nn) con K=3.

Región	Momentos Geométricos	Momentos de Hu	Descriptor de Venas
Palma Derecha 850 <i>nm</i>	13.33 %	1.67 %	22.50 %
Palma Izquierda 850 <i>nm</i>	11.67 %	0.83 %	21.50 %
Palma Derecha 940 <i>nm</i>	17.67 %	3.83 %	23.83 %
Palma Izquierda 940 <i>nm</i>	14.83 %	0.83 %	26.33 %

Tabla 5.24: Porcentajes de clasificación correcta usando Redes Neuronales Artificiales (RNA).

Los porcentajes de clasificación son mayores para:

- A) Las técnicas de procesamiento Clásicas y la Máscara de desenfoque.
- B) Descriptores de Venas.
- C) Redes Neuronales Artificiales.

5.6. Conclusiones

Con el propósito de evaluar la eficiencia de los algoritmos propuestos, se procesaron las imágenes en el infrarrojo cercano de la palma y dorso de la mano pertenecientes a 4 bases de datos diferentes. Las características generales de las bases de datos se resumen en el Cuadro 5.25:

Característica	UPT	PUMPD	VPD	CASIA
Region	DD, DI, PD, PI	PD y PI	DD, DI, PD, PI	PD y PI
Longitud de Onda (λ)	880 <i>nm</i>	***	***	460, 630, 700, 850 940 <i>nm</i> y Luz Blanca
Personas	38	250	50	100
Versiones	12	12	12	6
Total de imágenes	1824	6000	2400	7200
ROI	No	Si	No	No
Tiempo entre captura	1 Día	***	1 Semana	1 Mes
Edad	20 a 25	20 a 60	***	***
Tamaño de imagen	1024 × 768	128 × 128	1024 × 768	768 × 576

Tabla 5.25: Resumen de las características más importantes de las bases de datos analizadas. Simbología: (DD) Dorso Derecho, (DI) Dorso Izquierdo, (PD) Palma Derecha, (PI) Palma Izquierda, (***) Sin Datos.

En la mayoría de los casos, se obtuvieron los porcentajes de clasificación más altos cuando:

- A) El procesamiento se llevaba a cabo usando los métodos clásicos que se encuentran basados en la Ecuación del Histograma.
- B) Se usa como método de extracción de características, el propuesto, el denominado "Descriptor de Venas".
- C) Se utilizan Redes Neuronales Artificiales (RNA) como técnica de clasificación.

Los resultados reportados para la clasificación utilizando RNA, presentan una validación cruzada (k-fold cross validation) de 10 divisiones. Para el caso de la clasificación con K vecinos próximos, se utilizó un K=3 y se realiza una validación mediante Leave One Out (Dejando uno fuera).

La razón por la que los porcentajes de clasificación son bajos para la base de datos CASIA es debido a las variaciones en la iluminación que se encuentra presente en la imagen adquirida. Este cambio de iluminación afecta el correcto funcionamiento del algoritmo de segmentación de la ROI.

Lo anterior, provoca que la región de interés no sea la misma para todas las imágenes, ya que en algunas imágenes se encontrará desplazada del centroide de la imagen. El método propuesto alcanza porcentajes de clasificación correcta de hasta un 97.5 % para las 6000 imágenes de la Base de Datos PolyU Multispectral Palmprint de la Universidad Politécnica de Hong Kong.

Resultados similares son reportados en la literatura cuando a) se toma un conjunto de varias biometrías, b) la base de datos es pequeña y c) los porcentajes son bajos para bases de datos de dimensiones similares. En la Tabla 5.26, se muestra un resumen con los porcentajes de clasificación así como la cantidad de personas e imágenes que fueron utilizadas. Además se muestra la región de la mano utilizada como por ejemplo el dorso, los dedos o la palma.

Referencia	Número de Imágenes	Número de Personas	Número de Versiones	Porcentaje de Clasificación	Lugar
Jian-Da Wu <i>et al.</i>	500	25	10	99.2	Dedos
Qiushi Zhao <i>et al.</i>	10800	90	***	***	Palma y Dorso
Eui Chul Lee <i>et al.</i>	8160	102	***	62.18	Dedos
Byung Jun Kang <i>et al.</i>	6380	80	***	98.56	Dedos
David Zhang <i>et al.</i>	6000	500	12	98.8	Palma
Raúl Castro <i>et al.</i>	6000	500	12	97.5	Palma

Tabla 5.26: Comparativa de porcentajes de clasificación de diversos trabajos relacionados al reconocimiento de personas a través de la distribución del patrón de venas.

Capítulo 6

Conclusiones

Un sistema biométrico típico del patrón de venas, consta de 5 etapas de procesamiento individual: adquisición de imágenes, realce de imágenes, segmentación del patrón de venas, extracción de características y clasificación. A lo largo de esta Tesis, se analizaron diversas técnicas de procesamiento digital de imágenes para la extracción del patrón de venas de una persona, a partir del análisis de imágenes adquiridas en el infrarrojo cercano de diferentes regiones de la mano.

Para esto, se implementó un sistema óptico digital en el Laboratorio de Óptica y Sistemas de Visión de la Universidad Politécnica de Tulancingo, que adquiere imágenes de la palma y/o dorso de la mano de las personas. Dicho sistema esta basado en una configuración por reflexión.

Debido a las características de absorción de los principales componentes del infrarrojo en la sangre, se determinó utilizar una fuente de iluminación de LED's en el infrarrojo cercano, esta fuente en forma de anillo, cuenta con un pico de emisión a los 880 nm. Se utilizó además una cámara Multispectral 2CCD de escaneo progresivo, la cual es capaz de adquirir imágenes tanto en el espectro visible (400 nm - 700 nm) como en el infrarrojo (700 nm - 1000 nm).

Se desarrolló además, una interfaz para controlar la cámara Multiespectral a través de su puerto de comunicaciones Giga Ethernet utilizando para ello el entorno MATLAB. A través de ésta interfaz, es posible visualizar en tiempo real la imagen a adquirir. Posteriormente, se tiene la posibilidad de aplicar el preprocesamiento y observar el patrón de venas utilizando varios algoritmos para el mejoramiento, realce y segmentación.

A las imágenes adquiridas mediante la interfaz y el sistema óptico, se les calculó el centroide o centro de gravedad utilizando los Momentos Geométricos. A partir del centro calculado, se trazó un cuadro delimitador del área a analizar. Al proceso de delimitación de esta zona se le conoce como extracción de la Región de Interés, y a la zona o región ya delimitada como Región de Interés o ROI.

La ROI es utilizada ampliamente debido a que permite reducir considerablemente el tiempo de cómputo en las áreas de Mejoramiento y Realce, Segmentación, Extracción de características y por ende, la tarea de Clasificación. Cabe mencionar que las imágenes adquiridas son de un tamaño de 1024×768 píxeles. Obteniendo la ROI a estas imágenes, el tamaño se reduce a 258×258 píxeles.

Los algoritmos implementados para el mejoramiento y realce del patrón de venas están basados en: a) Ecualización del Histograma, Contraste Limitado de la Ecualización Adaptativa del Histograma (CLAHE) y Máscara de Desenfoque (USM).

El algoritmo de Ecualización del Histograma aumenta significativamente el contraste ya que redistribuye de manera más uniforme los niveles de gris de una imagen. Sin embargo, esta técnica no es muy útil cuando hay grandes variaciones de iluminación. Caso contrario sucede con las técnicas de CLAHE y USM.

El algoritmo de CLAHE redistribuye de manera mucho más uniforme los valores de intensidad de los niveles de gris por regiones de un tamaño $m \times n$, por lo que las imágenes con bajo contraste tienen una mejora considerable después de aplicar esta técnica.

La técnica de Máscara de Desenfoque, se puede utilizar para quitar ruido y para mejoras en el contraste de la imagen. Utiliza una imagen suavizada para crear una máscara de la imagen original.

Con el objetivo de remover o eliminar la información innecesaria de los pliegues y vello en la palma y dorso, respectivamente, que estuviese presente en las imágenes de la palma o dorso de la mano, se utilizó el filtraje espacial pasa bajas. Esta información innecesaria de los pliegues y el vello, generan ruido haciendo que el patrón de venas se vea distorsionado, provocando falsa información acerca de la posición de las venas.

Adicionalmente para la binarización y segmentación de las imágenes mejoradas se emplearon, la técnica de binarización adaptativa y las operaciones de morfología matemática con el propósito de obtener únicamente la distribución de venas. A partir de las imágenes segmentadas y binarizadas, es posible extraer un conjunto de descriptores de forma. Las técnicas de extracción de características están basadas en Momentos Geométricos, Invariantes de Hu y un método propuesto en éste trabajo de tesis denominado "Descriptor de Venas".

Los descriptores o momentos de una imagen, nos dieron información tanto global como local de la distribución de intensidades de una imagen. Como se puede apreciar en la Figura 3.1, los descriptores basados en momentos, son invariantes al cambio en la posición, orientación y escala de una imagen.

Los descriptores de venas son vectores binarios característicos generados a partir de la cantidad de información relevante que hay en una región dada.

Con este método se busca obtener un vector característico de cada imagen con una menor cantidad de datos en comparación a la imagen original pero manteniendo información discriminante del patrón de venas. Para extraer las características de la estructura del patrón de venas, se realiza un conteo de píxeles mediante un sub-muestreo por regiones. El sub-muestreo consiste en subdividir en pequeñas zonas o regiones de igual tamaño a la imagen que se desea obtener los descriptores.

Los descriptores obtenidos alimentan una red neuronal (perceptrón multicapa) la cual usa un algoritmo de aprendizaje supervisado para entrenamiento conocido como propagación hacia atrás (Back Propagation).

Además se realiza un estudio comparativo usando un clasificador lineal llamado k vecinos próximos (k nearest neighbors). Como se observa en la Figura 3.11 los resultados de clasificación dependerán del número de k vecinos a considerar.

Como se observó, los mejores resultados de mejoramiento y realce de información se obtienen a partir del uso de las técnicas de Contraste Limitado de la Ecuación Adaptativa del Histograma (CLAHE) y de la Máscara de Desenfoque (Unsharp Masking) como se aprecia en las Figuras 2.9 y 2.10.

Estas últimas no dependen de una iluminación uniforme de la escena. Con propósitos de evaluación de la eficiencia de los algoritmos, se analizaron las siguientes bases de datos:

- **Base de datos UPT (UPT).**
- **PolyU Multispectral Palmprint Database** perteneciente a la Universidad Politécnica de Hong Kong (PUMPD).
- **Vein Pattern Database** perteneciente al Instituto de Control e Ingeniería de la Información de la Universidad Tecnológica de Polonia (VPD).
- **CASIA Multi-Spectral Palmprint Image Database** perteneciente al Instituto de Automatización de la Academia de Ciencias China (CASIA).

En todos los casos, los algoritmos de mejoramiento basados en la Ecuación del Histograma, a pesar de que realzan la información del patrón de venas, existe la presencia de sombras que pueden llegar a distorsionar y afectar el resultado final.

De igual forma, en el caso del Contraste Limitado de la Ecuación Adaptativa del Histograma y la Máscara de desenfoque, se observa un aumento del contraste así como la eliminación

de sombras en las imágenes. Por lo tanto, estas últimas dos técnicas de mejoramiento de imágenes ayudan a redistribuir los valores de intensidad de los niveles de gris de manera más uniforme haciendo que las imágenes sean más contrastadas.

El análisis de los resultados de las comparaciones anteriores, se pueden observar en las Figuras 4.24, 4.30, 4.37 y 4.43. Los resultados reportados para la clasificación utilizando RNA, presentan una validación cruzada (k-fold cross validation) de 10 divisiones. Para el caso de la clasificación con K vecinos próximos, se utilizó un $K=3$ y se realiza una validación mediante Leave One Out (Dejando uno fuera).

En la mayoría de los casos, se obtuvieron los porcentajes de clasificación más altos cuando:

- A) El procesamiento se llevaba a cabo usando los métodos clásicos que se encuentran basados en la Ecuación del Histograma.
- B) Se usa como método de extracción de características, el propuesto, el denominado "Descriptor de Venas".
- C) Se utilizan Redes Neuronales Artificiales (RNA) como técnica de clasificación.

Además, es sabido que, conforme aumenta el número de personas, disminuye el porcentaje de clasificación. Sin embargo, como se puede apreciar en la Figura 6.1, usando como método de extracción el Descriptor de Venas, el porcentaje de clasificación disminuye desde un punto porcentual hasta solo algunas décimas.

Personas Procesadas	Descriptor de Venas
25	98.66 %
50	97.83 %
100	96.41 %
150	97.44 %
200	97.16 %

Figura 6.1: Ejemplo de aumento de personas y disminución del porcentaje de clasificación.

En las Figuras 6.2 y 6.3, se muestran los resultados generales. Como se puede observar, el método "Descriptor de Venas" tiene los porcentajes de clasificación correcta más altos.

K vecinos próximos = 3		Técnicas Clásicas			Contraste Limitado de la Ecuación Adaptativa del Histograma			Máscara de desenfoque		
		VeinCode	Geométricos	Hu	VeinCode	Geométricos	Hu	VeinCode	Geométricos	Hu
UPT	Palma Derecha	62.94	32.68	10.09	11.62	23.46	12.74	15.57	25.88	25.48
	Palma Izquierda	72.15	27.85	9.87	23.03	24.78	10.96	17.98	34.43	23.72
	Dorso Derecho	76.54	42.93	23.90	55.70	29.17	14.04	59.43	40.13	21.71
	Dorso Izquierdo	69.96	35.31	19.52	35.09	24.34	17.76	42.32	35.53	26.10
UPHK	Palmas I-D	96.17	11.32	0.65	65.33	6.20	0.42	87.23	11.73	0.55
VPD	Palma Derecha	40.00	20.83	10.50	30.33	11.83	6.00	39.17	19.67	4.50
	Palma Izquierda	39.00	18.67	10.00	29.17	9.17	3.17	36.67	15.33	7.67
	Dorso Derecho	41.83	24.00	7.83	34.33	11.33	5.00	41.67	13.17	8.33
	Dorso Izquierdo	42.00	18.83	10.67	28.33	11.00	6.00	41.33	12.83	6.83
CASIA	Palma Derecha 850	25.50	12.17	2.17	16.33	13.83	4.00	25.50	12.17	2.17
	Palma Izquierda 850	30.67	16.50	3.17	21.33	14.33	2.17	30.67	16.50	3.17
	Palma Derecha 940	26.50	15.33	2.17	18.00	13.33	4.67	26.50	15.33	2.17
	Palma Izquierda 940	26.33	16.17	1.83	21.00	14.33	5.17	26.33	16.17	1.83

Figura 6.2: Porcentajes de clasificación obtenidos para las 4 bases de datos con K Vecinos Próximos.

Redes Neuronales Artificiales		Técnicas Clásicas			Contraste Limitado de la Ecuación Adaptativa del Histograma			Máscara de desenfoque		
		VeinCode	Geométricos	Hu	VeinCode	Geométricos	Hu	VeinCode	Geométricos	Hu
UPT	Palma Derecha	70.18	52.63	17.32	24.78	10.96	12.50	25.66	13.38	17.11
	Palma Izquierda	72.59	45.83	16.67	34.43	36.62	13.38	27.41	20.18	17.11
	Dorso Derecho	78.73	63.60	32.46	62.50	41.45	15.35	72.15	59.21	24.12
	Dorso Izquierdo	70.39	50.00	24.56	42.11	37.72	19.74	50.22	54.82	30.26
UPHK	Palmas I-D	97.50	24.50	0.62	78.32	12.15	0.42	92.23	19.73	0.43
VPD	Palma Derecha	42.67	31.83	12.33	26.67	19.50	6.67	36.50	30.33	9.33
	Palma Izquierda	35.17	27.17	9.67	23.00	18.83	6.67	35.17	28.33	9.00
	Dorso Derecho	43.00	25.33	6.33	29.83	17.67	9.00	34.33	24.83	11.50
	Dorso Izquierdo	36.83	23.33	7.50	22.33	19.50	7.50	33.00	25.17	9.33
CASIA	Palma Derecha 850	27.83	11.33	0.33	22.83	14.00	2.33	22.50	13.33	1.67
	Palma Izquierda 850	35.83	8.83	1.17	23.50	14.17	1.17	21.50	11.67	0.83
	Palma Derecha 940	29.83	11.33	0.87	23.17	16.00	3.33	23.83	17.67	3.83
	Palma Izquierda 940	33.83	8.50	0.83	24.17	15.83	1.83	26.33	14.83	0.83

Figura 6.3: Porcentajes de clasificación obtenidos para las 4 bases de datos con Redes Neuronales Artificiales.

En las Figura 6.4, 6.5, 6.6, 6.7, 6.8, 6.9 y 6.10, se muestran los tiempos de cómputo de preprocesamiento de las imágenes para las 4 bases de datos analizadas, además se incluyen los tiempos de cómputo con una selección previa de la Region de Interés y utilizando las imágenes completas. Como se puede observar, los tiempos de cómputo aumentan considerablemente cuando se utilizan las imágenes completas.

Los tiempos de cómputo mostrados a continuación, están divididos por técnica de procesamiento (Clásicas, CLAHE y USM), por cantidad de imágenes (Base de datos completa, 1 imagen y 12 o 6 imágenes) y por región (Palma Derecha e Izquierda, Dorso Derecho e Izquierdo, Muñeca o Antebrazo Derecha e Izquierda). los tiempos de cómputo están expresados en minutos (color azul) y segundos (color rojo).

	Base de datos completa	1 imagen	12 imágenes
Contraste Limitado de la Ecuación Adaptativa del Histograma	11.3026146	0.634037	10.05021
Técnicas Clásicas	10.5156915	0.603619	3.956161
Máscara de desenfoque	12.2811385	0.678002	6.014585

Figura 6.4: Tiempos de cómputo para la base de datos PolyU Multispectral Palmprint.

	Base de Datos Completa	1 imagen	12 imágenes	Región
Contraste Limitado de la Ecuación Adaptativa del	3.247486	1.216872	7.183521	Dorso Derecho
	3.07613422	1.184241	5.606955	Dorso Izquierdo
	3.08068425	1.187461	5.116504	Palma Derecha
	2.93190768	1.197452	5.183548	Palma Izquierda
Técnicas Clásicas	2.7245544	1.158616	7.343136	Dorso Derecho
	2.69078465	1.149106	6.048399	Dorso Izquierdo
	2.72344657	1.112488	5.869355	Palma Derecha
	2.73728323	1.1626	5.252955	Palma Izquierda
Máscara de Desenfocaje	3.66931025	1.224004	7.258489	Dorso Derecho
	3.42730513	1.249363	6.87474	Dorso Izquierdo
	3.69804605	1.211372	7.087465	Palma Derecha
	3.89396155	1.237095	6.534586	Palma Izquierda

Figura 6.5: Tiempos de cómputo para la base de datos UPT con selección de la ROI.

	Base de Datos Completa	1 imagen	12 imágenes	Región
Contraste Limitado de la Ecuación Adaptativa del	13.9875774	2.877933	25.62046	Dorso Derecho
	13.8644154	2.838684	24.98324	Dorso Izquierdo
	14.3348841	2.883779	21.86729	Palma Derecha
	14.0170588	2.88612	22.37476	Palma Izquierda
Técnicas Clásicas	13.1904183	2.783119	23.52641	Dorso Derecho
	13.8358207	2.818754	22.12578	Dorso Izquierdo
	13.7694948	2.805147	20.95049	Palma Derecha
	13.232137	2.785587	21.09617	Palma Izquierda
Máscara de Desenfocaje	20.414585	3.573776	35.97382	Dorso Derecho
	21.1269211	3.588702	37.31709	Dorso Izquierdo
	18.2279426	3.422414	30.4833	Palma Derecha
	18.5408894	3.47805	30.35024	Palma Izquierda

Figura 6.6: Tiempos de cómputo para la base de datos UPT sin selección de la ROI.

	Base de Datos Completa	1 imagen	12 imágenes	Región
Contraste Limitado de la Ecuación Adaptativa del	4.253143	1.520538	6.542815	Palma Izquierda
	4.23819823	1.222766	5.787284	Palma Derecha
	4.22796003	1.136982	5.587842	Dorso Izquierdo
	4.11270312	1.095319	5.685402	Dorso Derecho
Técnicas Clásicas	4.01824992	1.041177	5.251924	Palma Izquierda
	4.17030893	1.104099	5.200339	Palma Derecha
	4.00322133	1.004733	5.602987	Dorso Izquierdo
	4.07637873	1.039979	5.267591	Dorso Derecho
Máscara de Desenfoque	4.83085218	1.219565	6.974946	Palma Izquierda
	4.91387603	1.249042	6.433621	Palma Derecha
	4.92367263	1.185934	6.623758	Dorso Izquierdo
	5.03440125	1.138088	6.322985	Dorso Derecho

Figura 6.7: Tiempos de cómputo para la base de datos Vein Pattern Database con selección de la ROI.

	Base de Datos Completa	1 imagen	12 imágenes	Región
Contraste Limitado de la Ecuación Adaptativa del	18.9140998	2.737844	20.95874	Palma Izquierda
	18.4047585	2.827867	21.46834	Palma Derecha
	19.9779014	2.74101	20.81113	Dorso Izquierdo
	23.0769722	2.762777	21.69622	Dorso Derecho
Técnicas Clásicas	21.525217	2.652534	20.25714	Palma Izquierda
	16.6278024	2.625821	20.01017	Palma Derecha
	16.3042166	2.587121	19.60587	Dorso Izquierdo
	20.5105702	2.58857	19.88301	Dorso Derecho
Máscara de Desenfoque	23.3623758	3.451939	30.14537	Palma Izquierda
	23.2338633	3.504579	30.1	Palma Derecha
	23.3032094	3.308829	29.74049	Dorso Izquierdo
	29.9051738	3.368456	29.0624	Dorso Derecho

Figura 6.8: Tiempos de cómputo para la base de datos Vein Pattern Database sin selección de la ROI.

	Base de Datos Completa	1 imagen	6 imágenes	Región
Contraste Limitado de la	3.76397922	0.988982	2.870548	Palma Izquierda 850
	3.5770055	0.987173	2.76715	Palma Derecha 850
Ecuación Adaptativa del	3.6322562	0.988021	2.76676	Palma Izquierda 940
	3.51267987	0.982832	2.941318	Palma Derecha 940
Técnicas Clásicas	3.40010773	0.892126	2.731087	Palma Izquierda 850
	3.31867537	0.876034	2.523435	Palma Derecha 850
	3.36080612	0.856782	2.598826	Palma Izquierda 940
	3.2261397	0.913495	2.56994	Palma Derecha 940
Máscara de Desenfocaje	3.7102362	1.031276	2.892481	Palma Izquierda 850
	3.5583615	0.940969	2.785258	Palma Derecha 850
	3.43693502	1.018167	2.709354	Palma Izquierda 940
	3.40271968	0.943746	2.815193	Palma Derecha 940

Figura 6.9: Tiempos de cómputo para la base de datos CASIA con selección de la ROI.

	Base de Datos Completa	1 imagen	6 imágenes	Región
Contraste Limitado de la	14.0225486	2.231545	8.273463	Palma Izquierda 850
	13.5392023	2.351333	6.689698	Palma Derecha 850
Ecuación Adaptativa del	14.2073546	2.298544	9.820392	Palma Izquierda 940
	12.4152551	2.048489	9.696946	Palma Derecha 940
Técnicas Clásicas	12.6356724	1.845253	6.701226	Palma Izquierda 850
	8.94444313	2.108369	6.11665	Palma Derecha 850
	8.3525004	1.590554	6.12713	Palma Izquierda 940
	8.56435027	1.613769	6.288656	Palma Derecha 940
Máscara de Desenfocaje	12.8716355	2.174337	14.47916	Palma Izquierda 850
	12.7969604	2.154369	12.30645	Palma Derecha 850
	12.7856588	2.137641	13.07434	Palma Izquierda 940
	15.4779553	2.167048	13.57122	Palma Derecha 940

Figura 6.10: Tiempos de cómputo para la base de datos CASIA sin selección de la ROI.

6.1. Trabajos derivados de la Tesis

1. Raúl Castro-Ortega *et al.* "***Analysis of the palm vein distribution for people recognition***", **Proceedings of SPIE Optical Engineering + Applications, Signal, Image and Data Processing**. Memoria de Congreso. Aceptado para Presentación, San Diego California (EUA) Agosto de 2015.
2. Raúl Castro-Ortega *et al.* "***Análisis de la distribución de venas de la mano para reconocimiento de personas***", Séptimo Encuentro de Investigación UPT". **Memoria de Congreso**. Volumen 1. ISBN: 04-2014-091714443200-102. Septiembre 2014.
3. Raúl Castro-Ortega *et al.* "***Biometric analysis of the palm vein distribution by means two different techniques of feature extraction***", **Proceedings of SPIE Optical Engineering + Applications, Memoria de Congreso**. Vol. 9217, ISBN: 9781628412444. **Signal, Image and Data Processing**. San Diego California (EUA) Agosto de 2014.
4. Raúl Castro-Ortega and Carina Toxqui-Quitl, "***Analysis of palm vein distribution from IR images for people recognition***", Sexto Encuentro de Investigación UPT. Octubre 2013.
5. Raúl Castro-Ortega, Francisco Solís-Villareal and Carina Toxqui-Quitl, "***Biometric analysis of palm vein distribution using neural networks***", Quinto Encuentro de Investigación UPT. Septiembre 2012.