



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA**

RECONOCIMIENTO ROBUSTO DE ROSTROS USANDO IMÁGENES TÉRMICAS

**TESIS PARA OPTAR AL GRADO DE DOCTOR EN
INGENIERÍA ELÉCTRICA**

GABRIEL ENRIQUE HERMOSILLA VIGNEAU

**PROFESOR GUIA:
JAVIER RUIZ DEL SOLAR SAN MARTIN**

**MIEMBROS DE LA COMISION:
SERGIO TORRES INOSTROZA
DOMINGO MERY QUIROZ
JORGE SILVA SANCHEZ**

**SANTIAGO DE CHILE
JUNIO 2012**

Resumen

El principal objetivo de este trabajo de tesis es el estudio del reconocimiento de rostros en el espectro térmico (8-12 μm). Como parte de este estudio se analizan y comparan el desempeño de métodos y algoritmos de reconocimiento de rostros en el espectro térmico. También, se propone una metodología para desarrollar robustamente el reconocimiento de rostros térmicos, y se analiza cómo el reconocimiento de rostros térmicos se ve afectado por los cambios que sufren las imágenes térmicas cuando son capturadas en diferentes periodos de tiempo debido a cambios en el metabolismo y condiciones ambientales.

Los métodos de reconocimiento de rostros fueron seleccionados por el desempeño obtenido en otros estudios comparativos, trabajar en tiempo real, requerir sólo una imagen por sujeto y operar en línea. El estudio analiza 3 métodos basados en calces locales: Histogramas LBP, Histogramas WLD y Descriptores Jets de Gabor, y 2 métodos basados en calces globales: el método SIFT y SURF. Los métodos son comparados utilizando la base de datos creada UCHThermalFace, que permite evaluar los métodos en condiciones reales ya que incluyen variaciones naturales de iluminación, imágenes capturas en interior y exterior, expresiones faciales, pose, accesorios y oclusión.

Se analiza el uso de las características vasculares del rostro, que pueden ser obtenidas utilizando una cámara térmica. Se propone una metodología basada en el uso combinado de características fisiológicas de la red vascular del rostro humano y la intensidad de los píxeles del rostro térmico, metodología llamada *VascularThermalFace*. Además, se aborda un estudio de reconocimiento de rostros cuando existe diferencia temporal en la adquisición de las imágenes térmicas. Se crearon y analizaron 2 bases de datos llamadas UCHThermalTemporalFace y UCHThermalFaceIndoorOutdoor, donde se visualizan variaciones en el metabolismo de los sujetos durante diferentes días.

Las principales conclusiones de este estudio muestran que: (i) el mejor resultado entre velocidad de procesamiento y altas tasas de reconocimiento es alcanzado por el método de Histogramas WLD, (ii) se valida el uso de la metodología *VascularThermalFace*, debido al buen desempeño obtenido por todos los métodos de reconocimiento y sus altas tasas de reconocimiento, (iii) experimentos con imágenes adquiridas en diferentes periodos de tiempo, muestran que el desempeño de los métodos es alto y solamente dependen de una buena configuración de la cámara térmica, (iv) experimentos con imágenes adquiridas en sesiones de interior-exterior y viceversa, muestran que el desempeño es bueno para métodos basados en calces locales y bajo para métodos basados en calces globales. De estos resultados, es posible diseñar un sistema de reconocimiento de rostros térmico que sea eficiente y robusto.

Abstract

The focus of the work presented here is the study the performance of methods and algorithms of face recognition at thermal spectrum (8-12 μm). A new methodology is proposed for the development of a robust thermal face recognition system. Besides, an analysis for variations in the performance of the algorithm when there are changes in the thermal images due to changes in metabolism and environmental conditions is performed.

The analysis methods were selected by considering their performance in former comparative studies, real-time operations, the requirement of just one image per person, and to being fully online. This study analyzed three local-matching methods; histograms of Local Binary Pattern (LBP) features; histograms of Weber Linear Descriptors (WLD); and Gabor Jet Descriptors (GJD), as well as two global image-matching methods based on Scale-Invariant Feature Transform (SIFT) descriptors, and Speeded Up Robust Features (SURF) descriptors. The methods are compared using the UCHThermalFace database, which allows for evaluating the methods in real-world conditions that include natural variations in illumination, indoor/outdoor setup, facial expression, pose, accessories and occlusions.

We have analyzed the use of vascular features of the face, which can be obtained using a thermal camera. We propose a methodology based on physiological features of the vascular network of the human face and the intensity of the pixels of the thermal face, a methodology called *VascularThermalFace*. In addition, we evaluate a study of face recognition when there is a time difference in the thermal image acquisition. Two databases were created and analyzed: UCHThermalFaceIndoorOutdoor and UCHThermalTemporalFace; where metabolic changes in the subjects could be observed across several days.

The main conclusions of this study are: (i) the best trade-off between high recognition rate and fast processing speed is obtained by WLD-based methods, (ii) validation of the *VascularThermalFace* methodology, due to the good performance achieved by all methods of recognition and the high recognition rates, (iii) experiments with images acquired at different periods of time show high levels of performance and only depend on good thermal camera calibrations, (iv) experiments where the test images are acquired in an outdoor setup and the gallery images are acquired in an indoor setup, or vice versa, the performance based on local-matching methods is high and the performance based on global-matching methods is low. From these results, it is possible to design an efficient and robust thermal face recognition system.

A mi familia, novia, amigos y profesores, que participaron y ayudaron en el desarrollo de mi doctorado. En especial a mi gato Miau, quien me acompañó durante la escritura de la tesis y me aportó muchos momentos de felicidad y que hoy me observa desde el cielo...

Tabla de contenidos

1	Introducción	1
1.1	Introducción al tema de tesis	1
1.2	Motivación.....	2
1.3	Objetivos.....	4
1.4	Hipótesis	6
1.5	Contribuciones.....	7
1.5.1	Publicaciones relacionadas con la tesis	8
1.6	Estructura del documento de tesis	9
2	Trabajo Relacionado	10
2.1	Antecedentes generales.....	10
2.2	Descripción de las cámaras térmicas	12
2.3	Cámaras térmicas usadas en este trabajo	16
2.4	Estado del arte en el reconocimiento de rostros térmicos	16
2.5	Descripción de metodologías y algoritmos.	23
2.5.1	Histogramas LBP (Local Binary Patterns).....	24
2.5.2	Descriptores Jets de Gabor	26
2.5.3	Histogramas de WLD.....	27
2.5.4	Scale Invariant Feature Transform (SIFT)	30
2.5.5	Speeded-Up Robust Features (SURF).....	34
3	Estudio Comparativo de Reconocimiento de Rostros usando Imágenes Térmicas....	39
3.1	Introducción.....	39
3.2	Métodos bajo comparación.....	39
3.3	Estudio comparativo usando la base de datos Equinox	41
3.4	Estudio comparativo usando la base de datos térmica UCHThermalFace	46
3.4.1	Descripción de la base de datos UCHThermalFace	46
3.4.2	Descripción de los experimentos.....	48
3.4.3	Resultados de reconocimiento.....	51
3.5	Desempeño computacional.....	61

3.6	Discusión y conclusión del capítulo	63
4	Estudio Comparativo de Reconocimiento de Rostros usando <i>VascularThermalFace</i>	65
4.1	Introducción.....	65
4.2	Algoritmo de extracción de la red vascular (<i>VascularNetwork</i>) [53].....	66
4.3	Análisis de la red vascular y métodos de reconocimiento.....	68
4.3.1	Estudio de la red vascular superficial.....	68
4.3.2	Estudio de la invariancia temporal de la red vascular.....	69
4.3.3	Estudio de los métodos de reconocimiento utilizando la red vascular.....	75
4.3.4	Factores que limitan la metodología <i>VascularNetwork</i>	79
4.4	Metodología propuesta: <i>VascularThermalFace</i>	83
4.5	Análisis de reconocimiento usando la base <i>UCHThermalFace</i>	85
4.6	Discusión y conclusiones del capítulo.....	100
5	Estudio Comparativo de Reconocimiento de Rostros en el Tiempo	102
5.1	Introducción.....	102
5.2	Análisis de reconocimiento usando imágenes térmicas con variabilidad temporal	103
5.2.1	Descripción de la base de datos <i>UCHThermalTemporalFace</i>	103
5.2.2	Descripción de los experimentos.....	106
5.2.3	Resultados de reconocimiento.....	107
5.2.4	Discusión.....	116
5.3	Análisis de reconocimiento usando diferentes galerías con variabilidad temporal.....	117
5.3.1	Descripción de la base de datos <i>UCHThermalFaceIndoorOutdoor</i>	118
5.3.2	Descripción del experimento.....	118
5.3.3	Resultados de reconocimiento.....	120
5.3.4	Discusión.....	124
6	Conclusiones	126
7	Referencias	131

Lista de Figuras

Figura 1. Sujeto capturado bajo distintas iluminaciones en el espectro visible y el espectro térmico.....	3
Figura 2. Imágenes visibles y térmicas de un individuo con distintas expresiones y diferentes iluminaciones.....	4
Figura 3. Espectro Electromagnético.	11
Figura 4. Diferentes aplicaciones para el uso de imágenes térmicas	12
Figura 5. Chip de un Focal Plane Array.....	13
Figura 6. Sistema de captura de un IRFPA	13
Figura 7. Imagen real captada con un IRFPA	14
Figura 8. Cámaras térmicas	16
Figura 9. Individuo capturado con diferentes tipos de iluminación y diferentes expresiones	18
Figura 10. Espectro de imágenes del infrarrojo.....	19
Figura 11. Red vascular del rostro humano.....	21
Figura 12. Imagen térmica de un sujeto y su correspondiente mapa de venas del rostro	22
Figura 13. Esquema del operador LBP	25
Figura 14. Algoritmo de Histograma LBP	25
Figura 15. Descripción del algoritmo WLD.....	28
Figura 16. Histograma WLD.....	30
Figura 17. Construcción de la pirámide Gaussiana y la pirámide de diferencias Gaussianas.....	32
Figura 18. Descriptor SIFT	33
Figura 19. Cálculo con la imagen integral.....	35
Figura 20. Filtros Gaussianos de segundo orden discretizados y las aproximaciones usadas.	36
Figura 21. Filtros Haar	37
Figura 22. Asignación de la orientación.....	37
Figura 23. Descriptor SURF.....	38
Figura 24. Ejemplos de la base de datos Equinox	42
Figura 25. Configuración del experimento.....	48
Figura 26. Ejemplo para la base UCHThermalFace.....	48
Figura 27. Ejemplos de rostros con diferentes tamaños de cortes (base de datos UCHThermalFace).	50
Figura 28. Ejemplos de imágenes con oclusión parcial (base de datos UCHThermalFace).....	50
Figura 29. Ejemplos de imágenes alineadas después de aplicar ruido en la posición de los ojos (base de datos UCHThermalFace).....	51
Figura 30. Ejemplos de calces correctos para algunos sujetos de la base de datos UCHThermalFace.....	54
Figura 31. Resultados de experimentos de ruido en la posición de los ojos (0%, 2,5%, 5%, 10%), base de datos UCHThermalFace.	58

Figura 32. Resultados de experimentos de ruido en la posición de los ojos (2,5%, 5%, 10%), base de datos UCHThermalFace.	58
Figura 33. Ángulo de elevación versus tasa de reconocimiento para los métodos seleccionados ...	61
Figura 34. Ejemplo del proceso de extracción de la red vascular <i>VascularNetwork</i> para un sujeto	68
Figura 35. Aplicación de la metodología <i>VascularNetwork</i> a una imagen térmica de un brazo.	69
Figura 36. Aplicación de la metodología <i>VascularNetwork</i> a una imagen térmica de una mano....	69
Figura 37. Imágenes de red vascular para el sujeto 3 de la base de datos UCHThermalTemporalFace.	71
Figura 38. Imágenes de red vascular para el sujeto 5 de la base de datos UCHThermalTemporalFace..	72
Figura 39. Imágenes de red vascular para el sujeto 3 de la base de datos UCHThermalTemporalFace.	73
Figura 40. Imágenes de red vascular para el sujeto 3 de la base de datos UCHThermalTemporalFace.	73
Figura 41. Imágenes de red vascular para el sujeto 5 de la base de datos UCHThermalTemporalFace.	74
Figura 42. Imágenes de red vascular para el sujeto 5 de la base de datos UCHThermalTemporalFace.	74
Figura 43. Representación del método LBP para la red vascular.....	76
Figura 44. Descriptores obtenidos con el método SIFT	77
Figura 45. Calces obtenidos con el método SIFT.....	77
Figura 46. Calces obtenidos con el método SIFT.....	78
Figura 47. Calces obtenidos con el método SURF.....	78
Figura 48. Ejemplos de imágenes térmicas obtenidas con la configuración de parámetros fijos	80
Figura 49. Ejemplos de imágenes térmicas obtenidas con la configuración de parámetros automáticos.....	81
Figura 50. Imagen térmica de un sujeto con lentes..	82
Figura 51. Imágenes térmicas y de red vascular para un sujeto del conjunto de rotación de la base UCHThermalFace.....	83
Figura 52. Sujeto con el rostro húmedo	83
Figura 53. Pseudo código del algoritmo <i>VascularThermalFace</i>	85
Figura 54. Imágenes de la metodología <i>VascularThermalFace</i>	85
Figura 55. Ejemplos de rostros con diferentes tamaños de cortes para la metodología <i>VascularThermalFace</i> (base de datos UCHThermalFace).....	87
Figura 56. Resultados de los experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace.	87
Figura 57. Resultados de los experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace.	88
Figura 58. Ejemplos de imágenes de oclusión parcial de rostros para la metodología <i>VascularThermalFace</i> (base de datos UCHThermalFace).....	89
Figura 59. Resultados de experimentos de oclusión parcial, base de datos UCHThermalFace... ..	90

Figura 60. Resultados de experimentos de oclusión parcial, base de datos UCHThermalFace...	90
Figura 61. Ejemplos de calces correctos para algunos sujetos de la base de datos UCHThermalFace para la oclusión parcial del rostro térmico.	91
Figura 62. Resultados de experimentos de ruido en la posición de los ojos (2.5%), base de datos UCHThermalFace.....	92
Figura 63. Resultados de experimentos de ruido en la posición de los ojos (5%), base de datos UCHThermalFace.....	92
Figura 64. Resultados de experimentos de ruido en la posición de los ojos (10%), base de datos UCHThermalFace.....	93
Figura 65. Resultados de experimentos de ruido en la posición de los ojos (2.5%), base de datos UCHThermalFace.....	94
Figura 66. Resultados de experimentos de ruido en la posición de los ojos (5%), base de datos UCHThermalFace.....	94
Figura 67. Resultados de experimentos de ruido en la posición de los ojos (10%), base de datos UCHThermalFace.....	95
Figura 68. Ejemplos de calces SIFT.....	96
Figura 69. Resultados de experimentos de galería de interior versus exterior, base de datos UCHThermalFace.....	97
Figura 70. Resultados de experimentos de galería de exterior versus interior, base de datos UCHThermalFace.....	97
Figura 71. Ejemplo de imágenes para la base de datos UCHThermalTemporalFace, tamaño 81x150 píxeles.....	105
Figura 72. Diagrama representativo de las sesiones de captura realizadas en el laboratorio.	105
Figura 73. Ejemplos de imágenes con oclusión parcial (base de datos UCHThermalTemporalFace).....	107
Figura 74. Ruido en la posición de los ojos.	108
Figura 75. Reconocimiento de rostros en el tiempo utilizando el enfoque de rostro térmico.	110
Figura 76. Reconocimiento de rostros en el tiempo utilizando el enfoque de <i>VascularNetwork</i>	110
Figura 77. Reconocimiento de rostros en el tiempo utilizando el enfoque de <i>VascularThermalFace</i>	111
Figura 78. Ejemplos de imágenes de la base de datos UCHThermalFaceIndoorOutdoor.	120

Lista de Tablas

Tabla 1. Descripción de los conjuntos de imágenes Equinox	43
Tabla 2. Resultados para la base de datos Equinox.	45
Tabla 3. Resultados para la base de datos Equinox.	46
Tabla 4. Resultados de experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace.	56
Tabla 5. Resultados de experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace.	57
Tabla 6. Resultados experimentos de oclusión parcial, base de datos UCHThermalFace.	57
Tabla 7. Resultados experimentos de oclusión parcial, base de datos UCHThermalFace.	57
Tabla 8. Resultados de experimentos de galería de interior vs exterior, base de datos UCHThermalFace.	59
Tabla 9. Resultados de experimentos de galería de exterior vs interior, base de datos UCHThermalFace.	59
Tabla 10. Resultados de experimentos de expresiones faciales, base de datos UCHThermalFace.	59
Tabla 11. Resultados de experimentos de distintas distancias, base de datos UCHThermalFace. ...	60
Tabla 12. Tiempo de procesamiento. Tiempo medido en milisegundos.	62
Tabla 13. Resultados de experimentos de expresiones faciales, base de datos UCHThermalFace.	98
Tabla 14. Mejoras obtenidas en los resultados de reconocimiento al aplicar la metodología <i>VascularNetwork</i>	99
Tabla 15. Mejoras obtenidas en los resultados de reconocimiento al aplicar la metodología <i>VascularThermalFace</i>	99
Tabla 16. Resultados de los experimentos de reconocimiento de rostros en el tiempo, base de datos UCHThermalTemporalFace.	112
Tabla 17. Resultados de los experimentos de oclusión parcial del rostro térmico, base de datos UCHThermalTemporalFace.	113
Tabla 18. Resultados de los experimentos de 2.5% de ruido en posición de los ojos del rostro térmico, base de datos UCHThermalTemporalFace.	114
Tabla 19. Resultados de los experimentos de 5% de ruido en posición de los ojos del rostro térmico, base de datos UCHThermalTemporalFace.	115
Tabla 20. Resultados de los experimentos de 10% de ruido en posición de los ojos del rostro térmico, base de datos UCHThermalTemporalFace.	115
Tabla 21. Resultados de los experimentos de 10% de ruido en posición de los ojos con oclusión parcial del rostro térmico, base de datos UCHThermalTemporalFace.	116
Tabla 22. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor.	122

Tabla 23. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor.	122
Tabla 24. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor.	123
Tabla 25. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor.	123

1 Introducción

1.1 Introducción al tema de tesis

El uso de imágenes térmicas en reconocimiento de rostros permite solucionar las limitaciones que poseen los sistemas clásicos basados en imágenes visibles como la dependencia a las condiciones de iluminación y a las expresiones faciales. Debido a las características propias que poseen las imágenes térmicas, lo que significa trabajar en un rango espectral de longitud de onda mayor que el espectro visible, es que este tipo de imágenes poseen menos dependencia a los efectos de iluminación que las imágenes capturadas en el visible, además de ser una técnica de captura no invasiva [1][2].

El objetivo principal de esta tesis es el estudio del reconocimiento de rostros en el espectro infrarrojo (IR), específicamente entre el rango de 8 a 12 μm o en el espectro infrarrojo térmico. Este estudio de reconocimiento de rostros concentra su análisis principalmente en algoritmos de reconocimiento de rostros usados en el espectro visible, los cuales son adaptados para el reconocimiento de rostros en el espectro térmico, además incorpora una nueva metodología de reconocimiento de rostros basada en la extracción de la red vascular del rostro humano y su fusión con el rostro térmico. Los principales métodos a usar en este estudio son algoritmos basados en características locales del rostro, como los métodos de histogramas de LBP [3], histogramas de WLD [4] y los Descriptores Jets de Gabor [5]. Además, son estudiados algoritmos basados en características globales de los rostros como los métodos de SIFT [6] y SURF [7]. La idea principal de este estudio es analizar cada uno de los métodos de reconocimiento de rostros, identificando sus ventajas y desventajas en el ámbito térmico, con el fin de ser una guía para poder desarrollar un sistema que sea capaz de reconocer de manera robusta los rostros de los sujetos sin importar la condición en que es adquirido el rostro, esto significa que el sistema debe ser capaz de manejar diversas situaciones como: condiciones ambientales controladas y no controladas, diferentes tipos de oclusión, distintas condiciones corporales de los individuos, variabilidad temporal, etc. Además, en este estudio de reconocimiento de rostros se analiza el uso de características vasculares del rostro térmico. Las características vasculares pueden ser obtenidas con una cámara térmica y representan la información de las características sanguíneas del rostro tales como la distribución de los capilares, venas y arterias en el rostro humano [8][9]. La idea de usar la representación de la red vascular es debido a que la distribución de las venas y arterias del rostro humano es única y puede ser captada por una cámara térmica, lo cual nos entrega una característica robusta para reconocer individuos. Biológicamente, la cantidad de calor que poseen las venas y arterias en el rostro es producto de la difusión calórica, la cual corresponde a la diferencia de temperaturas que existe por ejemplo entre el tejido y las venas. Esta conducción de calor crea un suave gradiente de

temperatura, el cual puede ser observado por una cámara térmica y puede ser determinado a través de un algoritmo basado en un análisis morfológico de la imagen térmica.

Además, el reconocimiento de rostros humanos en ambientes no controlados ha aumentado el interés de la comunidad científica e investigadores en los últimos años. Muchos estudios muestran que el uso de las imágenes térmicas pueden resolver las limitaciones que posee el reconocimiento de rostros en el espectro visible, tales como invariancia a variaciones en iluminación y robustez a variaciones en la pose, las cuales son dos de los mayores factores que afectan el desempeño de los sistemas de reconocimiento de rostros en ambientes no controlados [1][2][10][11].

Finalmente, este estudio de reconocimiento de rostros aborda un análisis de reconocimiento de rostros con captura de las imágenes en diferentes periodos de tiempo con el fin de validar el desempeño de los métodos de reconocimiento de rostros propuestos en esta tesis cuando existe una variabilidad temporal en la captura de las imágenes, incluyendo factores como efectos del metabolismo en los individuos, cambios de aspectos de los individuos y efectos del cambio de las condiciones ambientales, factores que afectan considerablemente la robustez y el rendimiento de los métodos de reconocimiento. También, el estudio de los métodos de reconocimiento de rostros en el tiempo analiza las mejoras obtenidas de la aplicación de los métodos de reconocimiento de rostros que se basan en la extracción de la red vascular y su fusión con el rostro térmico, metodología propuesta en esta tesis para el reconocimiento robusto de rostros térmicos.

1.2 Motivación

La idea que motiva el desarrollo de esta tesis es solucionar los problemas que existen en los sistemas de reconocimiento de rostros actuales, y también estudiar el espectro infrarrojo térmico identificando las ventajas y desventajas sobre el espectro visible en el ámbito del reconocimiento de rostros.

Un ejemplo simple para ver la motivación del uso de sistemas de reconocimiento de rostros térmicos es realizar el reconocimiento de rostros bajo completa oscuridad o cuando existe mucha iluminación. Con los sistemas actuales de reconocimiento en el espectro visible no es posible reconocer sujetos en ausencia de iluminación o cuando existen problemas de iluminación, en cambio con un sistema que posea algoritmos de reconocimiento de rostros en el espectro térmico, esto es posible debido a la naturaleza y propiedades del espectro térmico, lo cual nos permite captar individuos bajo ausencia de iluminación. Dada estas propiedades del espectro térmico y por ende de las cámaras térmicas, es posible desarrollar sistemas de vigilancia nocturna o sistemas de interacción humano-robot, en los cuales un robot domestico podría trabajar sin iluminación como por ejemplo para detectar y reconocer personas. Un ejemplo de imágenes térmicas captadas en ausencia de iluminación es mostrado en la Figura 1, se aprecia un sujeto el cual es capturado en el espectro visible y el espectro térmico, bajo 3 condiciones distintas de iluminación. Se observa al individuo bajo luz natural tenue, mucha luz artificial y sin iluminación. Las imágenes térmicas del individuo no muestran variaciones ante los cambios de iluminación.

En la actualidad, el desarrollo de sistemas de reconocimiento de rostros se basa principalmente en el estudio del espectro visible. El espectro visible posee diferentes características que lo hacen apropiado para evaluar sistemas de reconocimiento de rostros pero

inadecuado en diversas situaciones. Entre los problemas que posee el reconocimiento de rostros en el espectro visible se encuentran: la dependencia a las condiciones ambientales de iluminación, la dependencia a variaciones de la pose del rostro y a cambios de expresión facial [1][2][10][11]. Por los motivos anteriores y dado las características propias del espectro infrarrojo térmico, donde no existe dependencia de las condiciones ambientales de iluminación; es posible realizar la captura de una imagen térmica sin iluminación; la información de las características térmicas del rostro es única para cada persona, es que en este trabajo se estudian y analizan diversas metodologías de reconocimiento de rostros utilizando el espectro térmico. La Figura 2, muestra como el espectro térmico soluciona las limitaciones de variación de la condición de iluminación y los cambios expresión facial dadas en el espectro visible. Se observa a un individuo con 3 condiciones de iluminación distintas y con 3 expresiones diferentes. Las imágenes térmicas del individuo no son afectadas por la iluminación y se aprecia una leve variación en las expresiones.



Figura 1. Sujeto capturado bajo distintas iluminaciones en el espectro visible y el espectro térmico. Arriba: Imágenes visibles del individuo. Abajo: Imágenes térmicas del individuo. Note que en el caso de la condición de ausencia de iluminación, en la imagen visible no aparece el individuo, mostrando que sería imposible realizar el reconocimiento de rostros. En cambio en las imágenes térmicas se aprecia que no existe variación dada las 3 distintas iluminaciones.

Por otra parte, los sistemas de reconocimiento de rostros que usan imágenes visibles son fáciles de engañar usando elementos que realizan oclusión en algunas partes del rostro como: máscaras, bufandas, sombreros, disfraces, etc. Lo anterior motiva a estudiar diferentes métodos de reconocimiento de rostros con el fin de encontrar un método robusto a estas dificultades

(oclusión, diferentes artefactos), además de aprovechar las ventajas de la utilización del espectro infrarrojo térmico.

Además, el hecho de poder extraer características vasculares del rostro usando las cámaras térmicas, motiva la idea de estudiar los métodos de reconocimiento de rostros en base a estas características y la nueva información que entregan, con el fin de realizar análisis profundos y revisar si el uso de estas características entrega resultados óptimos en las tasas de reconocimiento de rostros del sistema.

Todos estos hechos motivan el desarrollo de un estudio de métodos de reconocimiento de rostros en el espectro térmico, con el fin de poder obtener una metodología que sea robusta, invariante y rápida, y que sirva como guía para la creación de un sistema de reconocimiento de rostros que utilice el espectro térmico.

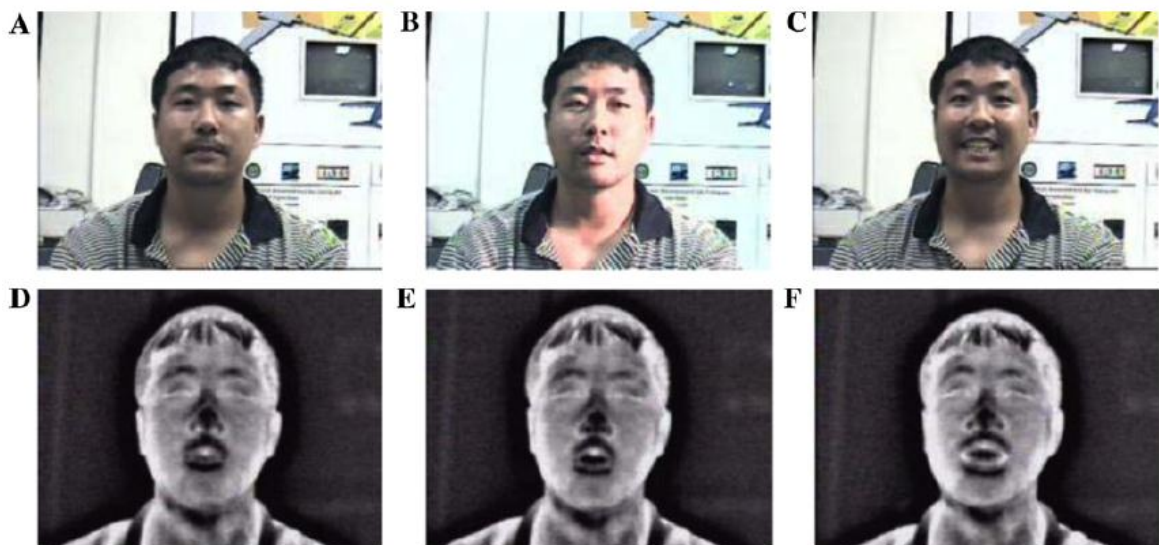


Figura 2. Imágenes visibles y térmicas de un individuo con distintas expresiones y diferentes iluminaciones. A) Imagen con expresión “enojado” y con iluminación artificial; B) Imagen con expresión “sorpresa” y con iluminación artificial lateral; C) Imagen con expresión “alegre” y con iluminación artificial; D) Equivalente térmico de A; E) Equivalente térmico de B; F) Equivalente térmico de C. Note que las imágenes térmicas no son afectadas por los cambios de iluminación y no existe una gran variación con las distintas expresiones (ejemplo imágenes D y E). Imágenes tomadas de [12].

1.3 Objetivos

Los objetivos de la tesis están relacionados con el estudio y análisis del reconocimiento de rostros en el espectro infrarrojo térmico (8-12 μm).

Objetivos generales

- Realizar un estudio sobre el reconocimiento de rostros utilizando el espectro térmico, con el fin de diseñar y construir sistemas de visión computacional que sean capaces de reconocer en forma robusta rostros humanos usando cámaras térmicas en el rango 8-12 μm . El estudio debe ser capaz de analizar limitaciones como expresiones faciales, oclusiones, iluminación y cambios de aspecto de los individuos. El estudio debe entregar como resultado una metodología para los sistemas de reconocimiento de rostros que sea invariante y robusta:

Invariancia:

- Invariancia a expresiones faciales: El sistema de reconocimiento de rostros debe ser capaz de reconocer a un sujeto aun cuando el sujeto cambie la expresión del rostro.
- Invariancia a oclusión: El sistema de reconocimiento de rostros debe ser capaz de reconocer a un individuo aunque el individuo posea algún elemento que realice oclusión de alguna parte del rostro como: bigotes, barba, disfraz, lentes, antifaz, etc.
- Invariancia a la iluminación: El sistema de reconocimiento de rostros debe ser capaz de reconocer a un sujeto en ambientes interiores y exteriores con tasas elevadas de reconocimiento, sin depender de las variaciones de iluminación entregadas por el ambiente.
- Invariancia a cambios de aspecto: El sistema de reconocimiento de rostros debe ser capaz de reconocer a un individuo aun cuando el individuo realice cambios de aspecto y alteraciones del rostro como: maquillaje y cirugía estética.

Robustez:

- Robustez a variaciones intrínsecas: El sistema de reconocimiento de rostros debe alcanzar tasas de reconocimiento elevadas aunque existe una amplia variabilidad de individuos y existan cambios propios en el aspecto de los mismos.
 - Robustez a variaciones extrínsecas: El sistema de reconocimiento de rostros debe alcanzar tasas de reconocimiento elevadas en ambientes dinámicos donde existan variaciones de fondo y de iluminación.
- Diseñar e implementar sistemas de visión computacional robustos que utilicen imágenes térmicas para realizar reconocimiento de rostros.

Objetivos específicos

- Diseñar bases de datos de imágenes térmicas para realizar el estudio de reconocimiento de rostros. Las bases de datos térmicas deben poseer diversas características, tales como: al menos 50 personas, debe contener muestras de los individuos en diferentes sesiones (variabilidad temporal) y debe poseer muestras en ambientes de interior y exterior.
- Seleccionar e implementar métodos de reconocimiento de rostros para realizar un estudio comparativo de estos métodos en el espectro térmico.
- Mejorar la estabilidad del reconocimiento de rostro con imágenes térmicas. Cuando se adquieren imágenes infrarrojas para el conjunto de galería y de prueba con una diferencia de tiempo muy largo (semanas, meses), se ha observado que la intensidad de las imágenes infrarrojas cambia debido a la variabilidad de la emisión térmica producto de este lapso de tiempo [13][14].
- Implementar un método que detecte las características vasculares del rostro en el espectro infrarrojo.
 - Se quiere diseñar un sistema que permita extraer las características vasculares o venas del rostro de una imagen térmica. Para esto se desarrollarán varios módulos de procesamiento, comenzando por realizar una segmentación del rostro térmica, la aplicación de filtrado y difusión térmica para extraer la red vascular.
- Idear una metodología que permita combinar las características vasculares con el rostro térmico con el fin de obtener una nueva representación que utilice la información de la red vascular y el rostro térmico.
- Mejorar problemas de diversa naturaleza que posee el reconocimiento de rostros basado en imágenes térmicas como: problemas basados en el tiempo de adquisición de los conjuntos de galería y prueba para las imágenes infrarrojas, donde existe variabilidad de los patrones térmicos obtenidos en los diferentes tiempos de adquisición, problemas relacionados con las oclusiones, debido a que algunos elementos como los lentes, opacan la radiación infrarroja y problemas relacionados a cambios en las condiciones ambientales, producto de la cantidad de calor ambiental la cual hace que la temperatura corporal del individuo varíe.

1.4 Hipótesis

Las hipótesis que presenta este trabajo de tesis son las siguientes:

1. Es posible obtener una metodología de reconocimiento de rostros con imágenes térmicas que sea invariante a expresiones faciales, oclusiones, iluminación y a cambios de aspectos de un individuo, que funcione adecuadamente con tiempos de procesamiento menores a un segundo y que alcance tasas de reconocimiento altas para bases de datos mayores a 50 individuos, basándose en un estudio comparativo de métodos de reconocimiento de rostros en el espectro térmico.

2. El uso de la red vascular del rostro humano permite entregar la información necesaria para realizar un reconocimiento robusto utilizando imágenes térmicas.
3. La fusión de la red vascular y el rostro térmico permiten a los métodos de reconocimiento de rostro obtener tasas de reconocimiento de rostros más altas que al usar sólo el rostro térmico.

1.5 Contribuciones

En el presente apartado se muestran las contribuciones del trabajo de tesis acerca del reconocimiento robusto de rostros usando imágenes térmicas. Las contribuciones son detalladas a continuación:

Creación de distintas bases de datos térmicas para el estudio de reconocimiento de rostros

- Base de datos UCHThermalFace: Base de datos creada con el fin de realizar un estudio de reconocimiento de rostros en el espectro térmico para ambientes controlados y no controlados. La base de datos posee diversos aspectos como distintas rotaciones y expresiones faciales, y fue capturada bajo condiciones ambientes de interior y exterior.
- Base de datos UCHThermalTemporalFace: Base de datos creada con el fin de realizar un estudio de reconocimiento de rostros en el espectro térmico, donde las imágenes obtenidas poseen distinta variabilidad temporal. La base de datos posee diversos aspectos como diferente tiempo de adquisición de las imágenes (distintas sesiones de captura en el tiempo), cambios de aspecto de individuos, cambio en el metabolismo de los individuos, y diferente condición ambiental donde se tomaba la muestra.
- Base de datos UCHThermalFaceIndoorOutdoor: Base de datos creada con el fin de realizar un estudio de reconocimiento de rostros en el espectro térmico cuando se usa como galería una sesión de interior y para la evaluación o test una sesión de exterior y viceversa. La base de datos posee imágenes capturadas en ambientes de interior y exterior obtenidas en 4 sesiones de tiempo diferente.

Desarrollo de un estudio comparativo de reconocimiento de rostros utilizando imágenes térmicas

- Se realizó un estudio comparativo de reconocimiento de rostros con el fin de obtener una metodología que sea robusta a variaciones intrínsecas y extrínsecas, que sea invariante a expresiones faciales, oclusiones, iluminación y cambios de aspecto para diseñar e implementar sistemas de reconocimiento de rostros robustos basados en visión computacional.

Diseño de una metodología robusta para el reconocimiento de rostros que utiliza la red vascular del rostro humano

- Se diseñó una metodología llamada *VascularThermalFace* que permite mezclar la información vascular del rostro humano con el rostro térmico y tiene por finalidad agregar información al rostro térmico acerca de la posición donde se encuentra su red vascular. Al utilizar esta metodología con un método de reconocimiento de rostros, tal como SIFT, SURF, WLD, etc., es posible mejorar las tasas de reconocimiento que obtienen estos métodos al usar sólo la información del rostro.

Método de reconocimiento WLD (Weber Linear Descriptor) es usado por primera vez en reconocimiento de rostros

- Una contribución interesante de este trabajo de tesis es que se utiliza por primera vez el método de histogramas WLD para realizar el reconocimiento de rostros, más aún, es primera vez que este método de reconocimiento es usado en el reconocimiento de rostros con imágenes térmicas.

Método de reconocimiento SURF es usado por primera vez en reconocimiento de rostros térmicos

- Se utilizó el método SURF para el reconocimiento de rostros utilizando imágenes térmicas.

Análisis de invariancia temporal para el reconocimiento de rostros con imágenes térmicas

- Una contribución importante es el análisis de reconocimiento de rostros usando imágenes térmicas con variabilidad temporal, con el fin de estudiar las limitantes del espectro térmico como cambios de aspecto de individuos, cambio en el metabolismo de los individuos, y diferente condición ambiental donde se tomaba la muestra. Este trabajo muestra que estas limitantes pueden ser solucionadas por medio de algunas metodologías de reconocimiento como es el caso del método de histogramas de WLD.

1.5.1 Publicaciones relacionadas con la tesis

La publicación principal, que incluye el desarrollo de la tesis de doctorado y corresponde a una publicación ISI, es mostrada a continuación [15]:

- G. Hermosilla, J. Ruiz-del-Solar, R. Verschae, and M. Correa, “A Comparative Study of Thermal Face Recognition Methods in Unconstrained Environments”, *Pattern Recognition*, vol. 45, no. 7, pp. 2445-2459, 2012.

Otras publicaciones relacionadas con la tesis y que tratan de estudios acerca de sistemas de reconocimiento de rostros utilizando imágenes térmicas son [16] y [17]. En [16] se realiza un estudio comparativo de métodos de reconocimiento de rostros en el espectro visible utilizando una base de datos de imágenes térmicas llamada Equinox. Esta tesis muestra una extensión de ese trabajo debido a que en [16], no se usan condiciones reales como ambientes en exterior, oclusiones, rotaciones, etc. En [17], se enfoca principalmente en la utilización del método SIFT [6] para realizar el reconocimiento de rostros con imágenes térmicas usando 2 enfoques: usando

el método original de SIFT [6] y usando un método SIFT con muchas etapas de rechazo probabilístico [18]. Además esta publicación es la primera que utiliza la red vascular y el método SIFT para el desarrollo de sistemas de reconocimiento de rostros térmicos.

Otra publicación relacionada con la tesis es [19]. En esta publicación se realiza un sistema de reconocimiento de rostros utilizando imágenes térmicas e imágenes visibles en un robot doméstico para realizar interacción de humano y robot. Se analiza el desempeño de reconocimiento e identificación de personas, realizado por el robot por medio de un esquema de decisión en base a la fusión las imágenes visibles y térmicas obtenidas de las diferentes cámaras.

1.6 Estructura del documento de tesis

En el Capítulo 2 se detalla el trabajo previo realizado por distintos autores en el ámbito del reconocimiento de rostros usando imágenes térmicas. Se muestra el estado del arte actual en el reconocimiento de rostros utilizando el espectro térmico y los trabajos más significativos en el campo de la visión computacional. Además, en el Capítulo 2 se muestran las metodologías a utilizar en este trabajo, donde se describen los algoritmos de reconocimiento de rostros usados en este trabajo. También, en el Capítulo 2 se describen las especificaciones de las cámaras térmicas usadas a lo largo de este trabajo de tesis.

En el Capítulo 3 se muestra el estudio comparativo de reconocimiento de rostros en ambientes controlados y no controlados, donde se realiza un análisis comparativo de varios métodos de reconocimiento de rostros, seleccionados e implementados por el autor. Se realiza el estudio utilizando 2 bases de datos, una base estándar de imágenes infrarrojas térmicas llamada Equinox y una base de datos de imágenes térmicas llamada UCHThermalFace, diseñada en el proceso de esta tesis, en la Universidad de Chile. Se muestra una serie de experimentos basados en diferentes aspectos que afectan a los sistemas de reconocimiento de rostros actuales. Se analizan los diferentes métodos y se obtienen resultados que permiten escoger los mejores métodos de reconocimiento de rostros en el espectro térmico.

En el Capítulo 4 se presenta una nueva metodología de reconocimiento de rostros llamado *VascularThermalFace* basado en la información vascular del rostro humano. La nueva metodología *VascularThermalFace* consiste en usar la red vascular del rostro y superponerlo con el rostro térmico, con el fin de agregar información al rostro térmico. En este capítulo se comparan 2 enfoques basados en la red vascular (*VascularNetwork*) y la metodología propuesta *VascularThermalFace* con los algoritmos de reconocimiento de rostros utilizados en el Capítulo 3, realizando los mismos experimentos de ese capítulo. Se analizan las ventajas y desventajas de usar esta nueva técnica y se muestra con se extrae la información vascular del rostro térmico.

En el Capítulo 5 se realiza un estudio de reconocimiento de rostros con imágenes térmicas en el tiempo, capturadas durante más de 2 meses. Se analizan las dificultades que posee el reconocimiento de rostros térmico cuando las sesiones son en diferente periodo de tiempo y cuando las imágenes térmicas son capturadas bajo diferentes sesiones de interior y de exterior. En el Capítulo 5 se capturaron muestras de los sujetos durante 70 días aproximadamente y se utilizan 4 sesiones de interior y exterior, creándose 2 nuevas bases de datos de imágenes térmicas llamadas UCHThermalTemporalFace y UCHThermalFaceIndoorOutdoor. Finalmente, las conclusiones y proyecciones de la tesis son presentadas en el Capítulo 6.

2 Trabajo Relacionado

2.1 Antecedentes generales

Los trabajos del astrónomo inglés Sir Frederick William Herschel fueron los primeros que ayudaron a demostrar la existencia de la radiación infrarroja [20]. Herschel descubrió la radiación infrarroja en el año 1800, cuando realizaba experimentos de descomposición de la luz solar mediante de un prisma a través del cual medía la temperatura de los diferentes colores de la luz visible para encontrar una relación entre la cantidad de calor y la radiación solar. Herschel se dio cuenta que más allá del rango visible, en particular por sobre el color rojo, existía otra forma de radiación, no visible pero que desprendía calor. Herschel denominó a su descubrimiento rayos caloríficos, los que más tarde se conocerían como irradiancia o radiación infrarroja.

El descubrimiento de la radiación infrarroja o radiación térmica se debe principalmente a la cualidad de los cuerpos de irradiar energía en longitudes de onda perteneciente al espectro infrarrojo por sobre una temperatura distinta al cero absoluto (0 K). Como todos los objetos en la tierra están por sobre el cero absoluto, todos los objetos emiten radiación infrarroja.

El espectro infrarrojo forma parte del espectro electromagnético y específicamente se encuentra localizado entre el espectro de las microondas y el espectro de la luz visible (Figura 3). Abarca el rango de longitudes de onda entre los $0.75 \mu\text{m}$ y los $1000 \mu\text{m}$ aproximadamente. Este espectro infrarrojo se divide en tres zonas, las cuales corresponden al infrarrojo cercano ($0.75 \mu\text{m}$ - $1.4 \mu\text{m}$), el infrarrojo mediano ($1.4 \mu\text{m}$ - $8 \mu\text{m}$) y el infrarrojo lejano ($8 \mu\text{m}$ - $1000 \mu\text{m}$) [21]. Además el infrarrojo mediano y lejano se subdivide en nuevas zonas dependiendo de sus diferentes aplicaciones. En la zona del infrarrojo mediano y lejano la emisión infrarroja es producto de calor o radiación térmica y se conoce como termografía. Las imágenes captadas en ese rango se llaman termogramas.

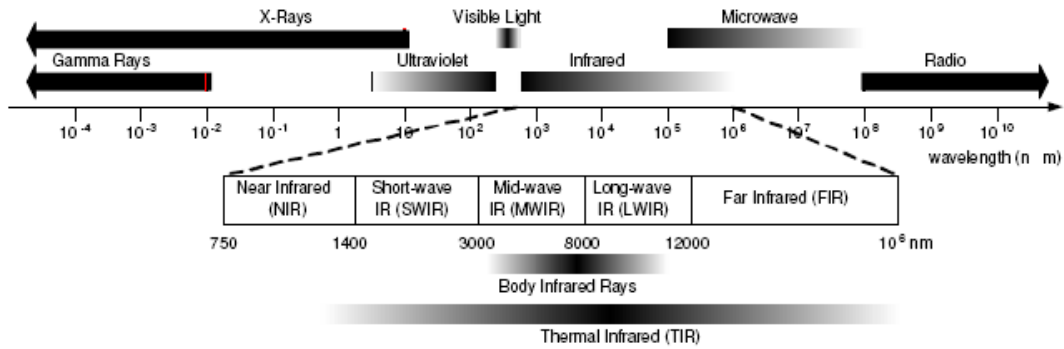


Figura 3. Espectro Electromagnético.

En la actualidad, diversas aplicaciones han aparecido como resultados de la utilización de la radiación infrarroja. En astronomía se están realizando nuevos descubrimientos en el universo con la utilización de esta radiación. En medicina, se utiliza como herramienta de diagnóstico, para detectar cáncer de mamás, desordenes hormonales, problemas a la tiroides, etc. Las cámaras infrarrojas son utilizadas en una gran variedad de aplicaciones como actividades policiales, vigilancia nocturna, percepción remota, aplicaciones militares, combate de incendios, protección de cultivo de salmones, detección de fiebre porcina en aeropuertos, en veterinaria, etc. [22]. Las imágenes infrarrojas se utilizan para detectar pérdidas de calor en edificios y probar sistemas electrónicos. Los satélites infrarrojos monitorean el clima terrestre, estudian modelos de vegetación, llevan a cabo estudios geológicos y miden las temperaturas oceánicas. Se utilizan en ingeniería para procesos industriales, en mantenimiento predictivo, análisis de circuitos integrados, análisis de distribución de calderas y hornos, etc. Finalmente nuevas aplicaciones están apareciendo en detección y reconocimiento de rostros, detección de personas y objetos calientes. En la Figura 4, se muestran imágenes infrarrojas de diversas aplicaciones.

El uso de esta tecnología ha aumentado en los últimos años teniendo muchas aplicaciones como las explicadas en el párrafo anterior. Esto se debe a las características que posee frente a la radiación visible como poder detectar características no observables en el espectro visible, no requerir fuentes de luz tales como sol o luz artificial, operar en la noche y en oscuridad, no poseer efectos adversos como sombra o cambios de luz [23][24]. Estas y otras características hacen que el espectro térmico sea un campo que se está explorando actualmente.

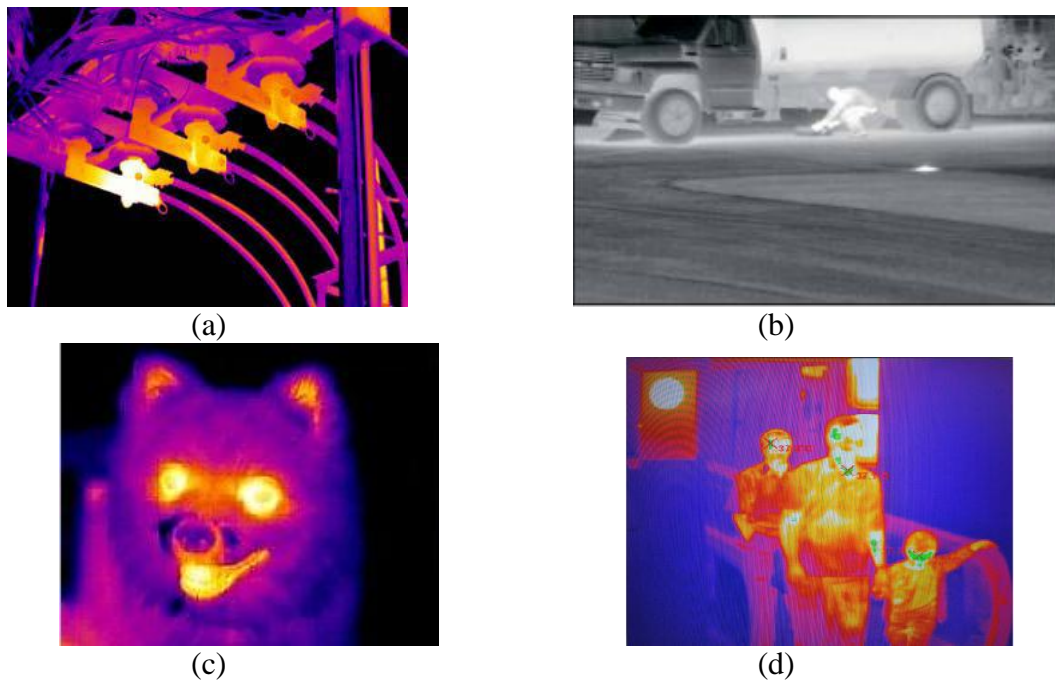


Figura 4. Diferentes aplicaciones para el uso de imágenes térmicas. (a) Sistema de distribución de potencia; (b) revisión de explosivos; (c) veterinaria; (d) detección de fiebre porcina en aeropuertos.

2.2 Descripción de las cámaras térmicas

Las cámaras térmicas permiten detectar la radiación infrarroja que emiten los cuerpos y que es invisible al ojo humano. El desarrollo de los sensores térmicos permite obtener una representación de la radiación térmica, convirtiéndola en una imagen dentro del espectro visible, la cual es vista como distintas intensidades de una escala de grises o de colores. El desarrollo de los detectores infrarrojos se ha realizado por medio de diferentes tecnologías, donde la más usada actualmente es la tecnología de detección térmica por medio de los microbolómetros [25]. La detección térmica absorbe la energía de la radiación incidente, la cual causa que la temperatura del material aumente por los cambios físicos del material. La temperatura relativa es medida eléctricamente para determinar el poder óptico en el detector. La detección térmica funciona convirtiendo el flujo de fotones en calor [25], el calor causa que la temperatura del detector aumente, lo que es detectado por un microbolómetro. El microbolómetro es un elemento que opera captando cambio en su resistencia cuando su temperatura es cambiada. Luego un circuito es usado para convertir una lectura en una señal de salida. La operación del microbolómetros es limitada por el requerimiento del tiempo de respuesta del detector, la cual debe ser compatible con la tasa de cuadros por segundos del sistema de imágenes. Los detectores térmicos pueden operar a temperatura ambiente, son más baratos que otras tecnologías pero poseen una modesta sensibilidad y una baja velocidad de respuesta.

Con el paso del tiempo los detectores infrarrojos individuales se han convertido en arreglos completos de matrices de detectores infrarrojos en un plano focal. Esta tecnología, comúnmente denominada *Focal Plane Array* (FPA) [26][27], consiste básicamente en un arreglo matricial de

detectores individuales que se ubican en el plano focal de un sistema óptico y que se construye sobre un chip, como se muestra en la Figura 5.

Cada pixel del *focal plane array* genera una señal que es relativa al flujo total por unidad de área (irradiancia). El nivel de gris del píxel depende del flujo de radiación por unidad de área y la eficiencia cuántica detectada, la cual es definida como la transferencia de razón señal a ruido, que depende de área de cada elemento del detector, del coeficiente de absorción y de su eficiencia de conversión [24].

Cuando se habla particularmente de un *Infrared Focal Plane Array* o IRFPA, se habla de un sistema completo de detección infrarroja, el cual se compone de una etapa de focalización óptica, luego de una matriz compuesta por los detectores infrarrojos, después una etapa electrónica en donde se realiza la amplificación de la señal y la conversión análogo/digital para luego pasar por la etapa del procesamiento y obtener así la interfaz visual del sistema, la Figura 6 muestra un sistema de captura de un IRFPA.

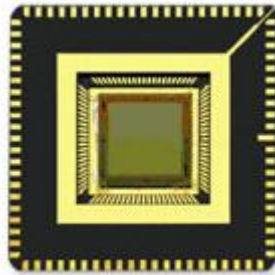


Figura 5. Chip de un Focal Plane Array. Tomado de [28].

Cada detector en un IRFPA realiza la conversión de la radiación infrarroja incidente a una señal eléctrica, el cual luego de las etapas electrónicas (amplificación y conversión A/D) entrega como resultado cada uno de los pixeles que conforman la imagen o frame de salida del sistema. Una imagen o frame se denomina al conjunto de valores obtenidos en un determinado tiempo al pasar por el sistema completo del IRFPA. Estos frames son entregados por el IRFPA a una tasa de varios frames por segundos por lo que al analizar los datos es conveniente realizar un esquema tridimensional con lo cual es posible observar de una mejor forma el resultado del IRFPA.

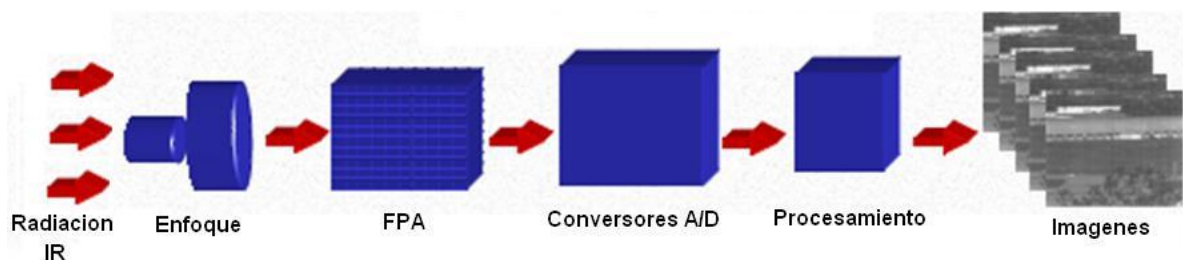


Figura 6. Sistema de captura de un IRFPA. Tomada de [29].

Todos los detectores del arreglo convierten la radiación infrarroja incidente en señales eléctricas las que más tarde a través de procesos electrónicos dan como resultado cada uno de los píxel que conforman la imagen total entregada por el sistema. Se esperaría que esta imagen sea

una imagen uniforme como las imágenes captadas por las cámaras que trabajan en el espectro visible. Esta situación no ocurre producto que cada uno de los detectores del arreglo funciona en forma independiente lo que genera diferencias en la fotorespuesta que producen los detectores del arreglo. Esto quiere decir que frente a un mismo nivel de incidencia de la radiación infrarroja en todo el plano focal se obtiene una respuesta no uniforme del sistema de imágenes. A este problema se le conoce como no-uniformidad de las imágenes infrarrojas (*Non-Uniformity*) [30][31] y se observa en las imágenes infrarrojas como una interferencia o ruido de patrón fijo (*Fixed Pattern Noise, FPN*) [32] que aparece sobre las imágenes como una rejilla o como si la imagen hubiese sido tomada con el lente sucio de la cámara. La Figura 7 muestra una imagen real captada por una cámara infrarroja que posee esta no uniformidad o ruido de patrón fijo en forma de enrejado, en la Figura 7 se aprecia como varía la no uniformidad en cada pixel de la imagen.

Además de la no uniformidad de las imágenes infrarrojas existe otro problema que poseen los detectores del arreglo. Este problema radica en que los detectores en conjunto con su electrónica asociada son altamente sensibles a la temperatura por ser construidos por materiales semiconductores, esto se ve reflejado en la variación de sus parámetros o *drift* por lo que es necesaria una calibración de los detectores para poder entregar la correcta información que reciben de la radiación de los cuerpos. Este problema del *drift* que poseen los detectores implica que la no uniformidad va cambiando constantemente en las imágenes obtenidas por lo que hace el problema de corrección de la no uniformidad resulta aún más difícil y más interesante.

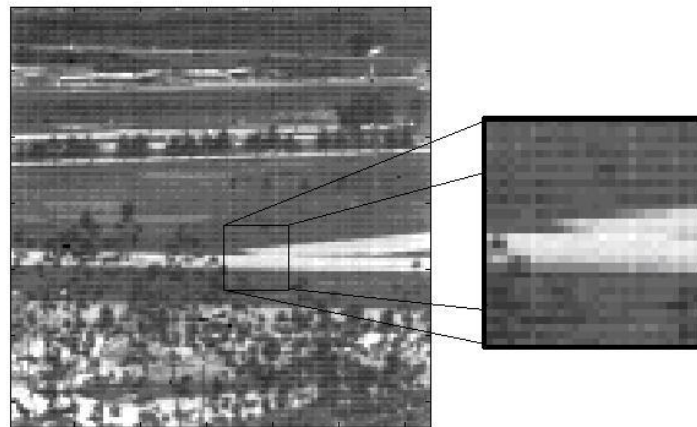


Figura 7. Imagen real captada con un IRFPA. Se observa el ruido de patrón fijo en forma de enrejado. Modificada de [33].

Para solucionar el problema de no uniformidad en imágenes infrarrojas muchos métodos se han creado e investigado en los últimos años [30][31][34]. Los métodos para solucionar este problema se dividen en dos categorías, las cuales son de calibración [35] y escena [36][37][33].

Por temas de costos y tiempo de procesamiento, los métodos más usados y actualmente investigados corresponden a los de escena. En este trabajo se nombran algunos de los métodos más conocidos y que son aplicados en la eliminación de no uniformidad. Estos métodos son: Estadísticas constantes [38][39], Scribner [40], RLS (*recursive least square*) [41], etc.

Algunos de estos métodos fueron analizados en la tesis con el fin de realizar un preprocesamiento de las imágenes térmicas. Sin embargo, es importante mencionar que la cámara

térmica usada para desarrollar este trabajo, es nueva (posee 2 años de fabricación) y cuenta con un módulo de corrección de no uniformidad incorporado. Esto hecho ayuda mucho a la hora de adquirir las imágenes térmicas, ya que utilizando este módulo interno, el problema de la corrección de no uniformidad es solucionado.

Hoy en día el precio de las cámaras térmicas ha decrecido significativamente y la tecnología infrarroja a mejorado, obteniendo mejor resolución y calidad, y el ruido de patrón fijo producido en las cámaras térmicas más antiguas ha sido eliminado usando algoritmos de técnicas de corrección de no uniformidad (NUC) incorporados dentro de las mismas cámaras térmicas. Así, el interés en el uso de la tecnología infrarroja, específicamente térmica, se ha incrementado en los últimos años. Sin embargo, aún existen algunas limitantes, dado que las imágenes de rostros térmicos tienen variaciones no deseables debido a: cambios en la temperatura ambiente; modificaciones en los procesos metabólicos de los sujetos; susceptibilidad de la cámara con factores extrínsecos tales como el viento, y la respuesta variable del sensor en el tiempo cuando la cámara está trabajando en largos periodos de tiempo [8][42][13].

2.3 Cámaras térmicas usadas en este trabajo

Las especificaciones de las cámaras térmicas usadas para realizar este trabajo de tesis y obtener las diferentes bases de datos para realizar el reconocimiento de rostros son mostradas a continuación:

- CEDIP JADE UC Infrared System [43] (Figura 8).

Rango espectral:	8 - 14 μm
Resolución:	320x240 pixeles
Sensibilidad térmica:	80 mK
Frame rate:	50-60 frames/s
Salida digital:	14 bits
Sensor:	Microbolómetro sin refrigeración (uncooled)
Precio:	40000 USD

- FLIR 320 TAU [44] (Figura 8)

Rango espectral:	7.5 – 13.5 μm
Resolución:	324x256 pixeles
Sensibilidad térmica:	75 mK
Frame rate:	30 frames/s
Salida digital:	14 bits
Sensor:	Microbolómetro sin refrigeración (uncooled)
Precio:	8000 USD



Figura 8. Cámaras térmicas: (izquierda) CEDIP JADE UC Infrared System, (derecha) FLIR 320 TAU.

2.4 Estado del arte en el reconocimiento de rostros térmicos

Con el paso de los años, diferentes metodologías se han creado para obtener la mayor información de la imagen infrarroja captada, nuevas técnicas de mejora de la imagen y nuevos algoritmos de procesamiento digital de imágenes, los cuales se han incluido en la visión computacional para realizar diferentes tareas relacionadas como reconocimiento y clasificación.

El reconocimiento de rostros es una manera natural de reconocer a las personas, por este motivo ha sido estudiado de diferentes maneras en el espectro visible. Sin embargo, los métodos de reconocimiento de rostros en el espectro visible poseen diferentes problemas cuando se intenta identificar un individuo: dependencia de las condiciones de iluminación [1], variación de la pose, cambios de expresión facial y oclusión del rostro [45]. En la actualidad se han desarrollado métodos que compensan parcialmente estos problemas de condición de iluminación [46], sin embargo la compensación de iluminación no funciona adecuadamente cuando los sistemas de reconocimiento de rostros se encuentran ubicados en condiciones no controladas o al aire libre (exterior) [47]. Para solucionar esta clase de inconvenientes, se ha ido desarrollando el reconocimiento de rostros utilizando imágenes térmicas, debido a que las imágenes infrarrojas térmicas no son dependientes a las condiciones de iluminación [1][2][10][11] y por lo tanto problemas de este tipo no existen. En la Figura 9, se puede apreciar una imagen de un individuo fotografiado con diferentes tipos de iluminación y diferentes expresiones, en ambos espectros, en el visible y en el infrarrojo térmico. Se puede observar de la Figura 9 que las imágenes infrarrojas del rostro del individuo son invariantes a cambios de iluminación y sufren pequeñas variaciones ante cambios relacionados con la pose y expresión del sujeto (similar al ejemplo mostrado en el Capítulo 1, Figura 2).

Antiguamente los sistemas de reconocimiento de rostros eran usados principalmente en el espectro visible y no en el espectro infrarrojo. La causa fundamental de esta situación era que las cámaras infrarrojas eran más costosas que las cámaras convencionales, por lo que su fabricación era principalmente para desarrollar aplicaciones militares y médicas [2][10]. Otras causas que contribuían a que las cámaras infrarrojas no fueran ocupadas para los sistemas de reconocimiento de rostros eran que las imágenes infrarrojas poseían baja resolución espacial y además las imágenes infrarrojas poseían un ruido llamado ruido de patrón fijo [38][39], el cual era inherente en la tecnología infrarroja y contribuía a factores negativos para los sistemas de reconocimiento de rostros.



Figura 9. Individuo capturado con diferentes tipos de iluminación y diferentes expresiones. Tomada de [2].

Con el paso del tiempo y debido a los avances tecnológicos en el espectro infrarrojo, es que hoy en día es posible considerar el uso de la tecnología infrarroja para el reconocimiento de rostros. Diversos factores que se mencionaron anteriormente han sido mejorados con el paso de los años, superando las limitaciones de resolución y ruido inherente que poseía la tecnología infrarroja. Esta mejora tecnológica trajo como consecuencia una extensión al desarrollo de los sistemas de reconocimiento de rostros del visible al infrarrojo.

Desde el punto de vista del espectro infrarrojo, es posible obtener diversos tipos de imágenes infrarrojas dependiendo de la zona o rango del infrarrojo en que trabaje la cámara infrarroja. Un ejemplo de las distintas imágenes generadas en el espectro infrarrojo es mostrado en la Figura 10. En la Figura 10 se puede observar el espectro de imágenes en el infrarrojo, el cual está dividido en 3 zonas conocidas como el infrarrojo cercano ($0.75 \mu\text{m}-1.4 \mu\text{m}$), infrarrojo mediano ($1.4 \mu\text{m}-8 \mu\text{m}$) y el infrarrojo lejano ($8 \mu\text{m}-12 \mu\text{m}$). Actualmente este es el rango de operación en el cual trabajan las diferentes cámaras infrarrojas.

Las principales ventajas de usar el infrarrojo térmico son mostradas a continuación:

- Los sensores infrarrojos recogen la energía térmica emitida por un cuerpo en vez de la luz reflejada [48].
- Los sensores infrarrojos exhiben un comportamiento invariante a cambios de iluminación, incluso operando en completa oscuridad.
- La piel humana tiene una alta emisividad contenida entre los $8-12 \mu\text{m}$, presentando una naturaleza térmica [48].

- La información anatómica y vascular del rostro humano que entregan los sensores infrarrojos es única para cada persona [42].

El espectro infrarrojo posee diferentes dificultades de diversa naturaleza, estos problemas son mencionados a continuación:

- Ruido de patrón fijo: Este problema es producto de la tecnología infrarroja, específicamente los sensores térmicos. Al adquirir la radiación infrarroja, cada uno de los detectores del arreglo capta una porción no uniforme de la radiación infrarroja, por lo que la imagen obtenida se aprecia como si estuviera con el lente sucio o con un ruido llamado ruido de patrón fijo. Para solucionar este problema de eliminación de ruido, existen métodos llamados métodos de corrección de no uniformidad (NUC).
- Oclusión: este problema se debe principalmente al uso de elementos que opacan la radiación infrarroja, principalmente lentes y anteojos [1][2][10].
- Diferencia temporal en adquisición de imágenes de galería y de prueba: este problema ocurre producto del lapso de tiempo en el cual se adquieren imágenes infrarrojas para el conjunto de galería y conjunto de prueba. Si la diferencia de tiempo es muy larga (semanas, meses), se ha observado en [13][14][8][42] que las imágenes infrarrojas de los individuos cambian debido a la variabilidad de la emisión térmica producto de este lapso de tiempo. Este lapso de tiempo cambia significativamente los patrones térmicos de los individuos, lo cual trae como consecuencia una disminución en el desempeño del reconocimiento de rostros.

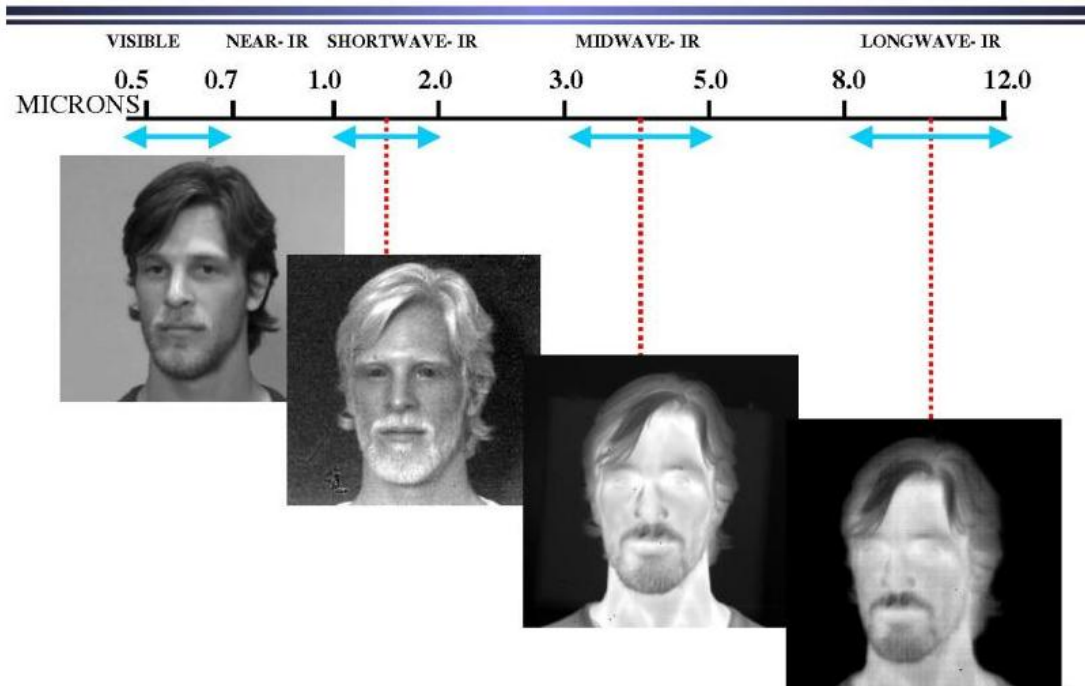


Figura 10. Espectro de imágenes del infrarrojo. Se observa la evolución de las imágenes de un sujeto desde el espectro visible (0.4-0.75 μm) hasta el infrarrojo lejano o térmico (8-12 μm). Tomada de [49].

En términos de algoritmos y métodos de reconocimientos, existen pocas diferencias en los algoritmos usados en el espectro visible y en el espectro infrarrojo térmico. Esto implica que algoritmos que son usados en el visible pueden ser usados en el espectro infrarrojo. El problema de utilizar métodos del espectro visible en el espectro infrarrojo es que estos métodos no garantizan la obtención de buenos resultados de desempeño para los sistemas de reconocimiento basados en los rostros térmicos. Esto se puede ver en que los algoritmos usados en espectro visible son usados para compensar las variaciones de iluminación que posee el espectro visible y para el espectro infrarrojo no es necesario realizar esta compensación debido a que las imágenes térmicas ya son invariantes a esta clase de fenómenos. Esta discusión se realiza producto que en la actualidad los algoritmos ocupados para el reconocimiento de rostros térmicos son casi los mismos usados en el espectro visible, donde no se realiza la distinción se usar métodos diferentes propios para el espectro infrarrojo. Por lo tanto, se deberían idear algoritmos basados en características térmicas para realizar el reconocimiento de rostros.

Muchos estudios comparativos acerca del reconocimiento de rostros en el espectro infrarrojo térmico han sido desarrollados en los últimos años [1][2][10]. Mucho de los enfoques desarrollados hacen uso de métodos basados en apariencia, tales como análisis de componentes de principales (PCA), análisis de discriminante lineal (LDA), y análisis de componentes independientes (ICA), los cuales proyectan las imágenes de los rostros a un sub-espacio donde el reconocimiento es llevado a cabo. Estos métodos alcanzan un ~95% en la tasa de reconocimiento pero no consideran condiciones reales (ambientes no controlados) [1][2][10] o cuando usan la base de datos Equinox [50]. La base de datos Equinox es de hecho una base de datos térmica estándar en el reconocimiento de rostros térmico. La base considera 91 individuos, capturados con 3 diferentes expresiones y 3 diferentes condiciones de iluminación (detalles en el Capítulo 3).

Otros enfoques reportados en el reconocimiento de rostros térmico son basados en el uso de calces locales: patrones locales binarios (LBP) [48] y Descriptores Jets de Gabor (GJD) [51][16]. Con la base de datos térmica Equinox se ha reportado que la tasa de reconocimiento es de ~97% para el caso del método LBP [48] y ~80% para el caso de GJD [51][16]. Metodologías basadas en calce global como SIFT (Scale Invariant Features Transform), han sido usadas también en el reconocimiento de rostros térmico [16][17]. Este enfoque es basado en el uso de descriptores de características locales que son invariantes a rotación, traslación y cambios de escala. Estos descriptores locales son usados para calces de pares de imágenes que consideran restricciones geométricas y probabilísticas. En [16], la metodología SIFT es usada para obtener descriptores directamente de la imagen térmica, mientras que en [17] se calculan los descriptores de imágenes vasculares generadas por procesamiento de las imágenes térmicas. Estos enfoques obtienen tasas de reconocimiento que dependen fuertemente de la base de datos usada; ~80% cuando se usa Equinox [16] y ~95% cuando se usa una base no pública [17].

Una metodología propia del espectro infrarrojo, que no puede ser usado en el visible, se basa en el estudio de las características vasculares del rostro humano, específicamente de la red vascular del rostro humano. La principal razón de la utilización de este método, que usa la información de las venas del rostro, es que resuelve problemas de invariancia a la iluminación y a expresiones faciales, debido a que las venas representan una estructura vascular única del rostro humano. De esta manera teniendo la arquitectura vascular del rostro se puede identificar a un sujeto.

En [8][9][52][53] describen la posibilidad de usar la extracción de características del rostro humano basándose en las características fisiológicas que posee el rostro. En la Figura 11 se puede apreciar una imagen del sistema vascular del rostro humano, se puede ver la configuración de la red vascular de la red de arterias, red de venas y el conjunto de venas y arterias. La idea que plantean en [8][9][52][53] es usar un enfoque basado en el uso de rostros térmicos, donde se aplican algoritmos de segmentación a cada rostro térmico para separar lo que es rostro de lo que no es, luego, usando algoritmos morfológicos, ellos extraen el mapa de las venas de los rostros térmicos. La red vascular extraída del rostro produce una forma de contorno que es única para cada individuo, además una mayor cantidad de venas siempre aparece en las mismas áreas, y su posición exacta puede determinar significativamente a un individuo de otro. Un ejemplo de un sujeto y su mapa de venas, tomado de [53], es mostrado en la Figura 12. La red vascular es una importante característica debido a que como está debajo de la piel es muy difícil de alterar. En los trabajos [8][9][52][53] muestran que la red vascular es procesada para obtener los puntos de minucias térmicas. Estos puntos de minucias térmicas son similares a los puntos de minucia producidos por las huellas dactilares. Luego, se presenta la idea de realizar calces con respecto a una base de datos de minucias térmicas para determinar la identidad del individuo. En [54] es presentado un enfoque eficiente para la extracción de características fisiológicas del rostro térmico. Las características representan la red vascular del rostro bajo la piel de la cara. Esta red vascular es única para cada individuo, y puede ser usada para sistemas de reconocimiento de rostros térmico. En [55] es presentado un enfoque similar basado en la huella facial térmica. Este enfoque usa un nuevo conjunto de características para representar el rostro: los puntos de bifurcación de los patrones térmicos y el centro de gravedad en la región de la cara térmica. Los estudios actuales no consideran métodos basados en información vascular, porque ellos aún necesitan ser mejorados para alcanzar el mismo desempeño que los métodos basados en calces locales.

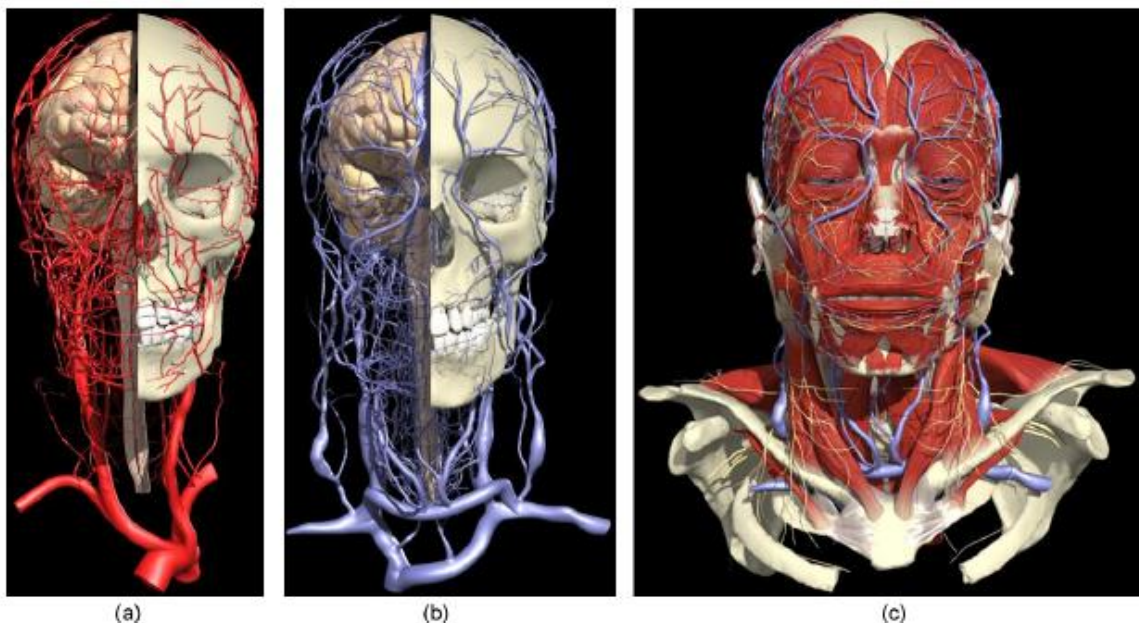


Figura 11. Red vascular del rostro humano. a) red arterial b) red de venas c) mapa superficial de venas y arterias de un rostro. Figura tomada de [8].



Figura 12. Imagen térmica de un sujeto y su correspondiente mapa de venas del rostro. Imagen tomada de [8].

Además, en [56][57][58] son propuestas metodologías basadas en la fusión de imágenes obtenidas del espectro visible y el espectro infrarrojo térmico. En [56][57] métodos estándar basados en apariencia son usados en conjunto con algoritmos genéticos para el análisis de fusión de datos visibles y térmicos. Los métodos alcanzan una tasa de reconocimiento de ~96% en la base de datos Equinox. En [59] son aplicados dos esquemas de fusión: fusión de datos y decisión de la fusión. El algoritmo es diseñado para detectar y reemplazar lentes con una plantilla de ojos en el caso de las imágenes térmicas. Esto es producto que los lentes opacan la radiación infrarroja, generando un hueco sin información. Un método comercial llamado *FaceIt* es usado para la evaluación del algoritmo de fusión. En [60] es descrito un método de fusión de imágenes visibles y térmicas usando la transformada discreta Wavelet (DWT). Los resultados de los experimentos muestran que el método de fusión es efectivo en términos de calidad visual comparado con los métodos convencionales de fusión. En [51], filtros Gabor son usados para extraer características faciales en el dominio visible y térmico. En [61] son analizados diferentes niveles de fusión entre datos visibles y térmicos. En [58] son analizadas las ventajas de combinar las imágenes en los espectros visibles y térmicos en sistemas de reconocimiento de rostros, el reconocimiento es realizado usando un clasificador de k-vecinos más cercanos.

Aunque, está fuera de la mira de este trabajo, vale la pena mencionar que las imágenes de rostros térmicas han sido usadas para el reconocimiento de expresión de rostros [62][63][64][65]. Alguna de estas metodologías de reconocimiento de rostros bajo el análisis de la tesis podría ser adaptada para el reconocimiento de expresión de rostros. En [66], las características son extraídas de la imagen térmica y la imagen visible y son usadas como entradas a una red neuronal que está a cargo del reconocimiento de expresión facial. En [65], la diferencia de temperatura local causada por el reajuste de los músculos faciales y cambios en la temperatura interior es usada como datos de entrada a una red neuronal usada para el reconocimiento de expresión facial. En [63] es propuesta la extracción no supervisada de características locales y globales, la clasificación es realizada usando maquinas de soporte vectorial (SVM). En [62] la temperatura de la piel del rostro es usada para reconocer y clasificar expresiones positivas y negativas. Los resultados de la clasificación sugieren que la temperatura de la piel del rostro puede ser usada para ayudar a distinguir entre expresiones positivas y negativas y puede asistir en la interpretación de estados afectivos. En [64] es usada una transformada de Wavelet para analizar multi-escala y cambios de multi-dirección de las texturas térmicas consideradas como características de textura en las imágenes.

En [16] el autor presenta un estudio comparativo preliminar de reconocimiento de rostros térmicos que no consideran condiciones de mundo real (sólo se usa la base de datos Equinox). Además no se usa el método WLD [4] recientemente propuesto. Este trabajo muestra una extensión de ese trabajo y soluciona todas esas limitaciones.

2.5 Descripción de metodologías y algoritmos.

Este apartado se concentra en el estudio de distintos métodos de reconocimiento que cumplen los siguientes requerimientos:

- Operación en línea: No existen etapas de enrolamiento fuera de línea (offline) y todos los procesos deben correr en línea (online). El sistema debe permitir la construcción de una base de datos que pueda ir incrementándose en el tiempo.
- Operación en tiempo real: El proceso de reconocimiento debe ser suficientemente rápido para permitir la interacción en tiempo real, como en el caso de aplicaciones de interacción humano-robot (Human-Robot Interaction, HRI) y como en el caso de búsquedas en grandes bases de datos en un tiempo razonable (pocos milisegundos o algunos segundos dependiendo de la aplicación y tamaño de las bases de datos).
- Una sola imagen por persona: Una imagen térmica por cada individuo debe ser suficiente para la identificación. La principal razón de esto es para compensar los costos computacionales y la posibilidad de no poder obtener más de una imagen por individuo en ciertas situaciones.
- Ambientes no controlados: No hay restricciones en cuanto a las condiciones del ambiente, tales como iluminación, conjuntos interiores o exteriores, expresiones faciales, escalas, poses, resolución, accesorios, oclusiones o fondo.

Este estudio utiliza 3 métodos basados en calces locales y 2 métodos basados en calces globales que son seleccionados considerando los requerimientos anteriores y debido a su buen desempeño en estudios comparativos de reconocimiento de rostros [67][46][68][5]. Dos métodos basados en calces locales, llamados histograma de características de LBP (Local Binary Pattern) [3] y los descriptores Jets de Gabor con clasificación de cuenta de Borda [5] son seleccionados por su desempeño en los estudios reportados en [46][5]. Un tercer método basado en calces locales, llamado histograma de características WLD (Weber Linear Descriptor), propuesto recientemente en [4], ha mostrado un buen desempeño en aplicaciones de detección de rostros, y es usado por primera vez en reconocimiento de rostros. El método de calce global SIFT (Scale-Invariant Feature Transform) [6], es incluido siguiendo su buen desempeño en el estudio de reconocimiento de rostros en [46]. Finalmente, el método basado en calces globales SURF (Speeded Up Robust Features) [7], cual es inspirado en el método SIFT, es incluido debido a que en muchas aplicaciones en las cuales el método SIFT es usado, SURF obtiene similar desempeño con mayor velocidad.

En resumen, los métodos usados en este trabajo son los siguientes algoritmos:

- Histogramas de patrones locales binarios (Histogramas LBP)

- Descriptores Jets de Gabor (GJD)
- Histogramas de WLD (Weber Linear Descriptor)
- Descriptores SIFT
- Descriptores SURF

2.5.1 Histogramas LBP (Local Binary Patterns)

El operador LBP es uno de los mejores descriptores de textura y ha sido utilizado en varias aplicaciones, primeramente fue diseñado para la descripción de texturas y luego se adaptó al reconocimiento de rostros [3]. Este operador posee ventajas como su invariancia a cambios monótonos en niveles de grises y además su eficiencia computacional, lo cual es importante en procesamiento de imágenes.

Principalmente la idea de usar el operador LBP para el reconocimiento de rostros es motivada en que los rostros pueden ser vistos como una descomposición de micropatrones los cuales pueden ser descritos por este operador.

Algoritmo

El operador LBP asigna una etiqueta (número) a cada pixel de la imagen comparando el pixel central con sus vecinos de una ventana de 3x3 generando como resultado un código binario. Así el histograma obtenido de las etiquetas puede ser usado como un descriptor de textura. El esquema clásico de la descripción del operador LBP es mostrado en la Figura 13.

Una extensión del método LBP llamada “patrones uniformes” es usada en este trabajo. A las etiquetas resultantes del operador LBP se le aplica la detección de patrones uniformes, la cual consiste en encontrar patrones binarios que contengan como máximo 2 transiciones de 0 y 1 o viceversa. Ejemplos de patrones uniformes son 00000000 (0 transiciones), 01110000 (2 transiciones) y 11001111 (2 transiciones). Sin embargo, los patrones 11001001 (4 transiciones) y 01010011 (6 transiciones) no son patrones uniformes. Luego estos patrones uniformes son usados de tal manera de construir un histograma, donde cada bin del histograma es llenado con un patrón uniforme y todos los patrones no uniformes son asignados en un sólo bin.

El histograma LBP contiene información acerca de la distribución de los micropatrones locales del rostro, tales como líneas, puntos y diferentes áreas del rostro. Para una representación eficiente del rostro, se debe mantener también la información espacial del rostro. Por este motivo es que la imagen del rostro es dividida en diferentes regiones rectangulares donde es calculado el histograma sobre los códigos LBP obtenidos en cada región. Finalmente, los histogramas de cada región son concatenados en un sólo histograma que representa la imagen del rostro. La Figura 14 muestra la descripción completa del LBP para un rostro.

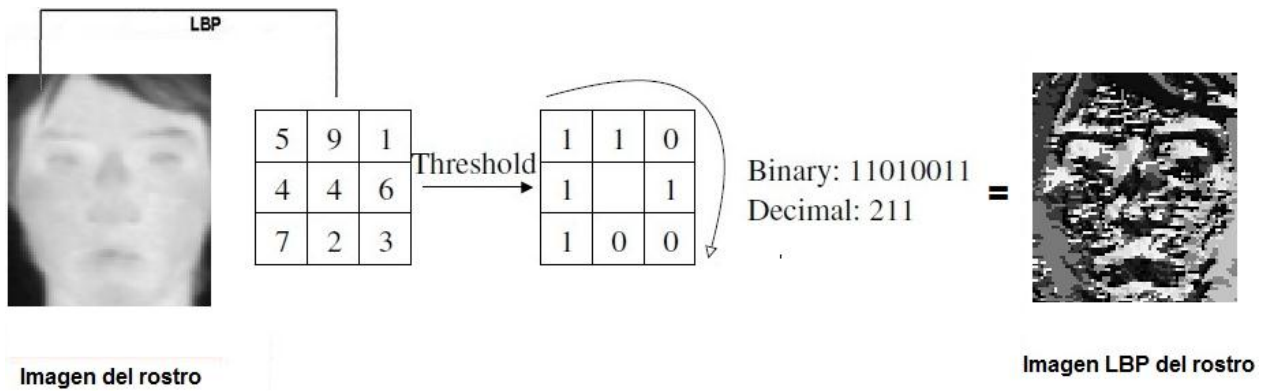


Figura 13. Esquema del operador LBP

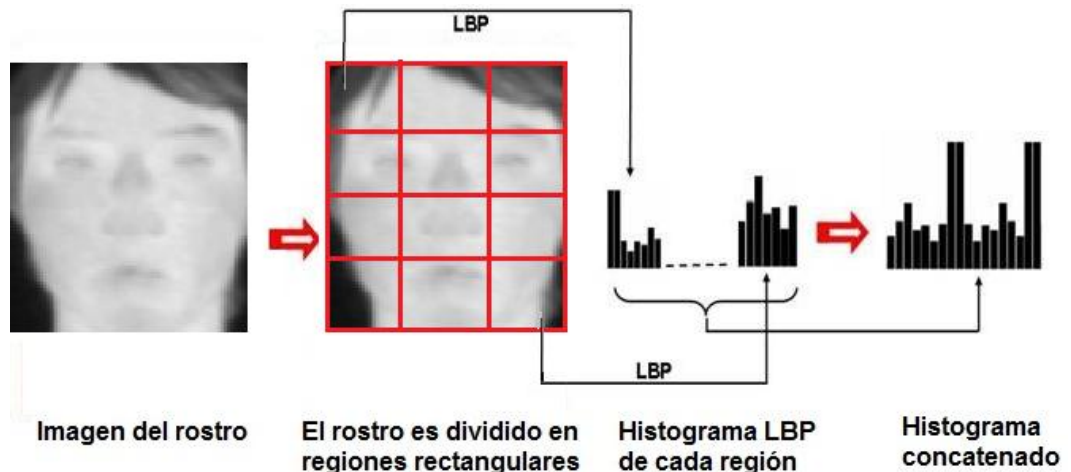


Figura 14. Algoritmo de Histograma LBP

Con este histograma, se tiene una descripción del rostro en base a 3 diferentes niveles de localidad. Las etiquetas del histograma contienen información acerca de los patrones pixel a pixel, las etiquetas entregan información sobre pequeñas regiones del rostro y los histogramas son concatenados construyendo una descripción global del rostro.

El reconocimiento es realizado usando un clasificador de vecinos más cercanos en el espacio de características. Se usan 3 medidas de disimilitud:

- Intersección de Histograma:

$$D(S, M) = \sum_i \min(S_i, M_i) \quad (1)$$

- Logaritmo de la verosimilitud estadística:

$$L(S, M) = -\sum_i S_i \log M_i \quad (2)$$

- Estadística Chi cuadrado (χ^2):

$$\chi^2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \quad (3)$$

2.5.2 Descriptores Jets de Gabor

Los filtros Gabor son ampliamente utilizados en el reconocimiento de rostros debido a que son comparables con la respuesta de los campos receptivos de las células simples de la corteza visual y poseen propiedades computacionales como la robustez a cambios de iluminación [5].

Los filtros Gabor son sinusoides espaciales localizadas por una ventana Gaussiana que permiten la extracción de características en las imágenes por medio de la selección de su frecuencia, orientación y escala.

Algoritmo

Un filtro Gabor se define matemáticamente como:

$$\psi_{\theta, \lambda}(z) = \frac{\|k_{\theta, \lambda}\|^2}{\sigma^2} e^{-\frac{\|k_{\theta, \lambda}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\theta, \lambda}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (4)$$

donde θ y λ se definen como la orientación y escala del filtro de Gabor respectivamente, $z = (x, y)$, $\| \cdot \|$ representa el operador norma, y el vector de onda $k_{\theta, \lambda} = \frac{k_{\max}}{\lambda} e^{i\theta}$, donde k_{\max} representa la frecuencia máxima. La escala es dada por $\lambda = 4\sqrt{2^n}$ con $n = 0, \dots, 4$ y la orientación es dada por $\theta = n\pi/8$ con $n = 0, \dots, 7$.

En la representación de características de Gabor, sólo las magnitudes de Gabor son usadas porque la fase del filtro cambia linealmente con pequeños desplazamientos. Distintos filtros de Gabor son construidos en base a distintos tamaños de σ . Utilizando $\sigma = \lambda$, se calculan diferentes tamaños de las ventanas para el filtro de Gabor, se usa $6\sigma + 1$ desde el centro de la ventana.

Se define un Gabor jet como un conjunto de 8 filtros de Gabor con la misma escala λ y posición (x, y) pero con distinta orientación θ . Cada Gabor jet es localizado uniformemente sobre la imagen, separado por una longitud de onda λ .

La idea principal en la extracción de características usando Gabor jet es construir una grilla en la imagen espaciada en λ . De esta forma se pueden calcular los filtros de Gabor centrados en cada nodo de la grilla y así obtener el jet de manera eficiente y rápida debido a que no se usa la imagen completa, sólo los puntos de la grilla. Cada jet es representado por 8 componentes en cada nodo. La cantidad de jet obtenidos en la imagen depende del tamaño de la región a usar, por

ejemplo si se usa un área de 37×37 , se encuentran 110 jets: 64 jets para $\lambda = 4$, 25 jets para $\lambda = 4\sqrt{2}$, 16 jets para $\lambda = 8$, 4 jets para $\lambda = 8\sqrt{2}$ y 1 jet para $\lambda = 16$.

Cada uno de los jets obtenidos anteriormente es comparado con todos los candidatos usando el producto interno normalizado y los resultados son combinados usando el método de recuento de Borda [5]. La idea de usar el recuento de Borda es obtener un puntaje final para usarlo como una medida de distancia. Si se quiere obtener cual es la imagen de la base de datos más parecida a una imagen de test, se compara el jet de cada nodo de la grilla de la imagen de test con el mismo nodo en cada una de las imágenes de la base de datos usando el producto interno normalizado. Usando el recuento de Borda se genera un ranking para cada uno de los nodos de la grilla. Finalmente se suman los resultados obtenidos en cada imagen de la base de datos y se obtiene un puntaje el cual es tomado como una medida de distancia. De esta forma es llevado a cabo el reconocimiento.

2.5.3 Histogramas de WLD

El método Weber Local Descriptor (WLD), es un simple, poderoso y robusto descriptor, el cual está basado en la ley de Weber [4]. Se basa en el hecho que la percepción humana de los patrones no depende solamente del cambio en el estímulo (tal como sonido o iluminación) sino también de la intensidad original del estímulo. Específicamente, WLD consiste en 2 componentes: excitación diferencial y orientación. La componente de excitación diferencial es una función del radio entre 2 términos: el primero es la diferencia relativa de intensidad del actual pixel contra sus vecinos y el otro es la intensidad del propio pixel actual. El componente de orientación es la orientación del gradiente del pixel actual. Para una imagen dada, se usan las 2 componentes del método WLD para construir y concatenar el histograma WLD.

Algoritmo

El algoritmo WLD utiliza el cálculo de 2 componentes: excitación diferencial y orientación. La Figura 15 muestra el cálculo del descriptor WLD.

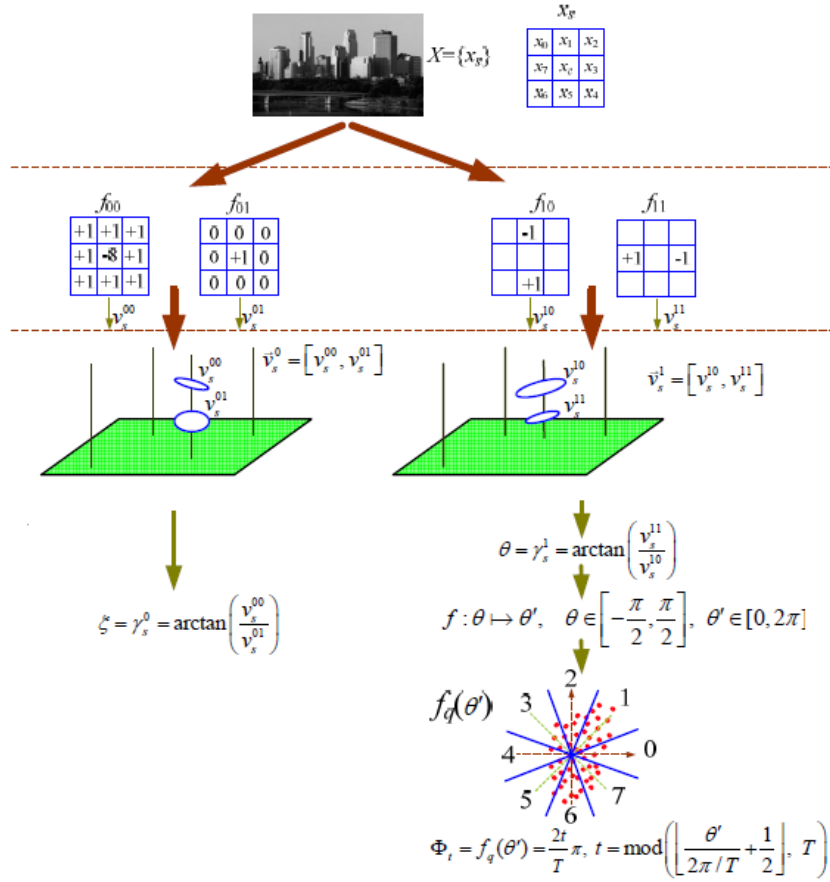


Figura 15. Descripción del algoritmo WLD. Tomado de [4].

Excitación diferencial: Se utilizan las diferencias entre el pixel actual y los vecinos como además los cambios del propio pixel actual. Esto significa que se espera encontrar las variaciones salientes dentro de la imagen con el fin de simular los patrones de percepción que siente el humano.

Así la excitación diferencial $\xi(x)$ se calcula como:

$$\xi(x_c) = \arctan\left[\sum_{i=0}^{p-1}\left(\frac{x_i - x_c}{x_c}\right)\right] \quad (5)$$

donde $x_i (i = 0, 1, \dots, p - 1)$ representa el vecino i del pixel central x_c y p representa el número total de vecinos.

Orientación: El componente orientación de WLD es la orientación del gradiente, la cual es calculada como:

$$\theta(x_c) = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \quad (6)$$

donde $v_s^{10} = x_5 - x_1$ y $v_s^{11} = x_7 - x_3$.

Por simplicidad, la orientación θ se cuantiza en T orientaciones dominantes. Antes de la cuantización, se realiza un mapeo de $\theta \in [-\pi/2, \pi/2]$ a $\theta' \in [0, 2\pi]$, de la siguiente forma:

$$\theta' = \arctan 2(v_s^{11}, v_s^{10}) + \pi \quad (7)$$

donde

$$\arctan 2(v_s^{11}, v_s^{10}) = \begin{cases} \theta, & v_s^{11} > 0 \text{ y } v_s^{10} > 0 \\ \pi - \theta, & v_s^{11} > 0 \text{ y } v_s^{10} < 0 \\ \theta - \pi, & v_s^{11} < 0 \text{ y } v_s^{10} > 0 \\ -\theta, & v_s^{11} < 0 \text{ y } v_s^{10} < 0 \end{cases} \quad (8)$$

De esta forma la función de cuantización es representada como:

$$\Phi_t = \frac{2t}{T} \pi, \quad t = \text{mod} \left(\left[\frac{\theta'}{2\pi/T} + \frac{1}{2} \right], T \right) \quad (9)$$

Histograma WLD: El cálculo del histograma es realizado como lo muestra la Figura 16. El histograma 2D $\{WLD(\xi_j, \Phi_t)\}$, con $j = 0, 1, \dots, N-1$, $t = 0, 1, \dots, T-1$, donde N es la dimensión de la imagen y T es el número de orientaciones dominantes, es construido colocando en cada columna la orientación dominante y en cada fila el histograma de la excitación diferencial con C bins.

Para obtener un descriptor más discriminativo, el histograma 2D $\{WLD(\xi_j, \Phi_t)\}$ es convertido a un histograma 1D H . Tal como muestra la Figura 16, cada columna del histograma 2D es proyectada a un histograma $H(t)$, $t = 0, 1, \dots, T-1$. Se reagrupan las excitaciones diferenciales ξ_j en T sub-histogramas $H(t)$, correspondiendo a cada una orientación dominante. Cada sub-histograma $H(t)$ es dividido en M segmentos $H_{m,t}$ para formar una matriz de histogramas. Cada columna de la matriz de histogramas corresponde a la orientación dominante y cada fila al segmento de excitación diferencial. Luego la matriz de histogramas es reorganizada como un histograma 1D H . Cada fila de la matriz es concatenada como un sub-histograma donde luego son concatenados para originar el histograma 1D $H = \{H_m\}$, $m = 0, 1, \dots, M-1$.

Note que cada sub-histograma $H(t)$ es eventualmente dividido en M segmentos, el rango de la excitación diferencial ξ_j también es dividido en M intervalos. Además cada segmento de sub-histograma $H_{m,t}$ es compuesto por S bins, es decir $H_{m,t} = \{h_{m,t,s}\}$, $s = 0, 1, \dots, S-1$. El número de celdas C en cada columna del histograma 2D $\{WLD(\xi_j, \Phi_t)\}$ es dado por $C = M \times S$.

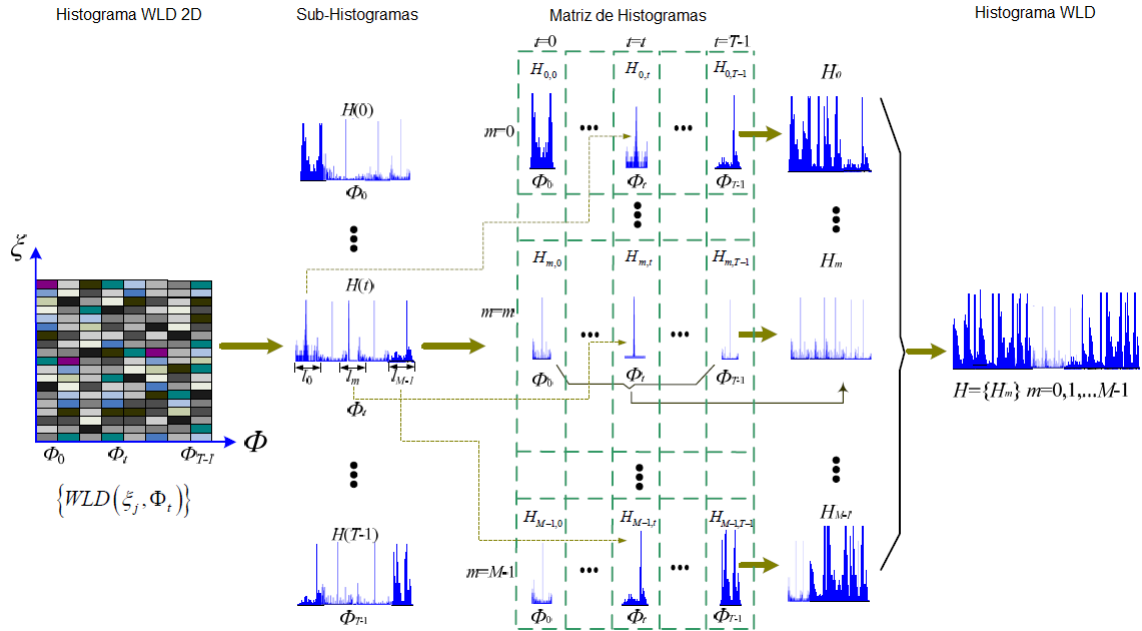


Figura 16. Histograma WLD. Tomado de [4].

El reconocimiento es realizado de la misma forma que en el caso del método LBP. Se usa un clasificador de vecinos más cercanos en el espacio de características. Se usan 3 medidas de disimilitud, ecuaciones (1), (2) y (3), mostradas anteriormente:

- Intersección de Histograma:

$$D(S, M) = \sum_i \min(S_i, M_i) \quad (1)$$

- Logaritmo de la verosimilitud estadística:

$$L(S, M) = -\sum_i S_i \log M_i \quad (2)$$

- Estadística Chi cuadrado (χ^2):

$$\chi^2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \quad (3)$$

2.5.4 Scale Invariant Feature Transform (SIFT)

La metodología SIFT [6] es una técnica que nos permite extraer un conjunto de vectores de características invariantes a escala y rotación de las imágenes, llamados descriptores locales. Cada uno de estos descriptores tiene asociado una posición, orientación y escala. Los descriptores son usados para ver cuán parecidas pueden ser distintas imágenes. El reconocimiento SIFT hace coincidir los descriptores individuales de una imagen de prueba con las de una base de datos

conocida, usando algoritmos de vecinos más cercanos para encontrar correspondencias entre los descriptores de dos imágenes, aplicando la transformada de Hough para identificar grupos de correspondencias que presenten posiciones, rotaciones y escalas coherentes y encontrando una transformación de coordenadas que relacione las dos imágenes a través del algoritmo de mínimos cuadrados. El enfoque de SIFT permite identificar objetos que se encuentren escalados, rotados y con oclusión, además de funcionar en tiempo real.

Algoritmo

Detección de puntos de interés usando el escala-espacio (Scale Space Detection): Se trata de identificar posibles puntos de interés usando una cascada de filtros. La idea de la detección de puntos de interés es encontrar las localizaciones y escalas que son repetibles bajo distintos puntos de vista en el mismo objeto, es decir se trata de encontrar las localizaciones que son invariantes a cambios de escala.

Se define el escala-espacio (Scale Space) como una función $L(x, y, \sigma)$, que es producida de la convolución de una variable de escala Gaussiana $G(x, y, \sigma)$, con una imagen $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (10)$$

donde $*$ es el operador convolución y $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$

La detección de la localización de puntos de interés puede ser eficientemente implementada usando una función de diferencia de Gaussiana (DoG):

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (11)$$

donde k es un factor multiplicativo.

La Figura 17 muestra la construcción de la diferencia de Gaussiana. La imagen inicial es convolucionada con el filtro Gaussiano con un factor $\sigma = \sqrt{2}$. La imagen Gaussiana resultante de la convolución se reduce en un factor de 2 es a 3. Luego este proceso se repite para la imagen Gaussiana hasta que la imagen sea pequeña. De este modo se genera una pirámide Gaussiana. Para cada tamaño de imagen, se calcula la diferencia entre la imagen antes y después de ser convolucionada. Al realizar esto para cada tamaño de imagen se genera la pirámide de diferencias de Gaussiana (DoG).

En cada imagen de la pirámide de DoG se buscan máximos y mínimos locales. Estos puntos se obtienen al comparar cada pixel de la imagen con sus 8 vecinos en la imagen actual y 9 vecinos en la imagen anterior y posterior con distinto σ .

Localización del punto de interés: Para localizar los puntos de interés, se realiza una interpolación parabólica entre cada punto de los máximos obtenidos y sus vecinos en las distintas escalas superior e inferior. Luego se elige el máximo de la parábola para la localización del punto.

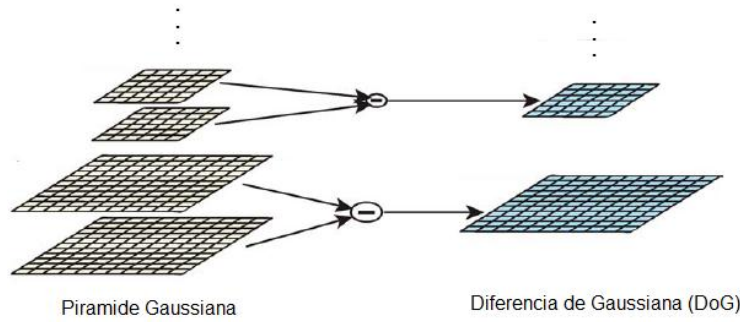


Figura 17. Construcción de la pirámide Gaussiana y la pirámide de diferencias Gaussianas. Modificado de [6].

Se eliminan los puntos que tengan un valor menor a 0.03 en el máximo (umbral) y además se eliminan los puntos que tengan vecinos mayores a 0.95 veces el valor central en un radio de 5 píxeles (sólo en la imagen actual).

Asignación de orientación: Una o más orientaciones son asignadas a cada punto de interés basada en la dirección del gradiente local de la imagen. Para cada imagen Gaussiana suavizada $L(x, y, \sigma)$, la magnitud del gradiente $m(x, y)$ y la orientación $\theta(x, y)$, es calculada usando:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (12)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (13)$$

El histograma de la orientación está formado por 36 bins cubriendo los 360 grados del rango de orientación. Cada muestra agregada al histograma tiene un peso que es ponderado por la magnitud del gradiente y por la distancia al centro de la región Gaussiana.

Las celdas con más votación en el histograma de orientación corresponden a la dirección dominante del gradiente local. Las celdas más votadas y cualquier celda sobre 80% de la votación del máximo son usadas para crear el punto de interés con esa orientación.

Descriptor local de la imagen: Las operaciones previas muestran como asignar a una imagen la localización, escala y orientación para cada punto de interés. Estos parámetros son repetibles en un sistema de coordenadas 2D y describen la región local de la imagen y además son invariantes a estos parámetros. Ahora es necesario calcular un descriptor local que sea altamente distintivo como invariante a otras variaciones tales como cambios de iluminación o distintas vistas 3D.

La Figura 18 muestra el esquema del descriptor para un punto de interés. El descriptor del punto de interés se crea calculando la magnitud del gradiente y las orientaciones de cada punto en la imagen en una región alrededor de la localización del punto de interés. Una ventana Gaussiana es usada para asignar un peso a la magnitud en cada punto, el cual es ilustrado en la Figura 18 como una ventana circular. Esta ventana previene cambios súbitos en el descriptor, debido a que da menos énfasis a los gradientes alejados del centro del descriptor. Los puntos son acumulados en un histograma de orientación, resumiéndolos en 4x4 sub-regiones, donde el largo de cada flecha corresponde a la suma de las magnitudes del gradiente dentro de la región. Note que la

Figura 18 muestra un arreglo del descriptor de 2×2 , sin embargo, el descriptor SIFT usado en este trabajo posee 4×4 regiones de histogramas con 8 orientaciones. Así el descriptor usado posee $4 \times 4 \times 8 = 128$ elementos para cada punto de interés.

Finalmente, el vector de características es modificado para reducir los efectos del cambio de iluminación, por lo que el vector de características es normalizado a la unidad. Además se reduce la influencia de magnitudes muy grandes para el gradiente, limitando los valores del vector de características normalizado a no más de 0.2 (luego se vuelve a normalizar a la unidad). Esto significa el vector de características posee más énfasis la distribución de orientaciones que en gradientes muy grandes.

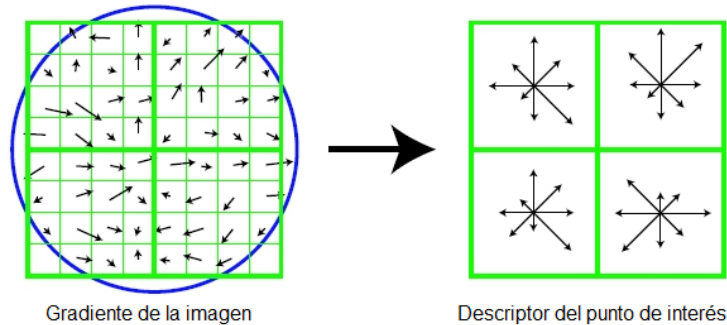


Figura 18. Descriptor SIFT. Tomado de [6].

Calces de los puntos de interés (Matching): El mejor candidato al calce para cada punto de interés se encuentra identificando el vecino más cercano de los puntos de interés de una base de datos de entrenamiento. El vecino más cercano es definido como el punto de interés que posee la mínima distancia euclidiana para el descriptor invariante obtenido anteriormente. Además se usa una medida más efectiva para realizar el calce, la cual consiste en usar el segundo vecino más cercano. Se evalúa que el radio entre las distancias del primer y segundo vecino más cercano sea mayor a 0.8 para aceptar el calce.

Verificación de calces: La etapa anterior permite eliminar calces falsos al usar la distancia del radio entre el primero y segundo vecino más cercano pero no permite identificar calces correctos o remover calces no validos de otros objetos. Muchos métodos de ajuste de parámetros como RANSAC o LMS (Least Mean Square) obtienen pobre desempeño a la hora de verificación de calces. Sin embargo la transformada de Hough entrega mejor desempeño al identificar grupos de características con una interpretación consistente que sean mapeados usando un mismo conjunto de parámetros. Cuando un grupo de características vota por la misma pose de un objeto, la probabilidad que la interpretación sea correcta es mucho más alta que cuando se usa sólo una característica.

Todos los grupos de características, identificados por la transformada de Hough votan, si en las celdas de votación hay 3 o más votos, es candidato a realizar la verificación geométrica. El proceso de verificación geométrica usa el algoritmo de mínimos cuadrados (LMS) con el fin de obtener los parámetros para la mejor transformación a fin relacionado a las 2 imágenes a comparar.

El calce final es aceptado basándose en un modelo probabilístico de rechazo. Un análisis bayesiano se usa para ver si el objeto está presente basándose en el número de calces encontrados. Se acepta el calce si el modelo entrega una probabilidad mayor a 0.98.

Es importante mencionar que en este trabajo se usan 2 enfoques para encontrar los calces utilizando la metodología SIFT. El primer enfoque es el realizado por el autor del método SIFT y que fue explicado anteriormente [6], donde el principal problema es que se detectan muchos falsos positivos. El segundo enfoque para los calces corresponde al realizado en el trabajo [18] donde proponen usar varias etapas de rechazo probabilísticas para eliminar el problema de la detección de los falsos positivos.

2.5.5 Speeded-Up Robust Features (SURF)

La metodología SURF presenta un rápido y novedoso esquema para calcular un detector y descriptor invariante a escala y rotación [7]. El método SURF entrega un detector con alta repetitividad, lo cual nos permite encontrar los mismos puntos de interés bajo distintos puntos de vista de manera rápida y robusta.

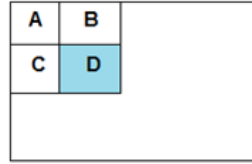
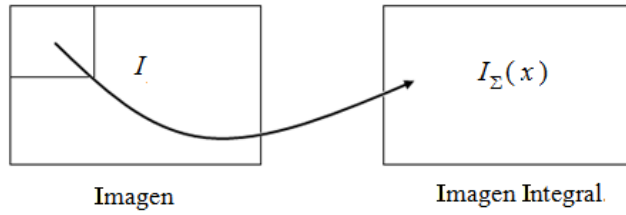
El descriptor SURF es eficiente debido a que usa la imagen integral que permite realizar la convolución con un cuadrado de manera muy rápida. Se usa una aproximación de la matriz Hessiana de Gaussiana para detectar los puntos de interés.

Algoritmo

Imagen integral: La imagen integral permite un rápido cálculo de la convolución para regiones cuadradas. La imagen integral $I_{\Sigma}(x)$ en la posición $x = (x, y)$ representa la suma de todos los pixeles en la imagen de entrada I dentro de una región rectangular formada por el origen de x .

$$I_{\Sigma}(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \quad (14)$$

Una vez que la imagen integral ha sido calculada, toma sólo 3 sumas calcular la suma de las intensidades sobre cualquier área rectangular. La Figura 19 muestra el esquema del cálculo de intensidades dentro de una región rectangular. Note que el tiempo de cálculo es independiente del tamaño del rectángulo.



$$\sum_D I(x, y) = I_\Sigma(D) + I_\Sigma(A) - I_\Sigma(B) - I_\Sigma(C)$$

Figura 19. Cálculo con la imagen integral. La suma de la imagen original dentro de cualquier rectángulo D de la imagen puede ser evaluada rápidamente usando la imagen integral: $\sum_D I(x, y) = I_\Sigma(D) + I_\Sigma(A) - I_\Sigma(B) - I_\Sigma(C)$.

Matriz Hessiana: La matriz Hessiana permite detectar estructuras de puntos en posiciones donde el determinante es máximo.

Dado un punto $x = (x, y)$ en una imagen I , la matriz Hessiana $H(x, \sigma)$ en x con escala σ se define como:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (15)$$

donde $L_{xx}(x, \sigma)$ es la convolución de una Gaussiana de segundo orden $\frac{\partial^2}{\partial x^2} g(\sigma)$ con la imagen I en el punto x , y de manera similar para $L_{xy}(x, \sigma)$, $L_{yy}(x, \sigma)$.

Los filtros Gaussianos son óptimos para el análisis del escala-espacio (Scale Space), pero en la práctica tienden a tener a ser discretizados y cortados. La matriz Hessiana es aproximada con filtros rectangulares debido a que la aproximación de Gaussianas de segundo orden puede ser calculada con muy bajo costo computacional usando las imágenes integrales. Además este cálculo es independiente del tamaño del filtro. En la Figura 20 se muestran ejemplos de filtros Gaussianos de segundo orden discretizados y cortados, y sus correspondientes aproximaciones de Gaussianas de segundo orden.

Se generan las imágenes D_{xx} , D_{xy} y D_{yy} a partir de la imagen I y los filtros Gaussianos de segundo orden. Luego se calcula el determinante aproximado de la matriz Hessiana dado por:

$$\det(H_{\text{aprox}}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (16)$$

donde el 0.9 aparece como un factor de corrección al usar cuadrados en vez de Gaussianas.

Este determinante aproximado de la matriz Hessiana representa la respuesta de la imagen en la localización x . Esta respuesta es analizada sobre diferentes escalas donde son detectados los máximos locales.

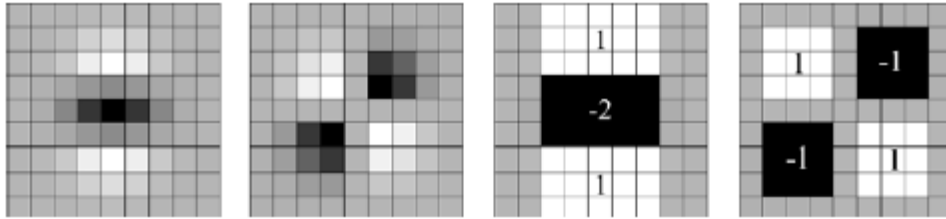


Figura 20. Filtros Gaussianos de segundo orden discretizados y las aproximaciones usadas. Tomado de [7].

Representación del escala-espacio: Los puntos de interés necesitan ser encontrados en escalas distintas debido a que la búsqueda de correspondencia entre imágenes debe ser vista a diferentes escalas. Las imágenes son repetidamente suavizadas con una Gaussiana y luego muestreadas para alcanzar un alto nivel en la pirámide.

Debido al uso de filtros cuadrados y la imagen integral, es posible aplicar los filtros cuadrados de cualquier tamaño y con la misma velocidad directamente en la imagen original, inclusive en paralelo. De esta manera el escala-espacio es realizado aumentando la escala del filtro en vez de reducir el tamaño de la imagen.

La construcción del escala-espacio comienza con un filtro de 9×9 , con el cual se calcula la respuesta de la imagen a una escala pequeña. Entonces son aplicados filtros de tamaños 15×15 , 21×21 , 27×27 , para tener todo el análisis a distintas escalas.

Localización de los puntos de interés: La localización de los puntos de interés es realizada en la imagen y en todas las escalas utilizando una vecindad cubo de $3 \times 3 \times 3$. Los máximos de la matriz Hessiana son interpolados en escala y en el espacio de imágenes.

El descriptor SURF describe la distribución de intensidad contenida dentro del punto de interés y sus vecinos. Se construye una distribución de Haar wavelets de primer orden en las direcciones x e y en vez de usar el gradiente, utilizando las imágenes integrales y usando 64 dimensiones para el vector descriptor. Con esto se tiene una reducción del tiempo de extracción de características y calces. Además el descriptor SURF incorpora un índice basado en el signo del Laplaciano, con el fin de mejorar aún más la robustez del descriptor y la velocidad del tiempo de calces.

Asignación de la orientación: Para que el descriptor sea invariante a la rotación de la imagen, se identifica una orientación reproducible para los puntos de interés. Primero se calcula la respuesta del filtro Haar wavelet en una posición x e y dentro de una vecindad circular de radio $6s$ alrededor del punto de interés, s representa la escala en la cual el punto de interés fue detectado.

Se usa la imagen integral para realizar un filtrado rápido. Los filtros Haar son mostrados en la Figura 21.

Una vez que las respuestas del filtro son calculadas, centrada en el punto de interés, son representadas como puntos en un espacio donde la respuesta horizontal es ubicada en la abscisa y la respuesta vertical en la ordenada. La orientación dominante es estimada como la suma de todas las respuestas dentro de una ventana móvil de tamaño $\pi/3$. La suma más grande sobre todas las ventanas define la orientación del punto de interés. La Figura 22 muestra la asignación de orientación.

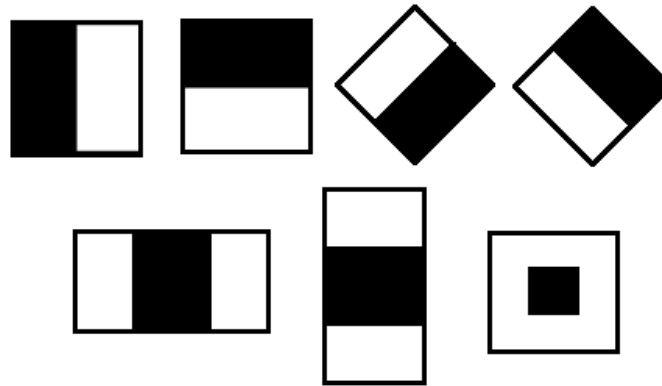


Figura 21. Filtros Haar. Corresponden a una onda cuadrada (rectángulos adyacentes) uno claro y otro oscuro. Se aplican en la imagen para determinar características, restando el promedio de los píxeles de la región oscura y los píxeles de la región clara. Modificado de [7].

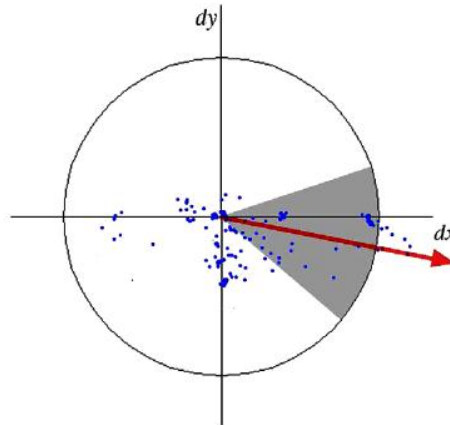


Figura 22. Asignación de la orientación. La orientación dominante es estimada como la suma todas las respuestas dentro de una ventana móvil. La mayor suma determina la orientación. Tomado de [7].

Descriptor: Para la extracción del descriptor, el primer paso consiste en la construcción de un cuadrado centrado en la región del punto de interés y orientado a lo largo de la orientación asignada anteriormente. El tamaño del filtro es de $20 s$.

La región cuadrada es dividida en pequeñas subregiones de 4×4 , para preservar la información espacial. Para cada subregión se calcula la respuesta del filtro Haar en 5×5 puntos

regularmente espaciados. Las respuestas horizontales d_x y verticales d_y del filtro Haar son sumadas sobre cada subregión y forman el primer conjunto del vector de características. Además, se extrae la información acerca del cambio de intensidad usando $|d_x|$ y $|d_y|$. Usando cada subregión se tiene un descriptor de 4 dimensiones, dado por un vector de intensidades $v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$. Concatenando para cada subregión de 4×4 se obtiene un vector descriptor de largo 64. La Figura 23 muestra la construcción del descriptor.

Calces: Para una rápida identificación en la etapa de calce, se incluye el signo del Laplaciano. El signo del Laplaciano permite distinguir entre estructuras de puntos con fondos oscuros o brillantes, aún si estas estructuras poseen similar forma. En la etapa de calces se comparan sólo las características que tienen el mismo tipo de contraste.

En el presente trabajo se usa la misma etapa de calces usada en el método SIFT, la cual fue explicada en la Sección anterior (2.5.4).

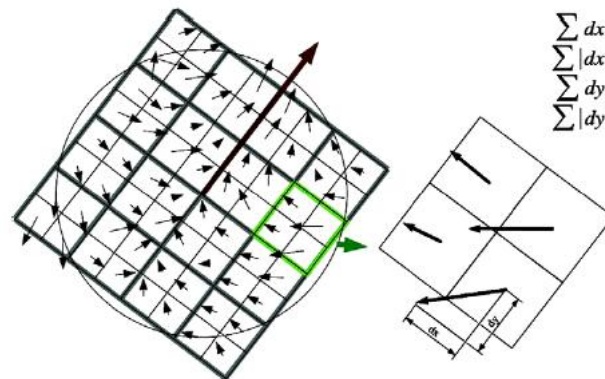


Figura 23. Descriptor SURF. El descriptor SURF es dado por un vector que posee 4 componentes: $v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$, donde $\sum d_x$ y $\sum d_y$ representan la suma de las respuestas horizontales y verticales del filtro Haar, y $\sum |d_x|$ y $\sum |d_y|$, corresponden al cambio de intensidad en esas direcciones. Tomado de [7].

3 Estudio Comparativo de Reconocimiento de Rostros usando Imágenes Térmicas¹

3.1 Introducción

El objetivo de este capítulo es realizar un estudio comparativo de métodos de reconocimiento de rostros en ambientes controlados y no controlados, es decir en espacios interiores tales como una oficina de trabajo o una sala de estar y en espacios exteriores tales como un patio o al aire libre, con diferentes tipos de iluminación, diferentes poses, oclusiones, etc. La principal motivación es la falta de comparaciones directas y detalladas de este tipo de métodos trabajando bajo las mismas condiciones. Los resultados de este estudio comparativo nos permiten desarrollar sistemas de visión computacional que sean capaces de reconocer en forma robusta rostros humanos usando cámaras térmicas en el rango 8-12 μm .

El estudio comparativo es llevado a cabo usando la base de datos Equinox y la base de datos UCHThermalFace. La base de datos Equinox fue seleccionada porque es una de las bases de datos más frecuentemente utilizada en el ámbito térmico, y además permite comparar los resultados obtenidos con otros estudios. La base de datos UCHThermalFace fue especialmente diseñada para el estudio del reconocimiento de rostros en ambientes controlados y no controlados en el dominio térmico. La base de datos incorpora imágenes térmicas adquiridas en condiciones interiores y exteriores, con variación de iluminación natural, expresiones faciales, pose, accesorios, oclusión y fondo. Por lo tanto, es adecuada para realizar el estudio y análisis de métodos de reconocimiento de rostros en ambientes no controlados.

3.2 Métodos bajo comparación

Como se menciono anteriormente, los métodos bajo comparación son seleccionados considerando una serie de requerimientos como: aplicación en tiempo real, operación en línea, sólo una imagen por persona, operación en ambientes no controlados, y su desempeño obtenido en estudios comparativos [67][46][68][5] y en detección de rostros [4]. En el Capítulo 2 se encuentra de manera detallada cada uno de los métodos usados en este trabajo. Sin embargo, se procede a realizar una explicación rápida de los métodos indicando los parámetros y características usadas para realizar este capítulo.

Histogramas LBP: El reconocimiento de rostros usando histogramas de características LBP (Local Binary Pattern) fue un método originalmente propuesto en [3] y ha sido usado por muchos

¹ Este capítulo está basado en la publicación [15] de la revista científica *Pattern Recognition - Elsevier*.

grupos de investigación desde entonces. En el enfoque original, tres niveles diferentes de localidad son definidos: nivel pixel, nivel regional, y nivel holístico. Los primeros 2 niveles de localidad son alcanzados dividiendo el rostro en pequeñas regiones donde son extraídas las características LBP y luego son creados histogramas para tener una representación de la información de texturas de manera eficiente. El nivel de localidad holístico, que corresponde a la descripción global del rostro, es obtenido con la concatenación de los histogramas de características LBP extraídos en cada región. El reconocimiento es realizado usando un clasificador de vecinos más cercano calculado en el espacio de características, usando una de las siguientes medidas de disimilitud: intersección de histogramas, distancia euclidiana, o estadística Chi cuadrado. El sistema de reconocimiento es implementado sin considerar preprocesamiento (ningún corte usando una máscara elíptica ni ecuilización de histograma, usados en [3]), y dividiendo las imágenes en 10 (2x5), 40 (4x10) o 80 (4x20) regiones.

Histograma WLD: El descriptor WLD (Weber Linear Descriptor) [4] es inspirado en la ley de Weber y calcula histogramas dimensionales de intensidad local (diferencia de excitación) y orientación. Primero, dado un pixel, del componente de excitación diferencial es calculado como el radio entre la diferencia relativa de intensidad de un pixel actual en contra de sus vecinos, y la intensidad del pixel actual. El componente de orientación es la orientación del gradiente del pixel actual. Luego es construido el histograma 2D de la excitación diferencial y orientación. Así, tal como LBP, WLD es un descriptor denso que es calculado sobre cada pixel, el cual hace la diferencia con SIFT el cual corresponde a un descriptor disperso. Además, WLD tiene más granularidad que SIFT, dado que es calculada en una región pequeña. En este trabajo el reconocimiento es realizado usando un clasificador de vecinos más cercano que opera en el espacio de características usando una de las siguientes medidas de disimilitud: intersección de histograma, distancia euclidiana o estadística de Chi cuadrado. Como en el caso de LBP, las imágenes son divididas en regiones de 10, 40 y 80 divisiones, y los histogramas obtenidos son concatenados para cada región para obtener el descriptor. El histograma 2D es cuantizado en 8 orientaciones con 64 valores de excitación diferencial.

Descriptores Jets de Gabor: Enfoques basados en calces locales para el reconocimiento de rostros en el espectro visible son comparados en [5]. El estudio analiza diferentes representaciones de características locales, métodos de clasificación, y alternativas de combinación de clasificadores. Tomando en cuenta los resultados de ese estudio, el autor implementa un sistema que integra las mejores elecciones posibles de cada análisis. El sistema usa descriptores Jets de Gabor (GJD) como características locales, las cuales son uniformemente distribuidas sobre la imagen, en una longitud de onda. En cada posición de la grilla de la imagen de test y la imagen de galería, y en cada escala (análisis multi-escala), el jet de Gabor es comparado usando el producto interno normalizado, ese resultado es combinado usando el método de cuenta de Borda. En la representación de características de Gabor, sólo las magnitudes de Gabor son usadas, y son utilizadas 5 escalas y 8 orientaciones para los filtros de Gabor. Se implementa el sistema usando todos los parámetros descritos en [5] (frecuencias del filtro, orientaciones, posición de la rejilla).

Descriptores SIFT: Enfoques basados en calces globales utilizando puntos de interés y descriptores se han vuelto muy populares y han experimentado un desarrollo impresionante en los últimos años. Típicamente, los puntos de interés son extraídos independientemente de la imagen de test y la imagen de referencia, ellas son caracterizadas por medio de descriptores invariantes,

donde finalmente estos descriptores son calzados hasta que sea encontrada una transformación entre las 2 imágenes. El sistema de Lowe [6], el cual usa el descriptor SIFT y etapas de rechazo probabilístico, es una elección popular para la implementación de reconocimiento de objetos dada sus capacidades de reconocimiento y su operación en tiempo real. Sin embargo, el principal inconveniente del sistema de Lowe es el largo número de detección de falsos positivos. Este inconveniente puede ser solucionado usando muchas etapas de rechazo probabilístico como el sistema L&R [18]. Este sistema ha sido usado para la construcción de un sistema robusto de verificación de huellas dactilares [18], para verificación de firmas (offline) [67] y para reconocimiento de rostros [46]. En este trabajo, se usa el sistema de Lowe y el sistema L&R para construir un sistema de reconocimiento de rostros. En el caso de SIFT-Lowe, se usa solo el número de calces para evaluar la tasa de desempeño en el reconocimiento. En el caso de SIFT-L&R, se consideran 2 variantes: el número de calces obtenidos y el número de calces obtenido después de usar muchas etapas de rechazo probabilísticas.

Descriptor SURF: El método SURF (Speeded Up Robust Features) [7] usa un procedimiento propuesto para obtener y calcular puntos de interés locales y descriptores con mayor velocidad que el enfoque SIFT. Esto es alcanzado usando: (i) el llamado detector rápido de Hessiana, el cual realiza una aproximación de filtros Gaussianos con filtros cuadrados y el proceso de la convolución se realiza muy rápidamente con el uso de la imagen integral, y (ii) el descriptor SURF que es más simple de calcular que el descriptor SIFT y corresponde a la respuesta de la distribución de filtros Haar-wavelets dentro de la región de la vecindad del punto de interés. Además, el descriptor SURF tiene un número menor de componentes que el descriptor SIFT. En este trabajo, se usa una implementación de SURF entregada por [69]. Como en el caso de SIFT-Lowe, se usa sólo el numero de calces como una medida de similaridad en orden de obtener la tasa de desempeño de reconocimiento.

Notación: Métodos y Variantes

Se usa la siguiente notación para referirse a los métodos y sus variaciones: A-B-C. La letra A describe el nombre del algoritmo de reconocimiento de rostros: LBP - Histograma de características LBP, WLD - Histograma de características WLD, GJD - Descriptores Jets de Gabor, SIFT-Lowe - Sistema original de Lowe con descriptores SIFT, SIFT-L&R - Sistema L&R con descriptores SIFT, SURF - correspondiente al método SURF. La letra B denota la medida de disimilitud o el tipo de clasificador: HI - Intersección de histograma, XS - estadística Chi cuadrado, EU - Distancia euclidiana, BC - cuenta de Bourda, excepto para el caso de SIFT y SURF donde no se usa una medida de distancia explicita. La letra C corresponde a los parámetros adicionales: el número de divisiones para el caso de los métodos basados en LBP o WLD, y el método de calces en el caso de los métodos SIFT y SURF (M: número de calces; S: número de calces con muchas etapas de rechazo probabilísticas).

3.3 Estudio comparativo usando la base de datos Equinox

El desempeño de los métodos es analizado usando la base de datos Equinox [50], la cual ha sido usada para validar métodos de reconocimiento de rostros usando imágenes térmicas. Esto permite una directa comparación de los métodos seleccionados con métodos de reconocimiento de rostros implementados en trabajos anteriores: en [1][2] son comparados métodos basados en LDA (análisis de discriminante lineal), LFA (análisis de característica local), y el método ICA

(análisis de componentes independientes). El método PCA (análisis de componentes principales) ha sido analizado en [56][2][10]. Otros métodos que han sido analizados usando esta base de datos son KPCA (Núcleo PCA) y KFLD (Núcleo de discriminante lineal de Fisher) en [57], LBP (Local Binary Pattern) en [48] y Wavelets en [56].

La base de datos Equinox consiste en 18.629 imágenes térmicas (8-12 μm) y 18.629 imágenes visibles de 240x320 píxeles de resolución. Todas las imágenes térmicas fueron radiométricamente calibradas (ver [10]). La base de datos contiene imágenes de 91 individuos obtenidas bajo 3 condiciones diferentes de iluminación (frontal, lateral izquierdo y lateral derecho), 3 expresiones diferentes (contento, enojado, sorprendido) y también una secuencia de 40 frames (4 seg.) en la cual el sujeto está hablando (vocales). La gente que usa lentes es capturada 2 veces, con y sin lentes. La Figura 24 muestra algunos ejemplos de la base de datos Equinox de imágenes térmicas y visibles obtenidas bajo diferentes condiciones. Para 3 de los 91 individuos, existen sólo imágenes donde el sujeto usa lentes. De esta manera, no es posible obtener una evaluación completa para esos sujetos. Es por esto que para hacer una evaluación justa, se remueven estos 3 sujetos de la base de datos, la cual finalmente contiene 88 individuos.

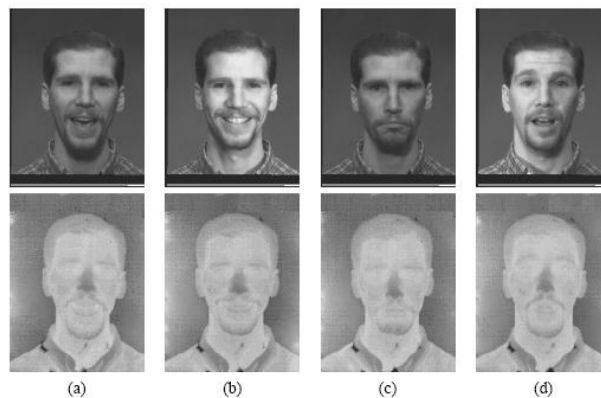


Figura 24. Ejemplos de la base de datos Equinox (tomado de [70]). Arriba: Visible; Abajo: Térmico. a) hablando (iluminación frontal), b) sonriendo (iluminación lateral derecho), c) enojado (iluminación frontal), d) sorprendido (iluminación lateral izquierda).

Siguiendo la terminología de Equinox, las imágenes 0, 3, 9 (de las 40 imágenes donde el sujeto está hablando) son llamados imágenes de vocales y las imágenes “sonriendo”, “enojado”, “sorprendido” son llamados imágenes de expresión. La metodología de evaluación para la base de datos Equinox crea múltiples conjuntos, que son mostrados en la Tabla 1, y son construidos y usados como conjuntos de test y conjuntos de galería. Para más detalles ver [10]. Así, cada uno de esos conjuntos (VA, EA, VF, EF, VL, EL, VG, EG, RR) pueden ser usados en diferentes experimentos, y ellos pueden ser parte tanto como conjunto de galería como conjunto de test o prueba. Los experimentos de reconocimiento son llevados a cabo siguiendo la metodología de Equinox [1][10] y para cada uno de los 54 conjuntos de galería-test es calculada la tasa de reconocimiento, utilizando el criterio del Top -1, que consiste encontrar el individuo que posea la distancia mínima entre un conjunto de datos de galería y un sujeto de prueba, y comparta el mismo identificador que el sujeto de prueba.

Tabla 1. Descripción de los conjuntos de imágenes Equinox

VA	Imágenes de vocales	todos los sujetos	todas las iluminaciones
EA	Imágenes de expresión	todos los sujetos	todas las iluminaciones
VF	Imágenes de vocales	todos los sujetos	iluminación frontal
EF	Imágenes de expresión	todos los sujetos	iluminación frontal
VL	Imágenes de vocales	todos los sujetos	iluminación lateral
EL	Imágenes de expresión	todos los sujetos	iluminación lateral
VG	Imágenes de vocales	todos los sujetos usando lentes	todas las iluminaciones
EG	Imágenes de expresión	todos los sujetos usando lentes	todas las iluminaciones
RR	500 imágenes	elegidas al azar	todas las iluminaciones

La media y desviación estándar para el reconocimiento Top-1 obtenidos de los 54 conjuntos de galería-test del experimento y los casos más difíciles para los pares de test/galería (VL/EF, EF/VL, VF/VG, VF/EG) son mostrados en la Tabla 2. Estos resultados son obtenidos usando las mismas condiciones reportadas en los trabajos anteriores [1][2][10]. Este experimento de reconocimiento consiste en usar una galería completa que incluye entre 3 y 9 imágenes por cada individuo en la base de datos. Se usa la metodología de evaluación propuesta por Equinox, la cual consiste en realizar una evaluación cruzada para todos los resultados del Top-1 de todos los pares de conjuntos. Sin embargo no se incluyen los subconjuntos de VA ni EA cuando ellos son usado como conjuntos de test, como en los resultados reportados en [1][2][10]. Algunos resultados presentados en este estudio incluyen estos conjuntos en la evaluación, pero el procedimiento correcto no debería considerarlos con el fin de prevenir el exceso de ajuste.

Los resultados obtenidos pueden ser resumidos en:

- Las variantes de WLD-X-X muestran el mejor desempeño, incluso en los casos difíciles (con lentes). La variante WLD-HI-80 obtiene el mejor resultado.
- Las variantes LBP-X-X obtienen el segundo mejor desempeño.
- Los métodos basados en SIFT, SURF y GJD no obtienen buenos resultados comparados con las variantes de WLD-X-X o LBP-X-X. Al parecer estos métodos son afectados por la calidad de la imagen del rostro, en este caso la baja calidad de las imágenes de la base de datos Equinox. Cuando estos métodos son usados en otras bases de datos, tal como los resultados presentados luego en la Sección 3.4, estos métodos muestran un alto desempeño.
- Los métodos basados en apariencia como PCA, LDA, etc., no obtienen buenos resultados comparados con WLD-X-X y LBP-X-X. Como puede ser observado, los métodos basados en apariencia no pueden manejar variaciones producto de las expresiones faciales o artefactos como lentes, donde existe pérdida de información debido a que el vidrio opaca la radiación térmica.

En un segundo experimento, los métodos son comparados con otros considerando una condición más real: 1 imagen por individuo en la galería y con imágenes de tamaño 150x81 pixeles de resolución. La misma metodología de evaluación descrita anteriormente es usada, pero en este caso la comparación es restrictiva a los 5 métodos bajo comparación removiendo los

métodos basados en apariencia debido a que no permiten trabajar adecuadamente con 1 imagen por individuo en la galería. Ellos necesitan al menos 2 ó 3 imágenes en orden de construir un modelo adecuado para el rostro. Los resultados son presentados en la Tabla 3. De la Tabla 3 se puede observar que:

- El mejor resultado es obtenido por la variante WLD-HI-80. Todas las variantes de WLD-X-X poseen un buen desempeño en este caso.
- Las variantes LBP-X-40 y LBP-X-80 obtienen buenos resultados, sobre el ~90% de la tasa de reconocimiento, seguidos por los otras variantes LBP y las variantes de SIFT.
- Las metodologías de GJD y SURF presentan los peores de los resultados.

En resumen, esta primera comparación de metodologías de reconocimiento de rostros en el espectro infrarrojo térmico incluye aspectos tales como variaciones de iluminación, variaciones en la expresión facial, variaciones faciales cuando la persona se encuentra hablando, y el uso de lentes, aspectos que afectan directamente el desempeño de un adecuado sistema reconocimiento de rostros. Los mejores resultados son obtenidos usando las variantes de WLD-X-X y LBP-X-X. Cuando se compara el desempeño con los otros métodos, GJD-BC, SIFT-X-X y SURF-X-X obtienen un desempeño bajo. La principal razón de este pobre desempeño para estas metodologías, parece ser la baja variabilidad en los valores de los píxeles de las imágenes térmicas de la base Equinox donde los valores de los píxeles están muy cuantizados. Incluso al aplicar técnicas de preprocesamiento de imágenes tales como ecualización de histograma o estiramiento de histograma, no mejoran el desempeño para estos métodos. Es importante mencionar que estos 5 métodos bajo comparación serán analizados a fondo en la siguiente Sección 3.4 usando la base de datos UCHThermalFace, la cual no posee problemas de contraste ni cuantización como la base de datos Equinox.

Tabla 2. Resultados para la base de datos Equinox. Experimentos de galería completa (las galerías incluyen entre 3 y 9 imágenes de cada individuo). La diagonal (donde los conjuntos de galería y test son iguales) y los casos en que el conjunto de test contiene subconjuntos de la galería no son considerados. SVD: desviación estándar. Los valores de media y desviación estándar usando la tasa de reconocimiento Top-1 para 54 pares de galería/test

Métodos	Media [%]	SDV	VL/EF [%]	EF/VL [%]	VF/VG [%]	EF/EG [%]
PCA [10]	96.0	3.2	86.8	91.2	99.7	98.3
PCA [1]	95.0	2.8	90.5	91.3	99.3	97.3
LDAG [1]	97.0	-	-	-	-	-
LDAt [1]	98.0	1.67	95.6	95.8	99.3	99.7
LFAe [1]	93.0	-	-	-	-	-
LFAb [1]	93.0	4.6	83.5	82.7	97.7	95.9
ICA [1]	94.0	3.9	87.9	85.7	99.3	96.8
PCA [2]	95.0	2.6	90.9	91.6	99.3	97.3
LDAt [2]	97.0	2.2	91.4	92.9	99.0	98.7
LDAG [2]	98.0	1.8	95.1	95.6	99.3	99.7
LFAb [2]	91.0	4.8	83.5	82.9	97.0	96.1
LFAe [2]	92.0	4.9	84.2	82.9	97.7	95.1
ICA [2]	94.0	3.9	86.8	86.2	99.3	96.6
Wavelets [56] **	93.5	4.4	91.2	90.4	-	-
PCA [56]**	92.9	4.3	91.2	90.4	-	-
KPCA [57]***	82.7	-	-	-	-	-
KFLD [57]***	96.3	-	-	-	-	-
LBP [48]****	97.3	1.8	99.2	96.6	-	-
LBP [48] con NUC****	93.3	4.3	93.6	90.2	-	-
LBP-EU-10	95.3	2.9	88.9	90.4	97.4	93.0
LBP-HI-10	98.4	1.0	96.2	97.4	99.3	97.1
LBP-XS-10	98.8	0.8	97.6	98.0	99.5	97.1
LBP-EU-40	97.2	1.8	92.7	95.7	98.8	94.7
LBP-HI-40	98.5	1.0	96.5	98.1	99.5	96.2
LBP-XS-40	98.3	1.2	95.5	97.8	99.8	95.2
LBP-EU-80	97.9	1.3	94.8	97.4	99.0	95.4
LBP-HI-80	98.7	0.9	96.7	98.5	99.5	96.6
LBP-XS-80	98.1	1.2	94.6	98.1	98.6	96.4
GJD-BC*	91.2	4.8	85.3	82.4	97.1	91.1
SIFT-L&R-M*	87.8	3.3	87.2	85.3	91.8	86.3
SIFT-L&R-S*	93.9	2.6	93.1	91.2	96.9	92.8
SIFT-Lowe-M*	91.9	3.3	84.9	92.0	96.2	92.8
SURF-M*	80.5	6.7	75.2	69.9	88.2	68.8
WLD-EU-10	98.1	1.4	96.5	96.8	100	96.2
WLD-HI-10	99.0	0.8	98.1	98.9	100	97.8
WLD-XS-10	98.9	0.8	98.3	98.6	100	96.9
WLD-EU-40	98.5	1.1	96.5	98.2	99.3	96.4
WLD-HI-40	98.9	0.7	97.9	99.1	99.5	97.8
WLD-XS-40	98.7	1.0	97.2	98.1	99.8	96.4
WLD-EU-80	98.9	0.7	97.4	98.7	99.3	97.6
WLD-HI-80	99.1	0.7	98.1	99.3	99.8	97.8
WLD-XS-80	98.9	0.8	97.4	98.7	99.8	96.6
NUC: Corrección de No-Uniformidad * Estiramiento de Histograma , rango 0-255 ** Se usan subconjuntos que no incluyen lentes ni algunas relaciones cruzadas. *** Se usan lentes y conjuntos de expresiones **** Se usan subconjuntos que no incluyen lentes						

Tabla 3. Resultados para la base de datos Equinox. Las galerías incluyen sólo 1 imagen por individuo. Todas las galerías, excepto VG, EG y RR son usadas. SDV: desviación estándar. Los valores de media y desviación estándar usando la tasa de reconocimiento Top-1 para 36 pares de galería/test. La mejor variante de cada método es mostrada en negritas.

Métodos	Media [%]	SDV	VL/EF [%]	EF/VL [%]	EL/VF [%]	VA/EF [%]
LBP-EU-10	77.1	6.0	75.1	66.7	76.4	75.1
LBP-HI-10	88.4	3.5	86.7	83.0	85.4	87.7
LBP-XS-10	91.4	3.7	92.6	85.9	87.2	91.9
LBP-EU-40	84.6	5.2	84.2	76.2	81.9	83.9
LBP-HI-40	91.0	3.6	90.9	85.6	86.8	90.9
LBP-XS-40	91.0	3.7	91.2	85.9	87.8	89.8
LBP-EU-80	88.1	4.2	83.0	86.3	82.1	85.5
LBP-HI-80	92.5	3.0	88.7	90.9	89.7	90.0
LBP-XS-80	89.8	3.6	85.6	88.4	85.0	86.9
GJD-BC*	71.7	7.4	63.9	63.4	66.3	70.9
WLD-EU-10	88.1	4.1	88.4	83.5	83.7	86.7
WLD-HI-10	92.2	2.6	92.3	89.4	88.2	91.6
WLD-XS-10	92.8	3.2	94.0	88.9	87.8	93.3
WLD-EU-40	91.5	4.1	91.6	87.2	87.2	93.3
WLD-HI-40	93.5	3.3	94.0	90.3	88.9	94.0
WLD-XS-40	93.2	3.6	94.4	89.4	89.2	94.0
WLD-EU-80	93.9	3.2	91.0	95.4	90.3	92.2
WLD-HI-80	95.0	2.9	91.7	95.8	92.9	92.9
WLD-XS-80	94.4	3.2	90.5	94.4	91.2	92.1
SIFT-L&R-M*	77.4	9.9	75.8	75.0	74.3	79.6
SIFT-L&R-S*	76.6	6.2	72.3	74.1	67.7	76.1
SIFT-Lowe-M*	82.2	5.2	70.9	82.1	80.9	79.3
SURF-M*	63.3	8.7	55.4	54.8	55.2	59.3

* Estiramiento de Histograma , rango 0-255

3.4 Estudio comparativo usando la base de datos térmica UCHThermalFace

Los métodos bajo estudio son analizados considerando condiciones reales que incluyen configuraciones en espacios controlados y no controlados y variaciones naturales en expresiones faciales, pose, accesorios, oclusión y fondo en la imagen.

3.4.1 Descripción de la base de datos UCHThermalFace

La base de datos UCHThermalFace está dividida en 3 conjuntos: Rotación, Palabras y Expresiones. Los conjuntos de Rotación y Palabras consisten en imágenes de rostros térmicos capturados en ambientes controlados y no controlados de 53 sujetos obtenidos bajo distintos ángulos de orientación (yaw) e inclinación (pitch), y un conjunto de imágenes capturadas mientras los sujetos se encontraban hablando. El conjunto de Expresión consiste en imágenes térmicas de 102 sujetos capturadas sólo en un ambiente controlado en interior. Las imágenes térmicas para los conjuntos Rotación y Palabras fueron adquiridas usando una cámara térmica FLIR 320 TAU [44], con sensibilidad en el rango de 7.5-13.5 μm , y una resolución de 324x256

pixeles. Para el conjunto de Expresiones se usa una cámara térmica Jade-UC IR Uncooled [43] de rango espectral 8-12 μm y resolución de 320x240 pixeles.

El conjunto de Rotación contiene 22 imágenes por sujeto, cada una corresponde a uno de los distintos ángulos de rotación adquiridos tanto para las sesiones en interior como en exterior (ver Figura 25 para ver la configuración del experimento). En ambos casos, sesión en interior como en exterior, la distancia del sujeto a la cámara térmica se encontraba fija en 120 cm, y la cámara térmica se encontraba situada en la posición P6 (ver Figura 25). Las imágenes de rostro térmicas fueron adquiridas mientras los sujetos observaban las posiciones 1-11 (ver Figura 25), la cual corresponde a los siguientes ángulos de rotación: R1: (orientación=-15°, inclinación=15°), R2: (orientación=0°, inclinación=15°), R3: (orientación=15°, inclinación=15°), R4: (orientación=-30°, inclinación=0°), R5: (orientación=-15°, inclinación=0°), R6: (orientación=0°, inclinación=0°), R7: (orientación=15°, inclinación=0°), R8: (orientación=30°, inclinación=0°), R9: (orientación=-15°, inclinación=-15°), R10: (orientación=0°, inclinación=-15°), R11: (orientación=15°, inclinación=-15°). La Figura 26 muestra un ejemplo de 11 imágenes de rostros térmico correspondiente a un individuo de la base de datos, adquirida bajo diferentes orientaciones e inclinaciones para una sesión en ambiente interior. Además de obtener el conjunto de Rotación, una secuencia de video fue capturada mientras cada sujeto observaba el punto P6 y decía la palabra “Pa-ra-le-le-pi-pe-do” tanto para la sesión en interior como la en la sesión exterior. Luego, 3 imágenes fueron seleccionadas de esta secuencia de video por cada individuo y para cada una de las sesiones interior e exterior. Estas imágenes forman el conjunto de Palabras, cual esencialmente contiene imágenes con diferente expresión facial y diferente gesticulación.

La configuración de la cámara térmica FLIR 320 TAU para el caso de las sesiones en interior y exterior fue realizada utilizando una configuración de parámetros de brillo y contraste fijos, elegidos manualmente por el autor de este trabajo. Estos valores de brillo y contraste fueron seleccionados con el fin de maximizar la información del rostro térmico, obteniendo imágenes no saturadas en el caso de interior y exterior cuando la cámara se encontraba en condiciones estables, es decir trabajando en su punto de operación. La cámara térmica trabaja con varios modos de operación, siendo el más usado, el modo de ganancia automática, que viene por defecto, debido a que usa un algoritmo eficiente de ecualización de histograma. Sin embargo, para estos experimentos se utilizaron parámetros fijos debido a los buenos resultados obtenidos en pruebas de laboratorio y porque se observó que al variar mucho el fondo de la imagen, los valores de los pixeles del rostro variaban mucho producto del modo ganancia automática que posee la cámara térmica. Este tema será tratado con más detalle en el Capítulo 4, Apartado 4.3.

El conjunto de Expresión fue capturado con una configuración diferente. En esta configuración los sujetos fueron capturados mientras observaban frontalmente la cámara térmica a una distancia fija de 150 cm. Primero, las imágenes fueron adquiridas mientras los sujetos mostraban 3 expresiones distintas “Feliz”, “Triste” y “Enojado”. Además una secuencia de video fue capturada mientras los sujetos se encontraban hablando diferentes vocales. Luego, 3 imágenes fueron seleccionadas al azar de cada secuencia de video por individuo.

En resumen, para los conjuntos de Rotación y Palabras, se crearon 14 conjuntos para interior y 14 conjuntos para exterior para llevar a cabo los experimentos de reconocimiento de rostros. Para el caso de interior y exterior, 11 conjuntos corresponden a las diferentes

combinaciones de ángulos de orientación e inclinación en el caso de conjuntos de Rotación (conjuntos de R1 a R11) y 3 conjuntos para las diferentes imágenes capturadas en los conjuntos de Palabras (conjuntos S1 a S3). Para el conjunto de Expresión son definidas 3 expresiones (“Feliz”, “Triste”, y “Enojado”) los que forman los conjuntos E1 a E3 y 3 conjuntos obtenidos de secuencia de las vocales (V1 a V3). Los experimentos reportados hacen uso de 34 conjuntos definidos. En cada uno de los experimentos es dado un conjunto el cual es usado como test, y un segundo conjunto el cual es usado como galería.

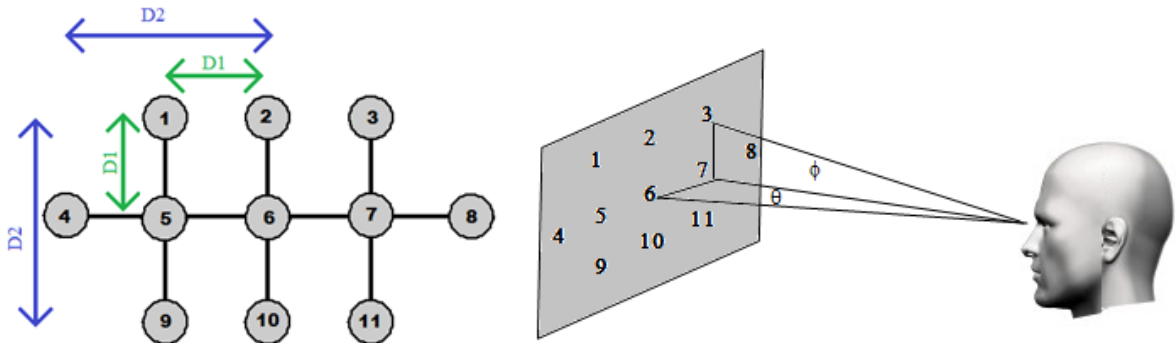


Figura 25. Configuración del experimento. Configuración experimental para la adquisición de datos con diferente ángulo de orientación (θ) y diferente ángulo de inclinación (ϕ). La distancia del individuo observando el punto P6 es 120 cm. D1 es 32.15 cm y D2 es 69.28 cm.



Figura 26. Ejemplo para la base UCHThermalFace. Ejemplo de 11 imágenes térmicas del conjunto de Rotación de un individuo de la base de datos UCHThermalFace capturado en la sesión en interior.

3.4.2 Descripción de los experimentos

Para evaluar y analizar los métodos de reconocimiento de rostros, 6 tipos de experimentos son llevados a cabo: (i) tamaño de la ventana, (ii) oclusión parcial del rostro, (iii) precisión en la detección de ojos, (iv) galería de interior versus exterior, (v) expresiones faciales, y (vi) distintas distancia. En todos los experimentos, excepto en el caso (v) de expresiones faciales, las caras

fueron alineadas con los ojos en la misma posición relativa, utilizando una distancia entre ojos de 42 píxeles. En el caso (vi) de los experimentos de distancia, la distancia entre ojos fue decreciendo de acuerdo con la reducción de la imagen. En el caso (v) de expresiones faciales, se usó un detector de rostros térmico basado en el enfoque de Viola-Jones [71], con el fin de incluir resultados sin alineación de ojos, además de las expresiones faciales.

La descripción de cada uno de los experimentos realizados es mostrada a continuación:

Tamaño de la ventana: En este experimento se analiza el desempeño de los métodos bajo análisis midiendo el efecto de usar diferentes tamaños de ventana en los rostros térmicos. Incrementar el tamaño de la ventana corresponde a agregar o remover diferentes cantidades de fondo en la región a ser analizada, sin cambiar la resolución de la imagen del rostro térmico. De esta manera, los experimentos son realizados incluyendo diferentes tamaños de ventana de 81x150 píxeles, 100x185 píxeles y 125x225 píxeles. La Figura 27 muestra algunos ejemplos.

Oclusión parcial del rostro: Para analizar el comportamiento de los métodos en respuesta a la oclusión parcial del rostro, las imágenes fueron divididas en 10 diferentes regiones (2 columnas y 5 filas), y una de estas regiones fue seleccionada al azar y sus píxeles llevados a 0 (ver ejemplos en la Figura 28). Este experimento analiza el comportamiento de los diferentes métodos en respuesta a la oclusión parcial.

Precisión en la detección de ojos: Muchos métodos de reconocimiento de rostros son muy sensitivos al alineamiento del rostro, el cual depende directamente de la precisión en el proceso de detección de ojos; la posición de los ojos es usualmente la primaria y a veces la única fuente de información para la alineación del rostro. Para analizar la sensibilidad de la precisión de la posición de los ojos para los diferentes métodos, se agrega ruido en la posición de la anotación de los ojos en las imágenes de test. El ruido fue agregado independientemente en la posición x y en la posición y y en cada ojo, usando el procedimiento descrito en [46]. En los distintos experimentos, el ruido puede tomar valores de 2.5%, 5%, ó 10% de la distancia entre ojos. Algunos ejemplos son mostrados en la Figura 29. Note que esto afecta la posición y la escala en la imagen del rostro térmico.

Galería de interior vs exterior: El desempeño de los métodos de reconocimiento de rostros depende fuertemente de las condiciones ambientales, particularmente bajo condiciones controladas y no controladas. En estos experimentos, las imágenes de test y las imágenes de galería corresponden a imágenes tomadas en sesiones en interior como en sesiones en exterior. Cuando las imágenes de test son imágenes de interior, entonces las imágenes de galería corresponden a imágenes de exterior y viceversa. Las imágenes de exterior fueron capturadas en verano (con temperaturas altas superiores a 30° Celsius), y a veces los rostros, como también la cámara, recibían directamente la luz del sol, provocando en algunos casos la saturación de las imágenes térmicas.

Expresiones faciales: En este experimento se analiza el efecto de tener diferentes expresiones faciales en los sujetos. Los experimentos consideran a los sujetos con diferentes expresiones faciales y además consideran diferente gesticulación debido a que los sujetos se encontraban hablando diferentes vocales.

Distintas distancias: En este experimento se analiza la sensibilidad de los métodos bajo distintas distancias entre los sujetos y la cámara. El efecto de tener distintas distancias entre el sujeto y la cámara es simulado disminuyendo la resolución de las imágenes térmicas. La resolución de la imagen es disminuida en un factor de $\sqrt{2}$ generando 5 nuevas distancias, desde 1.2 hasta 4.8 metros. Se usa para la evaluación de los métodos de reconocimiento de rostros, el conjunto de R6 para galería versus el conjunto de S1 como test.

En resumen los conjuntos usados para los experimentos de expresión corresponden al conjunto de Expresión (E1-E2) y (V1-V2), mientras que para todos los otros experimentos son usados los conjuntos de Rotación (R1-R11) y Palabras (S1-S3).



Figura 27. Ejemplos de rostros con diferentes tamaños de cortes (base de datos UCHThermalFace). (a) Sesión interior. Tamaños de ventana (en píxeles): 81x150, 100x185, 125x225. (b) Sesión exterior. Tamaños de ventana (en píxeles): 81x150, 100x185, 125x225.



Figura 28. Ejemplos de imágenes con oclusión parcial (base de datos UCHThermalFace). (Izquierda) Sesión interior. Tamaño de la ventana 125x225 píxeles. (Derecha) Sesión exterior. Tamaño de la ventana 125x225.



Figura 29. Ejemplos de imágenes alineadas después de aplicar ruido en la posición de los ojos (base de datos UCHThermalFace). Sesión interior. Tamaño de la ventana 125x225 píxeles. (a) Sin ruido; (b) 2.5% ruido; (c) 5% ruido; (d) 10% ruido.

3.4.3 Resultados de reconocimiento

El desempeño de los diferentes métodos es evaluado con la tasa de reconocimiento utilizando el criterio del Top-1. En todos los experimentos, el conjunto de Rotación R6, sin ninguna oclusión y sin incluir ruido en la posición de los ojos, es seleccionado como conjunto de galería debido que contiene una clara vista frontal para todos los sujetos de la base de datos. Naturalmente, en los experimentos en interior, es usado el conjunto R6 de interior, mientras que en el caso de exterior, es usado el conjunto R6 de exterior.

En las Tablas 4-11, la tasa de reconocimiento basada en el Top-1 es dada separadamente para los conjuntos de Rotación y Palabras. Para cada categoría una media es calculada: la media de las tasas de reconocimiento sobre todos los conjuntos de Rotación y la media de las tasas de reconocimiento sobre todos los conjuntos de Palabras. Además, el promedio entre estos dos resultados es calculado como la media total del sistema. La Figura 31 y la Figura 32 incluyen sólo el valor promedio. La Tabla 10 reporta los resultados en promedio obtenido usando el conjunto de Expresiones.

Los resultados son resumidos para cada uno de los experimentos mencionados anteriormente:

Tamaño de la ventana: La Tabla 4 muestra el desempeño de los diferentes métodos cuando diferentes tamaños de ventanas son aplicados en los conjuntos de test y galería. En todos estos experimentos, todos los conjuntos de galerías y test corresponden a imágenes obtenidas en interior.

Se puede concluir de los experimentos que el mejor tamaño de ventana depende de cada uno de los métodos. Para SIFT-X-X, SURF-M y GJD-BC el mejor tamaño de ventana es 125x225 píxeles, mientras que para WLD-X-X y LBP-X-X el mejor tamaño de ventana es de 81x150 píxeles. Estos resultados son consistentes con los resultados obtenidos en [46] para el caso de los métodos GJD-BC y LBP-X-X. Como se nota en ese estudio, GJD-BC trabaja mejor con ventanas que contienen una gran porción de fondo. La razón es debido a: (i) los filtros Gabor

codifican la información acerca del contorno de la cara, (ii) regiones grandes permiten el uso de filtros grandes, cuales codifican información para mayores escalas, y (iii) la clasificación de la cuenta de Bourda puede reducir el efecto de las regiones (o jets) que no son relevantes para el reconocimiento. Para el caso de las variantes de LBP-X-X, el hecho de agregar algún fondo no ayuda, y en la mayoría de los casos disminuye el rendimiento. Se puede observar que las variantes de WLD-X-X tienen similar comportamiento que las variantes de LBP en términos del tamaño de la ventana óptimo. Esto parece ser debido a la similitud en el tipo de análisis que usan ambos métodos con la información de la cara, cual es básicamente calcular el histograma de las características locales sobre regiones no superpuestas dentro de la región de la cara. En los casos donde los histogramas incluyen información del fondo (como cuando se usa un tamaño de la ventana grande), la habilidad de discriminar del método disminuye. Respecto a las variantes del método SIFT-X-X, en [46] pequeñas ventanas obtenían mejor desempeño en el caso de las imágenes visibles. La razón es que el fondo perturba el proceso de reconocimiento en el caso de imágenes visibles debido a que genera calces incorrectos entre puntos de interés correspondientes a regiones de la cara y el fondo. Sin embargo, este no es el caso de las imágenes térmicas, en las cuales el fondo es usualmente uniforme alrededor de la región de la cara, debido a que la temperatura de esta región es más baja que en el rostro o en la cabeza de los humanos. Esta es una ventaja adicional en el uso de las imágenes térmicas.

En términos del mejor rendimiento observado, las variantes de SIFT-X-X son las que obtienen los más grandes valores para la tasa de rendimiento, seguido del método SURF-M. El tercer mejor desempeño es logrado por las variantes del método WLD-X-X. En cuarto lugar se tienen los métodos GJD y LBP-X-X, los cuales obtienen similar rendimiento.

En la Tabla 4 se observa que las variantes de SIFT-X-X son más robustas contra las rotaciones. La principal razón parece ser que los calces de SIFT son robustos en contra de rotaciones de la cara fuera del plano hasta sobre los ~40%. Para ángulos de orientación e inclinación moderados (R2, R5, R7 y R10) muchos de los otros métodos se comportan bien, y en la mayoría de los casos, cuando el tamaño de la ventana es seleccionada adecuadamente, la tasa de reconocimiento obtenida es mayor a los ~90%.

Los métodos SIFT-X-X y SURF-M son más robustos en términos de las variaciones en expresiones faciales (en el conjunto de Palabras). En muchos casos estos métodos obtienen una tasa de reconocimiento de 100% en los experimentos reportados en la Tabla 4. La principal razón es que la forma en que los métodos SIFT y SURF realizan el paradigma de los calces, la cual básicamente calza pares de puntos de interés y encuentra una transformación coherente entre los conjuntos de calces (ver ejemplo en la Figura 30). Las variantes del método GJD-BC también son robustos a expresiones faciales. Una de las variantes de GJD-BC (100x185) obtiene una tasa de reconocimiento de 100% en los experimentos reportados en la Tabla 4. Esta robustez se alcanza gracias a la clasificación por la cuenta de Bourda, la cual realiza la votación de los Gabor-jets. Las variantes de WLD-X-X y LBP-X-X también muestran alguna invariancia a las expresiones faciales. Cuando el tamaño de la ventana es pequeño (81x150), ellos alcanzan una tasa de reconocimiento de 95.6% y 92.5% respectivamente.

Para cada uno de los métodos, el mejor resultado obtenido usando la tasa de reconocimiento más alta (considerando diferentes parámetros y distintos tamaños de ventana) es dado por las siguientes variantes:

- SIFT-Lowe-M, 125x225 pixeles: 97.6%
- SIFT-L&R-M, 125x225 pixeles: 95.1%
- SURF-M, 125x225 pixeles: 94.0%
- WLD-EU-80, 81x150 pixeles: 91.8%
- WLD-HI-80, 81x150 pixeles: 91.5%
- GJD-BC, 125x225 pixeles: 91.4%
- LBP-HI-80, 81x125 pixeles: 88.5%

En el caso de los métodos WLD y SIFT son seleccionadas 2 variantes. En el caso del método WLD, el desempeño de las 2 variantes es muy similar. En el caso del método SIFT, las variantes corresponden a 2 enfoques diferentes en la implementación del método. Para los siguientes experimentos, sólo el resultado de estas 7 variantes es considerado y analizado.

La Tabla 5 muestra el desempeño de las 7 variantes de los métodos seleccionados para los conjuntos de Rotación y Palabras en el caso de exterior. Los mejores resultados son obtenidos por las variantes de los métodos SIFT-X-X y SURF-M, seguido por el método WLD-HI-80. El cuarto lugar es obtenido por LBP-EU-80, y el quinto mejor desempeño es obtenido por el método GJD-BC. El método SIFT-Lowe-M muestra una alta tasa de reconocimiento (96.3%), la cual es muy similar a la obtenida en el caso de interior (97.6%). Los métodos SIFT-L&R-M y SURF-M muestran un comportamiento similar. En el caso de WLD-HI-80 y LBP-HI-80, el desempeño también es similar al obtenido en ambos casos de interior y exterior. Interesante es el caso de la variante WLD-EU-80, variante que usa la distancia euclidiana, donde el desempeño disminuye mucho en el caso de exterior. Lo mismo sucede con GJD-BC, donde el desempeño disminuye cerca de 11 puntos porcentuales con la configuración de exterior. Este resultado indica que GJD-BC no posee un comportamiento apropiado en condiciones de exterior.

Precisión en la detección de ojos: La Figura 31 muestra la sensibilidad de cada método con respecto a la precisión en la detección de ojos en el caso de interior, mientras que en la Figura 32 muestra la sensibilidad de los métodos en el caso de exterior.

Se puede observar que los valores más altos en la tasa de reconocimiento son obtenidos por SIFT-Lowe-M, donde se aprecia además que este método es muy robusto a variaciones en la posición de los ojos en ambos casos de interior y exterior. La tasa de reconocimiento varía desde la tasa de reconocimiento obtenida sin ruido entre 97.6% a 96.8% cuando es agregado un ruido de 10% en la posición de los ojos para los experimentos en interior y entre un 96.3% y 96.6% bajo el mismo nivel de ruido en los experimentos en exterior. Los métodos SIFT-L&R-M y SURF-M muestran un comportamiento similar. La principal razón de la robustez es la forma en que los métodos SIFT y SURF realizan los calces. (Ver explicación en el párrafo anterior y ejemplos en la Figura 30.)

En el caso de los experimentos en interior, el desempeño de los métodos LBP-HI-80, WLD-X-80 y GJD-BC decrece uniformemente con cada nivel de ruido en la posición de los ojos.

LBP-HI-80 y GJD-BC tienen el mismo comportamiento en los experimentos en exterior (una disminución máxima de ~12% en la tasa de reconocimiento). Sin embargo, las condiciones de exterior afectan un poco más las variantes de WLD-X-80, cuando es agregado un ruido de 10% en la posición de los ojos la tasa de reconocimiento disminuye en ~15%. No obstante, en la mayoría de los casos los métodos LBP-HI-80 y WLD-HI-80 obtienen el tercer mejor desempeño luego de los métodos de SIFT-X-X.

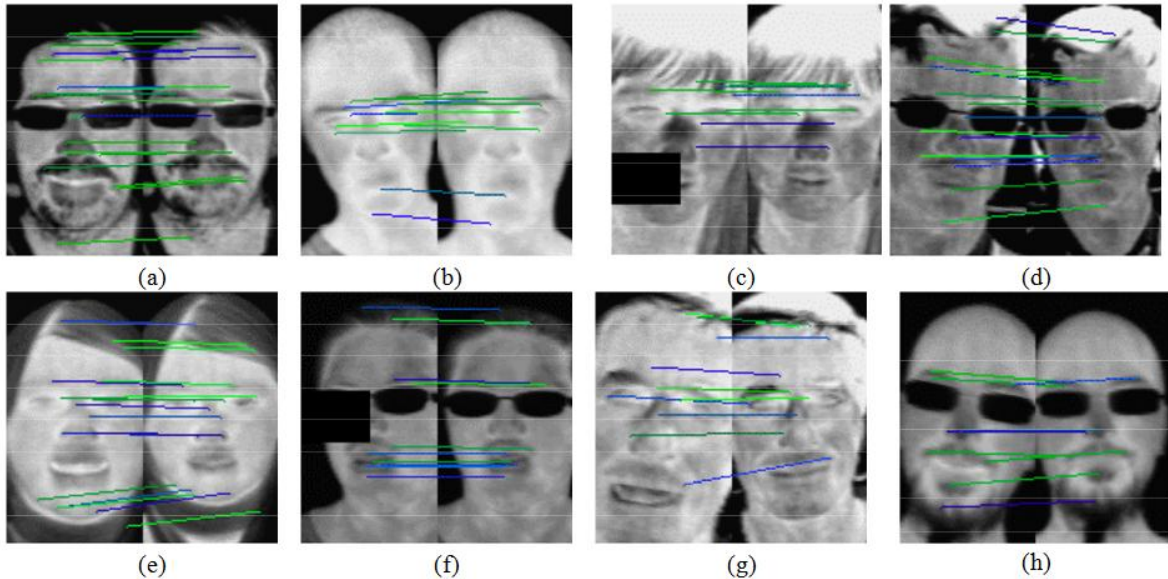


Figura 30. Ejemplos de calces correctos para algunos sujetos de la base de datos UCHThermalFace. En todos los casos la imagen de la derecha corresponde a la imagen de galería (R6) y en la imagen de la izquierda a la imagen de test. (a) Conjunto de Palabras S1. Sesión exterior; (b) Conjunto de Rotación R11. 10% de ruido en la posición de los ojos. Sesión interior; (c) Conjunto de Rotación R9. Oclusión. Sesión exterior; (d) Conjunto de Rotación R5. 10% de ruido en la posición de los ojos. Sesión exterior; (e) Conjunto de Palabras S3. Sesión interior; (f) Conjunto de Rotación R7. Oclusión. Sesión interior; (g) Conjunto de Palabras S3. 10% de ruido en la posición de los ojos. Sesión exterior; (h) Conjunto de Palabras S1. 10% de ruido en la posición de los ojos. Sesión interior.

Además, puede ser observado de las Figuras 31 y 32 que en algunos casos el desempeño de los métodos aumenta un poco con el ruido. Estas pequeñas variaciones no poseen significado estadístico y son producidas por la naturaleza estadística de los métodos, las que incluyen el cálculo de los histogramas de WLD y LBP, y el uso de las etapas de rechazo probabilísticas en el caso de las variantes de SIFT, como también el hecho de que algunas variaciones en la información de la posición de los ojos puede mejorar la alineación del rostro.

Oclusión parcial del rostro: La Tabla 6 muestra la sensibilidad de los métodos con respecto a la oclusión parcial de una región de la cara para el caso de interior, mientras que la Tabla 7 muestra la sensibilidad de los métodos en el caso de exterior.

En el caso de interior y en el caso de exterior se observa que el mejor desempeño es alcanzado por las variantes SIFT-X-X, y el segundo mejor desempeño es entregado por el método WLD-X-80. En ambos casos, las tasas de reconocimiento disminuyen cerca de 2-5% con una

oclusión del 10% de la cara, el cual es considerado un comportamiento muy bueno. Los métodos SURF-M, LBP-HI-80 y GJD-BC muestran un bajo desempeño, y son más afectados por la oclusión (las tasas de reconocimiento disminuyen cerca de 6-11%).

Galería de interior vs exterior: En la Tabla 8 son reportados los resultados del experimento de reconocimiento de rostros que usan el conjunto de galería con imágenes capturadas en interior y el conjunto de test con imágenes de exterior. A la inversa, en la Tabla 9 son mostrados los experimentos que usan el conjunto de galería con imágenes de exterior y el conjunto de test con imágenes de interior. Se puede observar claramente que en esos casos todos los métodos bajo comparación disminuyen su comportamiento drásticamente comparado con los experimentos anteriores. En todos los casos la tasa de reconocimiento es muy baja. Interesante es el desempeño alcanzado por GJD-BC. Otra interesante observación es que SURF-M supera las variantes de SIFT-X-X, mostrando robustez en las variaciones de interior-exterior.

Expresiones faciales: En la Tabla 10 son reportados los resultados del experimento de reconocimiento de rostros que usan el conjunto de Expresión. La Tabla 10 muestra el promedio de la tasa de reconocimiento para 6 experimentos. En cada experimento un conjunto es elegido como galería y uno como test, luego se crea una tabla para la validación cruzada de todas las tasas de reconocimiento obtenidas y se calcula la media total para cada uno de los métodos bajo comparación. Los conjuntos usados son E1-E3 y V1-V3. Se puede observar que el mejor resultado es obtenido por las variantes SIFT-X-X, con una alta tasa de reconocimiento (~99.6%), lo que muestra su invariancia a expresiones faciales. El segundo mejor resultado es obtenido por LBP-X-X, GJD-BC y WLD-X-X, y el tercer mejor resultado es obtenido por el método SURF-M. Estos resultados son consistentes con los obtenidos con el conjunto de Palabras (ver Tabla 4), excepto para el caso de SURF-M. La principal razón acerca de la disminución del desempeño de SURF-M parece ser al proceso de calces debido a las grandes variaciones en las expresiones faciales.

Distintas distancias: En la Tabla 11 son mostrados los resultados de los experimentos de reconocimiento de rostros que corresponden a distintas distancias entre el sujeto y la cámara. Las distancias usadas son 1.2, 1.69, 2.4, 3.39 y 4.8 metros. En todos los experimentos la resolución de la imagen era reducida con un factor de distancia igual a $\sqrt{2}$. Por ejemplo, en el caso de GJD-BC, cuando la distancia es 1.69 m, la resolución de la imagen correspondiente es de 88x150 píxeles. De la Tabla 11, puede ser observado que todos los métodos, excepto para el caso de SURF-M, son robustos a variaciones entre la distancia del sujeto a la cámara. Los métodos más robustos a las variaciones de distancia son las variantes de SIFT-X-X y LBP-X-X, note que LBP opera muy bien con tamaños de imágenes pequeños, es decir que es útil al operar con distancias grandes. Parece ser que el método SURF-M tiene una resolución mínima para trabajar adecuadamente, se observa que luego de los 2.4 m el rendimiento del método SURF-M decrece bastante.

Tabla 4. Resultados de experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace. Sesión interior. Tasa de reconocimiento. Conjuntos de Rotación y Palabras. Las mejores variantes de cada método bajo comparación son presentadas en negritas.

Métodos	Rotación												Palabras	Promedio
	R1 [%]	R2 [%]	R3 [%]	R4 [%]	R5 [%]	R6 [%]	R7 [%]	R8 [%]	R9 [%]	R10 [%]	R11 [%]	Media [%]	Media [%]	[%]
LBP-EU-80 81x150	54.72	77.36	54.72	50.94	81.13	100	90.57	39.62	73.58	92.45	84.91	72.73	81.76	77.24
LBP-HI-80 81x150	73.58	88.68	73.58	64.15	96.23	100	96.23	50.94	90.57	100	96.23	84.56	92.45	88.51
LBP-XS-80 81x150	64.15	90.57	73.58	62.26	88.68	100	92.45	41.51	86.79	100	81.13	80.10	89.94	85.02
GJD-BC 81x150	66.04	96.23	75.47	50.94	92.45	100	88.68	33.96	81.13	98.11	69.81	77.53	98.11	87.82
WLD-EU-80 81x150	77.36	96.23	90.57	64.15	96.23	100	92.45	71.70	90.57	98.11	90.57	87.99	95.60	91.80
WLD-HI-80 81x150	73.58	96.23	88.68	64.15	94.34	100	96.23	66.04	90.57	100	92.45	87.48	95.60	91.54
WLD-XS-80 81x150	73.58	94.34	90.57	64.15	88.68	100	92.45	62.26	92.45	96.23	84.91	85.42	94.97	90.20
SIFT-L&R-M 81x150	86.79	96.23	84.91	73.58	96.23	100	92.45	66.04	92.45	98.11	83.02	88.16	94.97	91.57
SIFT-L&R-S 81x150	86.79	100	81.13	56.60	98.11	100	96.23	50.94	86.79	98.11	81.13	85.08	98.11	91.59
SIFT-Lowe-M 81x150	92.45	100	84.91	73.58	98.11	100	96.23	66.04	94.34	100	90.57	90.57	99.37	94.97
SURF-M 81x150	69.81	94.34	67.92	50.94	86.79	100	71.70	30.19	83.02	92.45	71.70	74.74	89.94	82.19
LBP-EU-80 100x185	45.28	75.47	41.51	45.28	81.13	100	69.81	30.19	77.36	90.57	66.04	65.69	81.76	73.73
LBP-HI-80 100x185	67.92	86.79	67.92	54.72	98.11	100	92.45	39.62	94.34	100	84.91	80.62	93.08	86.85
LBP-XS-80 100x185	71.70	90.57	62.26	56.60	96.23	100	94.34	33.96	90.57	100	81.13	79.76	91.82	85.79
GJD-BC 100x185	83.02	98.11	84.91	50.94	94.34	100	96.23	37.74	79.25	98.11	79.25	81.99	100	91.00
WLD-EU-80 100x185	79.25	90.57	79.25	54.72	88.68	100	94.34	47.17	86.79	100	84.91	82.33	91.82	87.08
WLD-HI-80 100x185	69.81	88.68	81.13	56.60	92.45	100	96.23	45.28	90.57	100	86.79	82.50	92.45	87.48
WLD-XS-80 100x185	81.13	92.45	84.91	60.38	96.23	100	92.45	49.06	90.57	98.11	94.34	85.42	92.45	88.94
SIFT-L&R-M 100x185	94.34	98.11	83.02	77.36	98.11	100	96.23	71.70	92.45	100	81.13	90.22	94.34	92.28
SIFT-L&R-S 100x185	86.79	98.11	86.79	60.38	100	100	94.34	47.17	90.57	100	77.36	85.59	97.48	91.54
SIFT-Lowe-M 100x185	86.79	100	96.23	79.25	98.11	100	96.23	75.47	94.34	100	94.34	92.80	100	96.40
SURF-M 100x185	79.25	98.11	71.70	64.15	96.23	100	88.68	49.06	84.91	100	83.02	83.19	97.48	90.34
LBP-EU-80 125x225	39.62	73.58	37.74	24.53	66.04	100	62.26	24.53	58.49	84.91	49.06	56.43	81.76	69.10
LBP-HI-80 125x225	58.49	84.91	58.49	37.74	86.79	100	83.02	39.62	79.25	96.23	75.47	72.73	86.16	79.44
LBP-XS-80 125x225	60.38	90.57	58.49	39.62	90.57	100	90.57	39.62	84.91	98.11	79.25	75.64	90.57	83.11
GJD-BC 125x225	84.91	98.11	84.91	49.06	96.23	100	92.45	41.51	86.79	100	83.02	83.36	99.37	91.37
WLD-EU-80 125x225	62.26	92.45	69.81	49.06	88.68	100	84.91	33.96	79.25	96.23	81.13	76.16	89.94	83.05
WLD-HI-80 125x225	58.49	88.68	71.70	47.17	90.57	100	92.45	41.51	88.68	100	86.79	78.73	93.08	85.91
WLD-XS-80 125x225	67.92	94.34	73.58	50.94	92.45	100	94.34	47.17	88.68	98.11	86.79	81.30	95.60	88.45
SIFT-L&R-M 125x225	86.79	100	94.34	73.58	100	100	94.34	73.58	96.23	100	94.34	92.11	98.11	95.11
SIFT-L&R-S 125x225	92.45	100	90.57	67.92	100	100	94.34	56.60	98.11	100	92.45	90.22	99.37	94.80
SIFT-Lowe-M 125x225	90.57	98.11	96.23	88.68	100	100	98.11	83.02	98.11	100	94.34	95.20	100	97.60
SURF-M 125x225	83.02	100	84.91	67.92	100	100	96.23	56.60	98.11	100	88.68	88.68	99.37	94.03

Tabla 5. Resultados de experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

Métodos	Rotación												Palabras	Promedio [%]
	R1 [%]	R2 [%]	R3 [%]	R4 [%]	R5 [%]	R6 [%]	R7 [%]	R8 [%]	R9 [%]	R10 [%]	R11 [%]	Media [%]	Media [%]	
LBP-HI-80 81x150	86.79	86.79	83.02	60.38	98.11	100	96.23	73.58	88.68	98.11	90.57	87.48	90.57	89.03
GJD-BC 125x225	62.26	98.11	71.70	39.62	88.68	100	94.34	35.85	60.38	96.23	67.92	74.10	86.79	80.45
WLD-EU-80 81x150	62.26	73.58	79.25	47.17	84.91	100	84.91	58.49	77.36	84.91	64.15	74.27	72.96	73.62
WLD-HI-80 81x150	69.81	86.79	92.45	60.38	96.23	100	94.34	66.04	84.91	96.23	79.25	84.22	96.23	90.23
SIFT-L&R-M 125x225	94.34	100	92.45	81.13	100	100	100	75.47	98.11	98.11	88.68	93.48	98.11	95.80
SIFT-Lowe-M 125x225	92.45	100	96.23	84.91	100	100	100	81.13	94.34	100	90.57	94.51	98.11	96.31
SURF-M 125x225	94.34	100	96.23	86.79	96.23	100	98.11	75.47	94.34	100	88.68	93.65	97.48	95.57

Tabla 6. Resultados experimentos de oclusión parcial, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

Métodos	Rotación												Palabras	Promedio [%]
	R1 [%]	R2 [%]	R3 [%]	R4 [%]	R5 [%]	R6 [%]	R7 [%]	R8 [%]	R9 [%]	R10 [%]	R11 [%]	Media [%]	Media [%]	
LBP-HI-80 81x150	60.38	77.36	62.26	58.49	96.23	100	90.57	47.17	83.02	100	81.13	77.87	86.79	82.33
GJD-BC 125x225	71.70	86.79	66.04	35.85	71.70	100	83.02	26.42	66.04	92.45	64.15	69.47	93.71	81.59
WLD-EU-80 81x150	75.47	96.23	86.79	62.26	96.23	100	92.45	71.70	84.91	98.11	90.57	86.79	93.08	89.94
WLD-HI-80 81x150	75.47	94.34	86.79	62.26	94.34	100	94.34	71.70	90.57	100	90.57	87.31	96.23	91.77
SIFT-L&R-M 125x225	75.47	100	86.79	64.15	98.11	100	92.45	66.04	90.57	98.11	88.68	87.31	93.87	90.59
SIFT-Lowe-M 125x225	86.79	98.11	96.23	73.58	92.45	100	94.34	67.92	94.34	100	88.68	90.22	98.78	94.50
SURF-M 125x225	71.70	98.11	73.58	49.06	94.34	100	88.68	43.40	84.91	96.23	73.58	79.42	95.60	87.51

Tabla 7. Resultados experimentos de oclusión parcial, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

Métodos	Rotación												Palabras	Promedio [%]
	R1 [%]	R2 [%]	R3 [%]	R4 [%]	R5 [%]	R6 [%]	R7 [%]	R8 [%]	R9 [%]	R10 [%]	R11 [%]	Media [%]	Media [%]	
LBP-HI-80 81x150	73.58	73.58	75.47	52.83	90.57	100	79.25	52.83	81.13	92.45	67.92	76.33	81.13	78.73
GJD-BC 125x225	49.06	79.25	52.83	24.53	81.13	100	77.36	24.53	56.60	94.34	49.06	62.61	83.02	72.82
WLD-EU-80 81x150	52.83	71.70	71.70	39.62	81.13	100	81.13	47.17	64.15	81.13	50.94	67.41	71.70	69.56
WLD-HI-80 81x150	73.58	88.68	90.57	60.38	94.34	100	90.57	66.04	86.79	92.45	79.25	83.88	87.42	85.65
SIFT-L&R-M 125x225	84.91	98.11	86.79	77.36	100	100	98.11	69.81	94.34	98.11	88.68	90.57	94.34	92.46
SIFT-Lowe-M 125x225	81.13	98.11	92.45	77.36	98.11	100	100	75.47	94.34	100	84.91	91.08	97.48	94.28
SURF-M 125x225	79.25	98.11	92.45	73.58	92.45	100	92.45	54.72	94.34	94.34	73.58	85.93	97.48	91.71

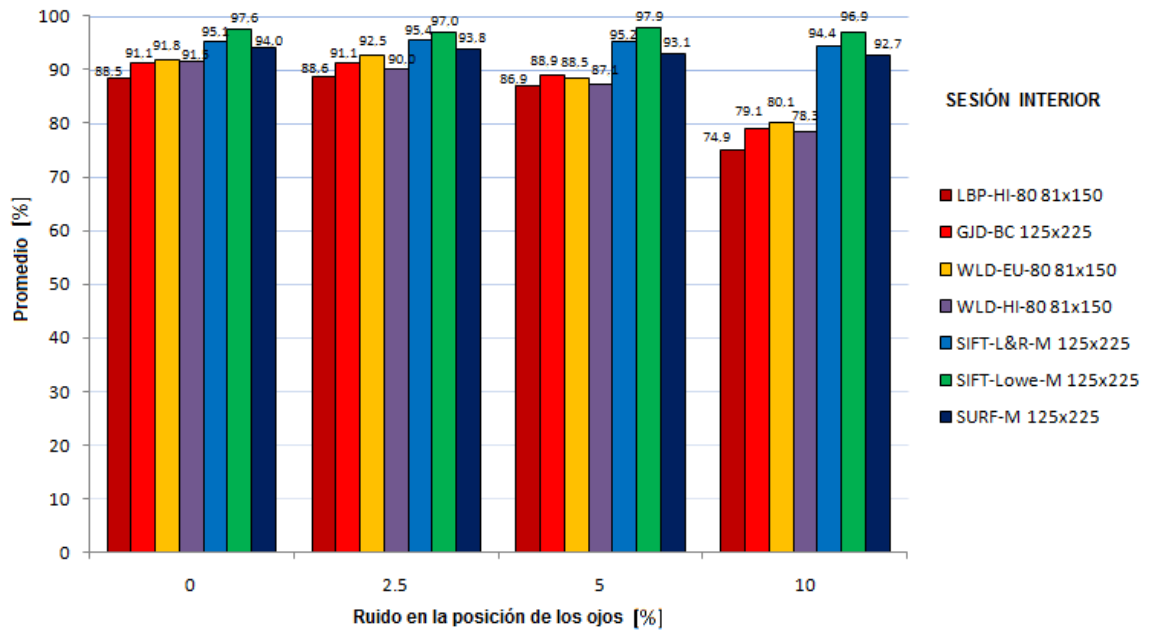


Figura 31. Resultados de experimentos de ruido en la posición de los ojos (0%, 2,5%, 5%, 10%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

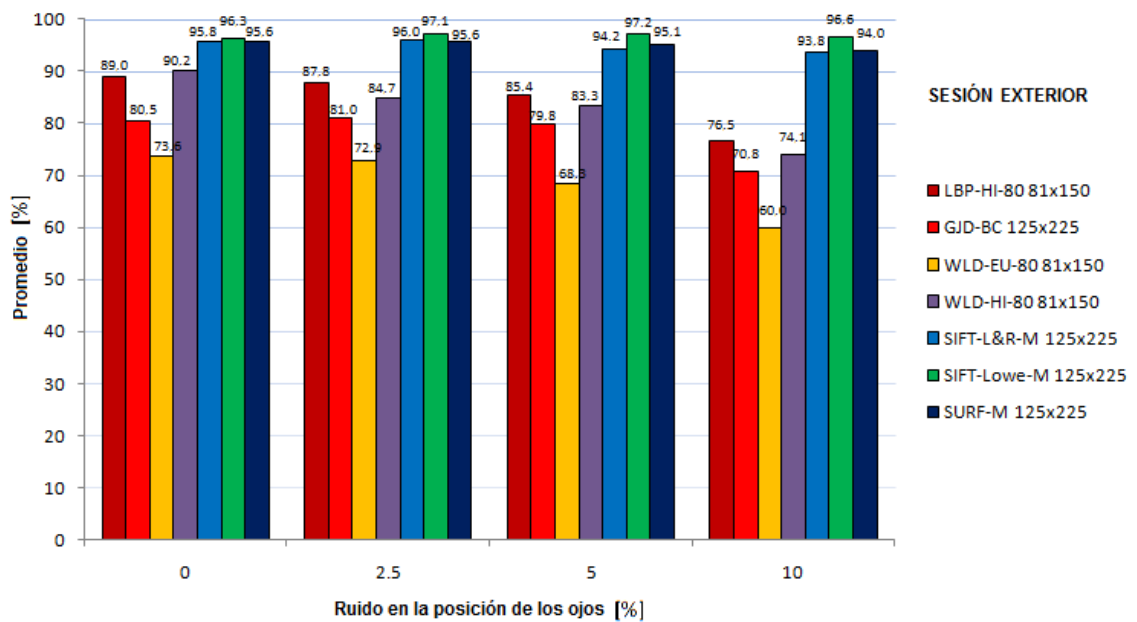


Figura 32. Resultados de experimentos de ruido en la posición de los ojos (2,5%, 5%, 10%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

Tabla 8. Resultados de experimentos de galería de interior vs exterior, base de datos UCHThermalFace. Mejores variantes de los métodos. Imágenes de interior para el conjunto de galería. Imágenes de exterior para el conjunto de test. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

Métodos	Rotación												Palabras Media [%]	Promedio [%]
	R1 [%]	R2 [%]	R3 [%]	R4 [%]	R5 [%]	R6 [%]	R7 [%]	R8 [%]	R9 [%]	R10 [%]	R11 [%]	Media [%]		
LBP-HI-80 81x150	24.53	30.19	16.98	20.75	32.08	32.08	16.98	9.43	22.64	20.75	11.32	21.61	22.01	21.81
GJD-BC 125x225	30.19	39.62	24.53	18.87	35.85	54.72	30.19	7.55	28.30	33.96	16.98	29.16	42.77	35.97
WLD-EU-80 81x150	22.64	28.30	16.98	9.43	15.09	20.75	11.32	13.21	16.98	16.98	11.32	16.64	22.01	19.33
WLD-HI-80 81x150	28.30	35.85	28.30	13.21	33.96	33.96	18.87	11.32	26.42	35.85	18.87	25.9	32.08	28.99
SIFT-L&R-M 125x225	9.43	7.55	15.09	9.43	16.98	13.21	9.43	5.66	7.55	11.32	5.66	10.12	11.32	10.72
SIFT-Lowe-M 125x225	13.21	13.21	9.43	0.00	15.09	16.98	11.32	1.89	5.66	11.32	3.77	9.26	14.47	11.87
SURF-M 125x225	24.53	30.19	22.64	11.32	24.53	35.85	13.21	7.55	22.64	26.42	5.66	20.41	22.01	21.21

Tabla 9. Resultados de experimentos de galería de exterior vs interior, base de datos UCHThermalFace. Mejores variantes de los métodos. Imágenes de exterior para el conjunto de galería. Imágenes de interior para el conjunto de test. Tasa de reconocimiento. Conjuntos de Rotación y Palabras.

Métodos	Rotación												Palabras Media [%]	Promedio [%]
	R1 [%]	R2 [%]	R3 [%]	R4 [%]	R5 [%]	R6 [%]	R7 [%]	R8 [%]	R9 [%]	R10 [%]	R11 [%]	Media [%]		
LBP-HI-80 81x150	13.21	15.09	3.77	11.32	24.53	24.53	13.21	7.55	24.53	24.53	15.09	16.12	20.13	18.13
GJD-BC 125x225	22.64	28.30	13.21	5.66	22.64	35.85	24.53	5.66	24.53	33.96	22.64	21.78	35.85	28.82
WLD-EU-80 81x150	15.09	20.75	13.21	15.09	30.19	24.53	20.75	11.32	22.64	22.64	15.09	19.21	25.79	22.50
WLD-HI-80 81x150	15.09	16.98	11.32	9.43	18.87	18.87	20.75	15.09	16.98	22.64	18.87	16.81	20.13	18.47
SIFT-L&R-M 125x225	18.87	20.75	18.87	7.55	11.32	20.75	9.43	5.66	9.43	16.98	9.43	13.55	16.04	14.80
SIFT-Lowe-M 125x225	9.43	7.55	11.32	3.77	9.43	15.09	11.32	7.55	13.21	9.43	7.55	9.61	16.35	12.98
SURF-M 125x225	7.55	18.87	7.55	3.77	13.21	24.53	15.09	3.77	15.09	24.53	11.32	13.21	27.04	20.13

Tabla 10. Resultados de experimentos de expresiones faciales, base de datos UCHThermalFace. Mejores variantes de los métodos. Tasa de reconocimiento. Conjuntos de Expresiones (E1-E3, V1-V3).

Métodos	Promedio [%]
LBP-HI-80	94.9
GJD-BC	94.4
WLD-EU-80	93.9
WLD-HI-80	94.1
SIFT-L&R-M	98.9
SIFT-Lowe-M	99.6
SURF-M	87.2

Tabla 11. Resultados de experimentos de distintas distancias, base de datos UCHThermalFace. Mejores variantes de los métodos. Tasa de reconocimiento. La resolución de la imagen es reducida con la distancia en un factor de $\sqrt{2}$.

Métodos	1.2 m	1.69 m	2.4 m	3.39 m	4.8 m
	Media [%]	Media[%]	Media[%]	Media[%]	Media [%]
LBP-HI-80. Resolución inicial 81x150	92.5	94.3	94.3	92.5	92.5
GJD-BC. Resolución inicial 125x225	100	100	100	94.3	88.7
WLD-EU-80. Resolución inicial 81x150	96.2	98.1	92.5	86.8	66.0
WLD-HI-80. Resolución inicial 81x150	96.2	96.2	94.3	94.3	67.9
SIFT-L&R-M. Resolución inicial 125x225	100	98.1	98.1	96.2	86.8
SIFT-Lowe-M. Resolución inicial 125x225	100	98.1	100	100	92.5
SURF-M. Resolución inicial 125x225	100	92.5	79.2	41.50	7.5

Las Figuras 33(a) y 33(b) muestran el ángulo de orientación (-30° , -15° , 0° , 15° , 30°) versus la tasa de reconocimiento obtenida por cada una de las metodologías seleccionadas y para rostros con un ángulo de elevación de 0° (un ángulo de orientación de 0° corresponde a una imagen frontal). Los ángulos de orientación (-30° , -15° , 0° , 15° , 30°) con un ángulo de 0° de inclinación corresponden a los conjuntos de Rotación {R4-R8}. Estos resultados resumen una parte de los resultados mostrados en las Tablas 4 y 5. Acá se pueden observar que para todas las metodologías el desempeño de los métodos decrece al aumentar el ángulo de orientación. Para bajas rotaciones ($\pm 15^\circ$), en el caso de sesión interior (Figura 33(a)), el desempeño de todas las metodologías es muy similar, con las variantes de la metodología de SIFT obteniendo los mejores resultados, mientras que en el caso de la sesión en exterior, GJD-BC y WLD-EU-80 claramente muestran un bajo desempeño en comparación con los otros métodos, los cuales tienen una similar tasa de reconocimiento. Para largas rotaciones ($\pm 30^\circ$), en el caso de la sesión interior, sólo SIFT-Lowe tiene un buen desempeño, mientras que en el caso de la sesión exterior, ambas variantes de SIFT, SIFT-Lowe y SIFT-L&R, tienen un desempeño de aproximadamente un 80% en la tasa de reconocimiento. Todos los demás métodos muestran un peor desempeño para grandes rotaciones, con WLD-HI-80 trabajando un poco mejor que los otros métodos, y en ambas configuraciones de interior y exterior.

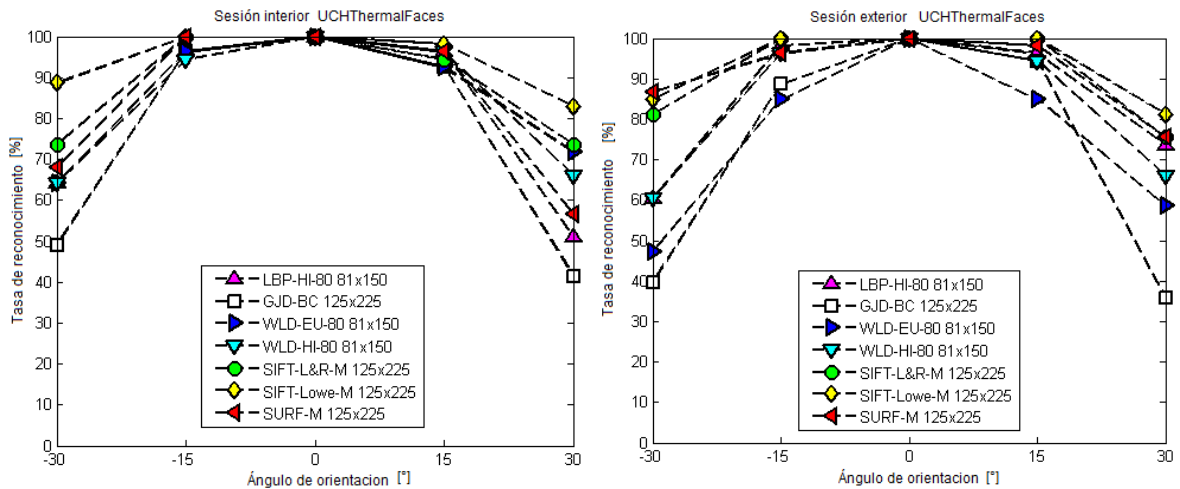


Figura 33. Ángulo de elevación versus tasa de reconocimiento para los métodos seleccionados. El ángulo horizontal corresponde a la posición R4, R5, R6, R7 y R8 de la Figura 25. (a) sesión interior; (b) sesión exterior.

3.5 Desempeño computacional

La velocidad en el proceso de reconocimiento es un factor importante en muchas aplicaciones de reconocimiento de rostros, por ejemplo en aplicaciones de interacción Robot-Humano o en la identificación usando grandes bases de datos de rostros. Por esta razón, se presenta un análisis comparativo de los métodos seleccionados en términos del tiempo de procesamiento. Para realizar esto, se ha evaluado el tiempo que requiere la extracción de características (FET-Feature Extraction Time), el tiempo requerido para calzar dos vectores de características (MT-Matching Time), y el tiempo total de procesamiento (PT- Processing Time) requerido para reconocer el rostro independientemente del tamaño de la base de datos. Note que en el caso de GJD, el tiempo total de procesamiento no es lineal sino de $n \log n$ del tamaño de la base de datos. Esto es debido a la manera en que trabaja la clasificación de cuenta de Bourda. Todos los otros métodos son lineales al tamaño de la galería. Note además que en este análisis se ha considerado sólo el tiempo requerido durante la operación, y no el tiempo requerido para crear la base de datos, ni el tiempo de cálculo de la extracción vascular ni de la fusión, donde estos tiempos son menores al milisegundo.

Los experimentos son llevado a cabo en un computador corriendo Windows 7 Ultimate (64-bits) con un procesador Intel Core 2 duo CPU T5870 @2.00 GHz GHz (4GB RAM). Para todos los métodos, con excepción de SIFT-Lowe, se uso una implementación de C/C++ compilada con una aplicación de 32-bits. Para SIFT-Lowe se uso la implementación² original en MATLAB, usando MATLAB 7.10.0 (R2010a). Si se tuviera una implementación de SIFT-Lowe en C/C++ el tiempo de procesamiento debería ser similar al tiempo obtenido por SIFT-L&R.

La Tabla 12 muestra el tiempo de procesamiento calculado para todos los métodos bajo comparación en términos de extracción de características, tiempo de calces y el tiempo total de procesamiento. En términos de extracción de características, LBP-X-80 es el método más rápido,

² <http://www.cs.ubc.ca/~lowe/keypoints/>

seguido por WLD-X-80. El tercer método más rápido es SURF-M. El cuarto método más rápido es GJD-BC, que tiene un tiempo de extracción de características dos veces más rápido que SIFT-L&R-M, tres veces más rápido que SIFT-Lowe-M y más de un orden de magnitud más lento que LBP-HI-80 y WLD-X-80.

En términos de tiempo de calces (tiempo por pares de imágenes), los métodos más rápidos son WLD-X-80, GJD-BC, LBP-X-80 y SURF-M, todos ellos con un tiempo de calces (MT) menor a 1 ms. El método SIFT-Lowe es el método más lento, con un tiempo de calces de más de 8 veces más lento que WLD-X-80, GJD-BC y LBP-X-80, y 4 veces más lento que SIFT-L&R-M. Cuando se considera el tiempo total de procesamiento de todos los métodos, el método con el tiempo más pequeño de procesamiento es LBP-X-80, independiente del tamaño de la base de datos. El segundo método más rápido son las variantes WLD-X-80. Para grandes bases de datos (ejemplo 1000 imágenes en galería), WLD-EU-80 y WLD-HI-80 son casi 2 veces más lentos que LBP-X-80. El método más lento es SIFT, con la variante SIFT-L&R-M siendo de 1.5 a 4 veces más rápido que SIFT-Lowe-M, dependiendo del tamaño de la base de datos. Para pequeñas bases de datos, SIFT-Lowe-M es el método más lento, es de dos órdenes de magnitud más lento que LBP-X-80, mientras que para grandes bases de datos se convierte en 70 veces más lento que LBP-X-80. El método GJD-BC es de 2 a 10 veces más rápido que las variantes de SIFT dependiendo del tamaño de la base de datos, y de 4 a 20 veces más lento que LBP-X-80. El método SURF es más rápido que las variantes de SIFT para todos los tamaños de la base de datos. En resumen, existe una clara distinción en cuanto a la velocidad de los métodos, con LBP y WLD siendo los métodos más rápidos, el método de GJD-BC y SURF en posición intermedia, y las variantes de SIFT siendo los métodos más lentos.

Tabla 12. Tiempo de procesamiento. Tiempo medido en milisegundos. Los experimentos son llevados a cabo en un computador con un procesador Intel Core 2 duo CPU T5870 @2.00 GHz (4GB RAM). FET/MT: tiempo de extracción de características/ tiempo de calces [ms]. PT: Tiempo total de procesamiento [ms]. Tamaños de la base de datos desde 1, 10, 100, y 1000 rostros.

Métodos	FET [ms]	MT [ms]	PT (FET+MT) [ms]			
			1	10	100	1000
LBP-X-80 81x150	2.6	< 1	3.6	3.6	12.9	122.3
GJD-BC 125x225	95.94	< 1	96.9	96.9	127.9	532.9
WLD-EU-80 81x150	3.9	< 1	4.9	5.9	23.9	202.9
WLD-HI-80 81x150	3.9	< 1	4.9	6.9	26.9	229.9
WLD-XS-80 81x150	3.9	< 1	4.9	4.9	16.9	130.9
SIFT-L&R-M 125x225	189.01	2.2537	191.26	211.55	414.38	2442.7
SIFT-Lowe-M 125x225	290.22	8.4386	298.66	374.61	1134.1	8728.8
SURF-M 125x225	36.00	<1	37.0	46.0	136.0	1036.0

3.6 Discusión y conclusión del capítulo

En este capítulo se presentó un estudio comparativo basado en el análisis de distintos métodos de reconocimiento de rostro en el espectro térmico. Los métodos analizados fueron seleccionados considerando la conveniencia en requerimientos predefinidos, tales como: operación en tiempo real, sólo una imagen por persona, operación en línea (sin entrenamiento), y un comportamiento robusto en condiciones de ambientes no controlados, además de su desempeño otros estudios. El estudio comparativo fue llevado a cabo usando dos bases de datos: Equinox y UCHThermalFace. La base de datos Equinox, fue usada como una línea de comparación, se realizaron experimentos para comparar sus resultados con los resultados presentados en trabajos previos basados en imágenes obtenidas bajo condiciones controladas. La base de datos UCHThermalFace incluye aspectos tales como rotaciones con diferentes ángulos de orientación e inclinación, variaciones en la condición del ambiente (interior/exterior) y expresiones faciales. Además, fue simulada una mala alienación de las imágenes y fue simulado también una oclusión parcial en las imágenes. Los métodos bajo comparación usados fueron: histogramas de LBP, descriptores jets de Gabor, descriptores SIFT, histogramas de WLD y SURF. Comentarios de los principales resultados de este estudio y algunas conclusiones son mostradas a continuación:

Comentarios respecto al tamaño de la ventana: Aunque los resultados de [46] son para imágenes visibles, la dependencia de los métodos con el tamaño de la ventana de la imagen no es grande y para los mejores métodos analizados, como SIFT y WLD, el efecto es más bien bajo. Esto parece ser debido a que la mayoría del fondo observado es uniforme para las imágenes térmicas, lo cual permite a los métodos de manera adicional usar más información acerca del contorno de la cara.

Comentarios acerca del alineamiento, oclusiones y expresiones: De los experimentos se puede concluir que en gran medida sólo algunos de los métodos analizados son robustos a alineamientos imprecisos, oclusión de rostros y a variaciones en la expresión. Aceptando estos factores que afectan el proceso de reconocimiento, su influencia en el desempeño de los métodos es mucho más bajo que en condiciones de exterior o variaciones en la pose.

Comentarios con respecto a las condiciones interior/exterior: Muchos de los métodos tienen comportamiento muy bueno en condiciones naturales de interior, como también en condiciones de exterior, con excepción de GJD-BC, el cual tiene un desempeño que disminuye considerablemente bajo condiciones de exterior. Este aspecto podría ser analizado con experimentos adicionales. En los experimentos donde las imágenes de test son adquiridas en una configuración de exterior y las imágenes de galería son adquiridas en interior, o viceversa, el desempeño de todos los métodos es muy bajo. La razón parece ser a la diferencia en el rango de los píxeles de las imágenes térmicas adquiridas en interior comparadas con las imágenes adquiridas con la configuración en exterior. La saturación observada en las imágenes adquiridas en condiciones de exterior, debido al calor en el ambiente, puede ser la principal razón para esto. Esto eventualmente podría mejorarse con una mejor calibración de la cámara, o por el uso de algoritmos de normalización, como en el caso de imágenes en el espectro visible [67]. Además la propia condición de los sujetos, es decir su temperatura corporal va cambiando dependiendo del ambiente de interior o exterior donde se encuentre el individuo. Estos aspectos de las condiciones de interior y exterior son analizados en el Capítulo 5 del presente trabajo.

Conclusiones acerca del desempeño de los métodos. La pregunta, ¿cuál de todos los métodos bajo análisis es el mejor? es muy difícil de responder. Sin embargo, se puede decir que los métodos basados en histogramas de WLD son una excelente elección si es necesaria la operación en tiempo real como bien sus buenos resultados en las tasas de reconocimiento. WLD es el segundo método más rápido, y su desempeño es mejor que LBP, cual es el método más rápido. Si el tiempo de procesamiento no es una consideración importante, es claro que las variantes de SIFT deberían ser el método a usar, dado por la robustez a errores de alineación, a rotaciones y a cambios de expresiones faciales. En particular, SIFT-L&R parece ser una buena elección para bases de datos de tamaño medianas o grandes, dado su corto tiempo de procesamiento en vez que SIFT-Lowe y su razonable buen desempeño en esos casos.

Se demuestra el hecho que las variantes de SIFT sean un método a elegir para algunas configuraciones es uno de los principales resultados de este trabajo, además hasta ahora los métodos basados en SIFT no son ampliamente usados en reconocimiento de rostros con imágenes visibles ni térmicas.

4 Estudio Comparativo de Reconocimiento de Rostros usando *VascularThermalFace*

4.1 Introducción

El objetivo de este capítulo es presentar una nueva metodología de reconocimiento de rostros que combina la utilización de características vasculares del rostro humano y la intensidad de píxeles de la imagen del rostro térmico. Esta metodología, llamada *VascularThermalFace*, permite obtener una nueva representación del rostro térmico, la que aplicada en métodos de reconocimiento de rostros, como los estudiados en el Capítulo 3, permite aumentar el desempeño y la tasa de reconocimiento a los sistemas de reconocimiento de rostros.

La nueva metodología presentada, *VascularThermalFace*, permite el uso combinado de características fisiológicas obtenidas de un algoritmo de extracción de la red vascular del rostro térmico (*VascularNetwork*), el cual fue diseñado en [53], en conjunto con la intensidad de píxeles de la imagen de rostro térmico obtenido por medio de una cámara térmica. Los autores de la metodología de *VascularNetwork* [53] describen que la red vascular extraída es única y repetible, y que puede ser usada como una característica biométrica de cada sujeto. Además la red vascular al estar debajo de la piel humana es difícil de alterar por lo que sería una característica biométrica robusta para realizar el reconocimiento de rostros.

El algoritmo *VascularNetwork*, diseñado para la extracción de la red vascular, considera varias etapas de procesamiento. La primera consiste en la aplicación de una etapa de segmentación del rostro térmico del resto de la imagen, con el fin de eliminar impurezas y zonas ajenas al rostro, y concentrar toda la atención en éste. Luego una etapa de filtrado de difusión, con el fin de obtener una mejora en las zonas donde se concentra más el calor. Al realizar este filtrado se realizan estas posiciones en el rostro donde se concentra el calor. Luego se utiliza una serie de operaciones morfológicas a la imagen filtrada, aplicadas en cadena, con el fin de localizar la red vascular superficial del rostro térmico. Finalmente al sustraer la imagen filtrada con esta imagen procesada morfológicamente es posible determinar dónde se encuentra la red vascular.

El algoritmo *VascularNetwork* es evaluado cualitativamente en este capítulo por medio de un análisis visual, con el cual es posible observar la representación de la red vascular para zonas en la que ésta es visible en el cuerpo humano. Además se estudia la invariancia temporal que poseen los resultados de la aplicación de este algoritmo, mostrando que la red vascular obtenida por el algoritmo *VascularNetwork* se mantiene en el tiempo. Se estudia y analiza el comportamiento de los métodos de reconocimiento de rostros seleccionados del Capítulo 3 con imágenes de red vascular, indicando qué métodos obtienen ventajas y desventajas al usar el enfoque *VascularNetwork*.

La nueva metodología *VascularThermalFace* es evaluada y validada utilizando la base de datos térmica UCHThermalFace. Se realizan experimentos similares a los efectuados en el Capítulo 3 con el fin de analizar las mejoras producidas al combinar los algoritmos de reconocimiento de rostros escogidos en el Capítulo 3 con la nueva metodología *VascularThermalFace* y validar el uso de ésta.

4.2 Algoritmo de extracción de la red vascular (*VascularNetwork*) [53]

El algoritmo *VascularNetwork* fue diseñado en [53] y es una herramienta fundamental para extraer la red vascular de un rostro térmico. Este algoritmo permite extraer la red vascular de las imágenes térmicas y utiliza 2 partes fundamentales: segmentación del rostro y extracción de la red vascular. El algoritmo *VascularNetwork* explicado a continuación muestra una modificación del algoritmo original [53], en la etapa de segmentación del rostro térmico. En [53] se utiliza el algoritmo EM (*Expectation Maximization*) para entrenar un modelo bayesiano y estimar los parámetros para realizar la segmentación, mientras que en esta implementación se utiliza un modelo de distribución de probabilidad de piel y no piel basado en las ideas de la publicación [72].

Segmentación del rostro térmico: Se crea un modelo de detección de piel térmica modelado cómo una distribución de probabilidad de piel térmica y una distribución de probabilidad de no piel térmica. Ambas distribuciones de probabilidad son modelados como distribuciones Gaussianas:

$$P_{PIEL\ TERMICA}(x) = \frac{1}{\sqrt{2\pi}\sigma_S} \exp\left(-\frac{1}{2} \frac{(x - \mu_S)^2}{\sigma_S^2}\right) \quad (17)$$

$$P_{NO-PIEL\ TERMICA}(x) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{1}{2} \frac{(x - \mu_N)^2}{\sigma_N^2}\right) \quad (18)$$

donde x representa el valor del pixel en la posición (i, j) de la imagen, μ_S y μ_N corresponden a las medias de piel térmica y no piel térmica respectivamente, y σ_S y σ_N son las desviaciones estándar de piel térmica y no piel térmica respectivamente. Los valores de las medias y de las desviaciones estándar son calculados con anterioridad por medio de un entrenamiento con muchas imágenes que poseen piel térmica y no piel térmica, a las cuales se les calculan los valores de media y desviaciones estándar.

La aplicación de este esquema para las imágenes térmica permite el cálculo de una máscara térmica, la cual es obtenida de la siguiente manera:

$$I_{MASCARA}(i, j) = \begin{cases} 1 & \text{if } P_{PIEL\ TERMICA}(I(i, j)) > P_{NO-PIEL\ TERMICA}(I(i, j)) \\ 0 & \text{if } P_{PIEL\ TERMICA}(I(i, j)) < P_{NO-PIEL\ TERMICA}(I(i, j)) \end{cases} \quad (19)$$

La imagen segmentada es obtenida por la multiplicación de la imagen de rostro térmica y la máscara obtenida anteriormente:

$$I_{SEGMENTADA}(i, j) = I_{TERMICA}(i, j)I_{MASCARA}(i, j) \quad (20)$$

Es importante destacar que parámetros generales para realizar la segmentación no pueden ser calculados debido a que dependen de la respuesta de cada cámara térmica. En este trabajo se obtuvieron los siguientes valores para las distribuciones de probabilidad de piel térmica y no piel térmica: $\mu_S=163.95$, $\mu_N=33.45$, $\sigma_S=64.85$ y $\sigma_N=41.10$, utilizando una cámara térmica FLIR TAU 320 (ver especificaciones en la Sección 2.3). Estos valores fueron obtenidos en una etapa offline de adaptación de parámetros (entrenamiento). Se tomaron muestras de imágenes térmicas de sujetos que poseían piel térmica, como rostros y manos, y de objetos fríos y calientes que no poseían piel térmica, como latas de bebida, ropa de individuos, etc. Se marcaron cuadrados en las imágenes manualmente donde había piel térmica y no había piel térmica. De esta manera para todos los píxeles marcados de piel térmica y no piel térmica, se calculan las medias y desviaciones estándar para obtener los valores mostrados anteriormente.

Extracción vascular: Se realiza un filtrado de difusión a la imagen segmentada con el fin de resaltar las zonas donde se encuentra la mayor información de calor, la cual es dada por la siguiente ecuación [53]:

$$I_{t+1}(i, j) = I_t(i, j) + \frac{1}{4} * (c_{N,t}(i, j)\nabla I_{N,t}(i, j) + c_{S,t}(i, j)\nabla I_{S,t}(i, j) + c_{E,t}(i, j)\nabla I_{E,t}(i, j) + c_{O,t}(i, j)\nabla I_{O,t}(i, j)) \quad (21)$$

Los 4 coeficientes de difusión y 4 gradientes corresponden a las 4 direcciones (norte, sur, este y oeste) con respecto a la posición (i, j) . Los coeficientes de difusión son todos calculados de la misma manera, siguiendo el siguiente ejemplo para el coeficiente en la dirección norte (c_N):

$$c_{N,t} = \exp\left(-\frac{\nabla I_{N,t}^2(i, j)}{k^2}\right) \quad (22)$$

donde $\nabla I_{N,t}(i, j) = I_t(i, j+1) - I_t(i, j)$ y k es una constante.

A la imagen filtrada se le aplica una serie de operaciones morfológicas para determinar la red vascular, que es relativamente baja en contraste, en comparación con el tejido de alrededor. Se usa primero una combinación de operadores morfológicos, donde se usa una etapa de erosión seguida de una dilatación. Este procedimiento es llamado “segmentación *Top Hat*”. El “*Top Hat*” es una técnica empleada para mejorar el brillo o “zonas calientes” de un objeto, las cuales, en este caso, corresponden a la red vascular. El proceso del “*Top Hat*” es mostrado a continuación [53]:

$$I_{ABIERTO} = (I_{FILTRADO} \ominus S) \oplus S \quad (23)$$

$$I_{VASCULAR} = I_{FILTRADO} - I_{ABIERTO} \quad (24)$$

donde \ominus y \oplus , corresponden a los operadores erosión y dilatación respectivamente, S corresponde a un elemento estructural de 3x3 definido por:

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad (25)$$

Luego se aplica un umbral a la imagen vascular resultante con el fin de obtener una imagen binaria para la posición de las venas. La Figura 34 muestra el resultado de aplicar el algoritmo de extracción de la red vascular al rostro de un sujeto.

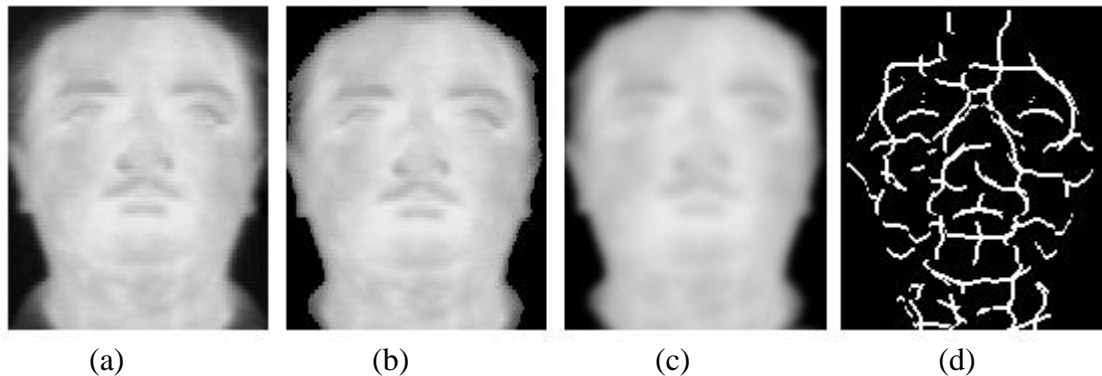


Figura 34. Ejemplo del proceso de extracción de la red vascular *VascularNetwork* para un sujeto. (a) Imagen original, (b) Imagen segmentada, (c) Imagen filtrada difusa y (d) Red vascular extraída usando la segmentación “Top Hat”.

4.3 Análisis de la red vascular y métodos de reconocimiento

En este apartado se realiza un análisis de la red vascular de manera cualitativa, mostrando ejemplos y situaciones que demuestran el uso de la metodología *VascularNetwork*. Se realizan diferentes experimentos para demostrar gráficamente que el algoritmo *VascularNetwork* obtiene la información de la red vascular del rostro humano. Se utiliza la zona del antebrazo de un humano para validar la metodología debido a que es posible visualizar directamente la red vascular en este caso. Luego se presenta un análisis basado en la invariancia temporal que posee la metodología, con un conjunto de imágenes obtenidas durante un largo periodo de tiempo. Además se realiza un análisis visual para cada uno de los métodos de reconocimiento seleccionados en el Capítulo 3, con el fin de analizar las ventajas y desventajas de utilizar esta representación en conjunto con éstos. Finalmente se muestran los factores que limitan el desempeño de esta metodología de *VascularNetwork*, tales como la configuración de la cámara térmica, el uso de lentes en los individuos, las rotaciones en las imágenes y casos extremos como transpiración, rostros mojados, etc.

4.3.1 Estudio de la red vascular superficial

Se presenta un análisis visual para demostrar que el algoritmo de extracción de la red vascular *VascularNetwork* permite determinar de manera exacta la ubicación de la red vascular

superficial. Para este análisis se utiliza un antebrazo y una mano, partes del cuerpo humano donde es posible visualizar la red vascular sin la utilización de elementos externos. La Figura 35 y Figura 36 muestran el resultado de aplicar el algoritmo *VascularNetwork* a un brazo, donde es posible visualizar en la imagen térmica la red vascular del antebrazo.

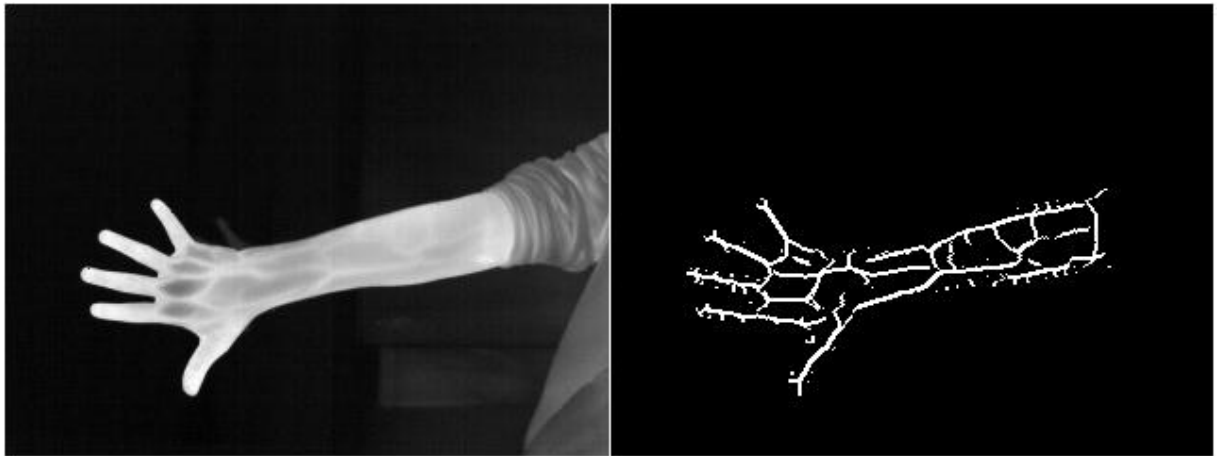


Figura 35. Aplicación de la metodología *VascularNetwork* a una imagen térmica de un brazo. Izquierda: Imagen térmica de un antebrazo. Derecha: red vascular, metodología *VascularNetwork*.

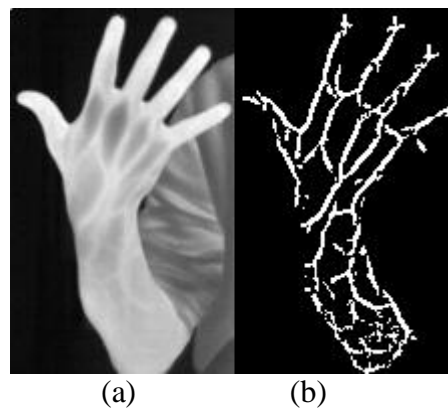


Figura 36. Aplicación de la metodología *VascularNetwork* a una imagen térmica de una mano. (a) Imagen térmica de una mano. (b) red vascular obtenida al aplicar la metodología *VascularNetwork*.

Es posible observar de las Figuras 35 y 36 que la metodología *VascularNetwork* permite determinar la posición de la red vascular superficial, validando cualitativamente de esta manera el algoritmo presentado en [53]. Además, esta metodología nos permite determinar adecuadamente y con una buena precisión la ubicación de la red vascular, lo cual hace que esta metodología sea una excelente alternativa para realizar reconocimiento basado en la red vascular.

4.3.2 Estudio de la invariancia temporal de la red vascular

Como fue observado en el apartado anterior, el uso de la metodología *VascularNetwork* permite determinar la red vascular de alguna parte del cuerpo humano utilizando una cámara

térmica. La pregunta ahora es si esta red vascular es capaz de mantenerse en el tiempo con el fin de ser usada como una característica robusta para su utilización en sistemas de reconocimiento de rostros. Por este motivo, es que este apartado tiene la finalidad de demostrar la invariancia temporal que posee el uso de la metodología *VascularNetwork*.

Para demostrar la invariancia temporal de la red vascular obtenida con la metodología *VascularNetwork*, se utilizan imágenes de la base de datos UCHThermalTemporalFace (que será explicada con más detalle en el Capítulo 5). Se utilizan 2 sujetos de esta base de datos para demostrar la invariancia temporal, se usan las primeras 28 imágenes, de las 50 adquiridas en diferentes periodos de tiempo, correspondientes a un lapso de tiempo entre el día 1 y el día 35, de 69 días que posee la base de datos UCHThermalTemporalFace. Para demostrar la invariancia temporal, se utiliza la primera imagen de un sujeto obtenida de la base de datos versus las 28 imágenes del conjunto. La idea es demostrar que no existe invariancia en la red vascular obtenida durante este lapso temporal para la primera imagen versus todas las demás. El análisis se efectúa dibujando la primera imagen del sujeto en color rojo, luego las imágenes adquiridas en los siguientes instantes de tiempo se dibujan en color azul. La superposición de ambas se dibuja en color magenta y representa las características que se mantienen en el tiempo. La Figura 37 muestra la invariancia vascular obtenida para el sujeto número 3 de la base de datos UCHThermalTemporalFace. La Figura 38 muestra la invariancia vascular obtenida para el sujeto número 5 de la base de datos UCHThermalTemporalFace.

Más aún, para revisar la invariancia temporal de la red vascular obtenida con la metodología *VascularNetwork*, se utiliza la imagen obtenida en la segunda sesión versus las primeras 4 imágenes y las últimas 4 imágenes de la base de datos UCHThermalTemporalFace, y luego se utiliza la penúltima imagen versus las primeras 4 imágenes y las últimas 4 imágenes de la base de datos. Las últimas 4 imágenes corresponden a la sesión 47, 48, 49 y 50 de la base de datos UCHThermalTemporalFace, es decir se evalúa la invariancia hasta el último día de la sesión, día 69. Si la red vascular obtenida de las primeras imágenes versus las últimas imágenes de la base de datos UCHThermalTemporalFace se mantiene constante, se demostraría que la red vascular es invariante en el tiempo. Los resultados de este análisis visual son mostrados en las Figuras 39 y 40 para el sujeto 3, y en las Figuras 41 y 42 para el sujeto 5.

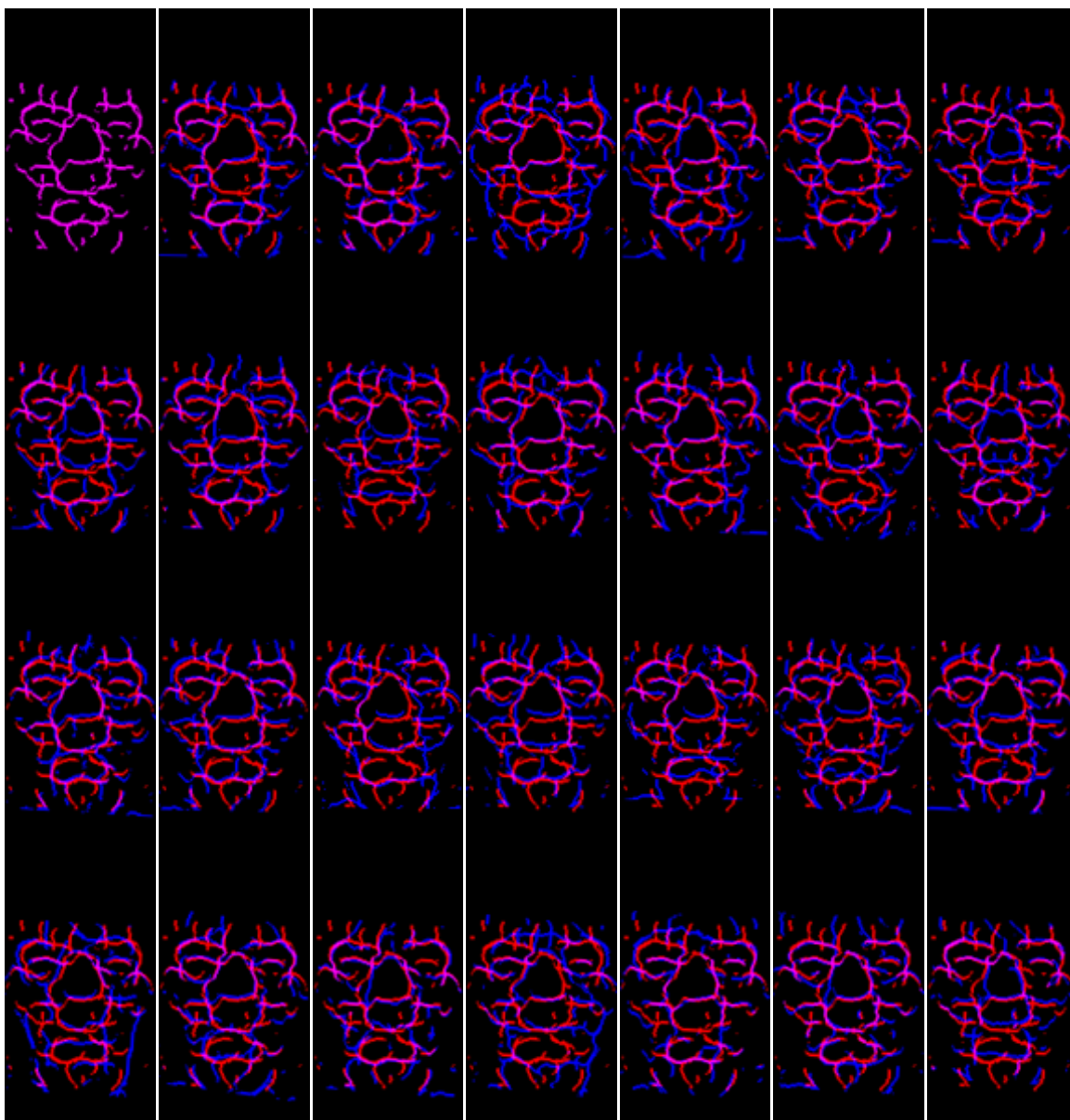


Figura 37. Imágenes de red vascular para el sujeto 3 de la base de datos UCHThermalTemporalFace. Color rojo indica imagen de galería en la primera sesión. Color azul indica imagen de test de la sesión (1-28). Color magenta indica pixeles rojo y azul en la misma posición. Se compara la imagen de la primera sesión versus las 28 sesiones siguientes.

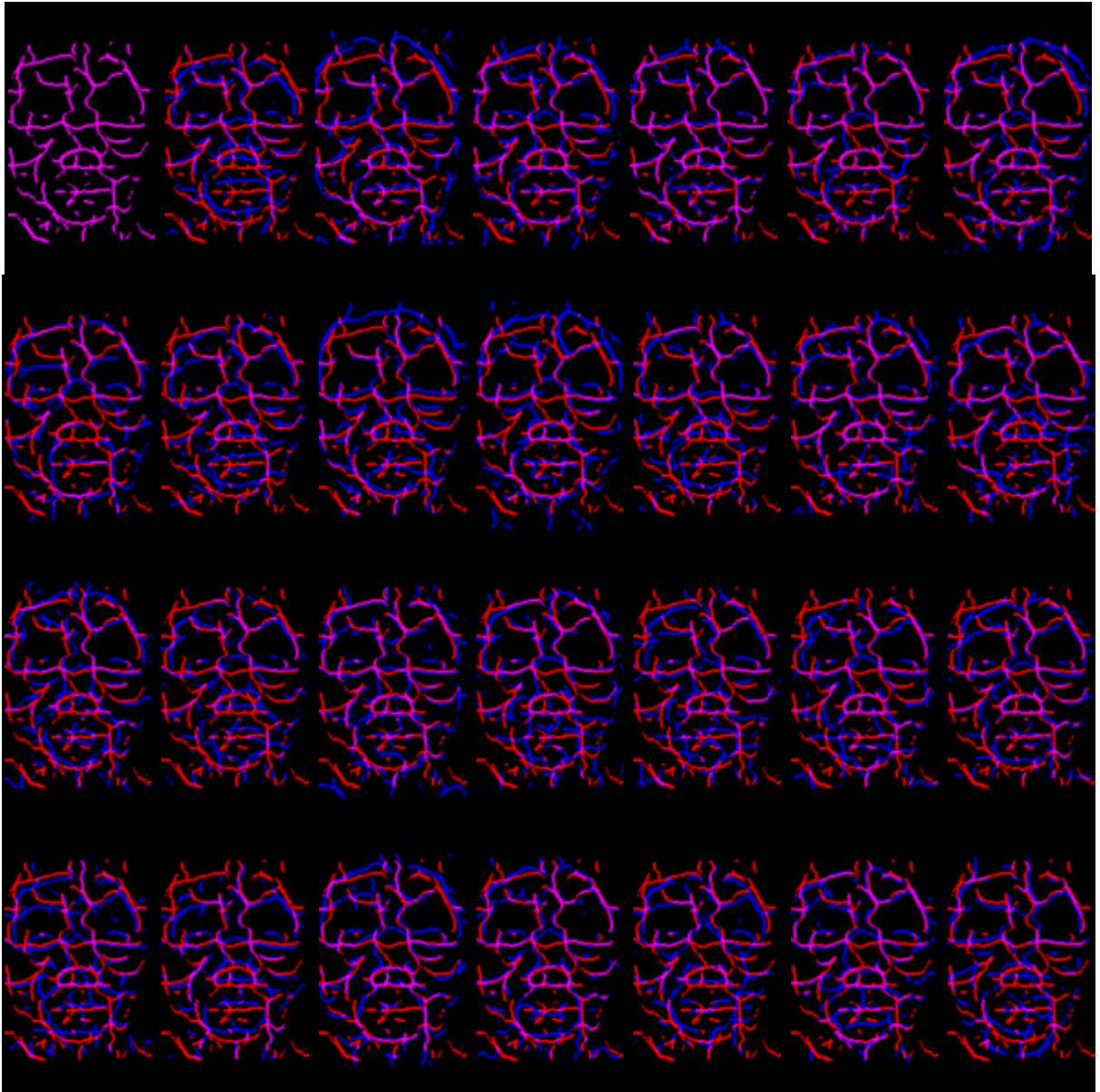


Figura 38. Imágenes de red vascular para el sujeto 5 de la base de datos UCHThermalTemporalFace. Color rojo indica imagen de galería en la primera sesión. Color azul indica imagen de test de la sesión (1-28). Color magenta indica pixeles rojo y azul en la misma posición. Se compara la imagen de la primera sesión versus las 28 sesiones siguientes.

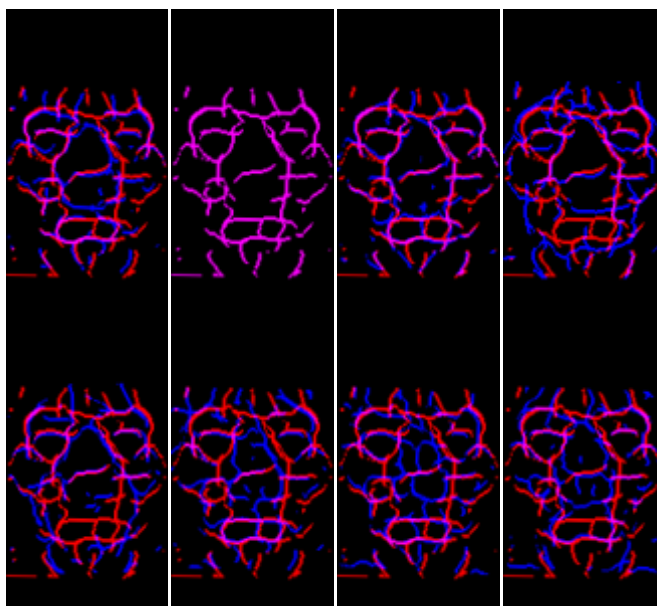


Figura 39. Imágenes de red vascular para el sujeto 3 de la base de datos UCHThermalTemporalFace. Color rojo indica imagen de galería en la primera sesión. Color azul indica imagen de test de la sesión. Color magenta indica pixeles rojo y azul en la misma posición. Arriba primeras 4 sesiones. Abajo últimas 4 sesiones. Se compara la imagen de la segunda sesión versus las 4 primeras y 4 últimas sesiones.

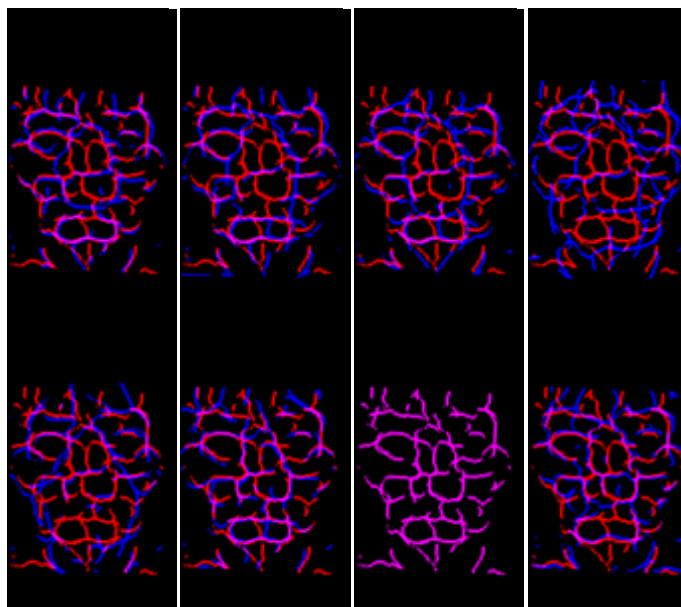


Figura 40. Imágenes de red vascular para el sujeto 3 de la base de datos UCHThermalTemporalFace. Color rojo indica imagen de galería en la primera sesión. Color azul indica imagen de test de la sesión. Color magenta indica pixeles rojo y azul en la misma posición. Arriba primeras 4 sesiones. Abajo últimas 4 sesiones. Se compara la imagen de la penúltima sesión versus las 4 primeras y 4 últimas sesiones.

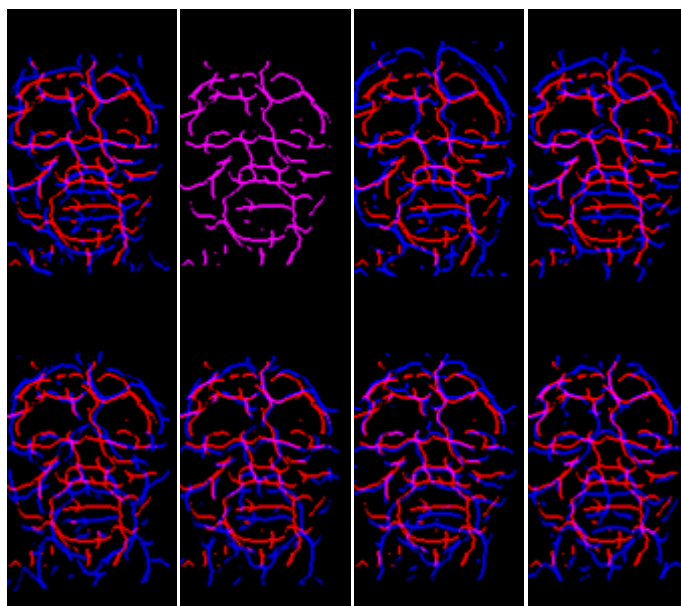


Figura 41. Imágenes de red vascular para el sujeto 5 de la base de datos UCHThermalTemporalFace. Color rojo indica imagen de galería en la primera sesión. Color azul indica imagen de test de la sesión. Color magenta indica pixeles rojo y azul en la misma posición. Arriba primeras 4 sesiones. Abajo últimas 4 sesiones. Se compara la imagen de la segunda sesión versus las 4 primeras y 4 últimas sesiones.

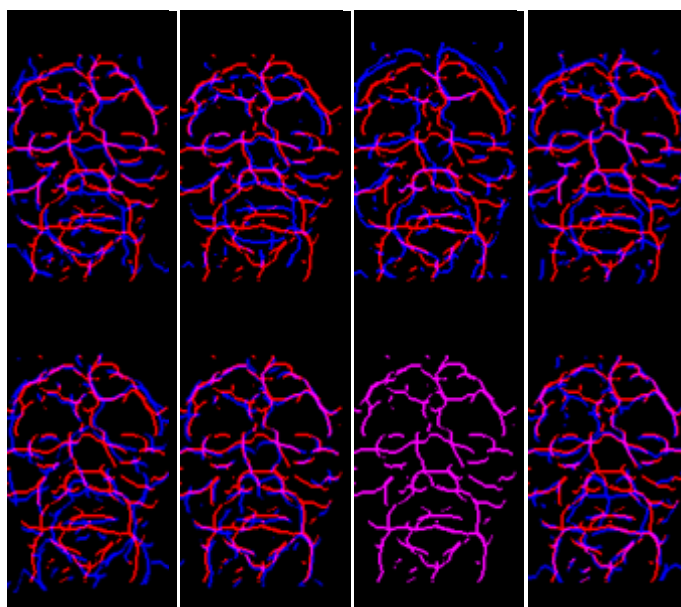


Figura 42. Imágenes de red vascular para el sujeto 5 de la base de datos UCHThermalTemporalFace. Color rojo indica imagen de galería en la primera sesión. Color azul indica imagen de test de la sesión. Color magenta indica pixeles rojo y azul en la misma posición. Arriba primeras 4 sesiones. Abajo últimas 4 sesiones. Se compara la imagen de la penúltima sesión versus las 4 primeras y 4 últimas sesiones.

Es posible visualizar de las Figuras 37 y 38, que la metodología *VascularNetwork* entrega una característica que se mantiene en el tiempo, que es observada como una estructura que se

mantiene durante las todas las sesiones de tiempo. Comparando las Figuras 39 y 41, se observa que la red vascular obtenida en la segunda sesión se mantiene para las 4 imágenes finales obtenidas después de varios meses de trabajo. Además, analizando las Figuras 37, 38, 39, 40, 41, y 42 se aprecia que existe un patrón de forma que es diferente entre los 2 individuos, el cual puede ser usado para discernir entre éstos. También, se puede apreciar que existen algunas imágenes donde la estructura de la red vascular es similar pero no coincide con la red de la primera sesión, lo cual se debe primordialmente a variaciones en la captura de la muestra y en el comportamiento del cuerpo humano, producto del calor corporal, que afecta la captura de la imagen térmica por su termorregulación. Sin embargo, aunque estos patrones se encuentren movidos, corresponden a la misma distribución de la red vascular. Otro elemento que es posible de ver en las Figuras (37-42), es que en algunas imágenes existen zonas donde la red vascular desaparece, esto también es producto de la captura de la cámara térmica debido a variaciones en el ambiente y condición corporal. La imagen térmica va variando en el tiempo, lo cual hace que en algunas zonas aparezca más o menos red vascular dependiendo de las condiciones ambientales y de las condiciones del propio individuo, sin embargo la red vascular se mantiene en el tiempo en la mayoría de los casos. Los factores producidos por la captura de cámara térmica serán analizados en la Sección 4.3.4.

Se concluye que es posible utilizar la metodología *VascularNetwork* para obtener la red vascular del rostro térmico y con ello generar una característica biométrica que es invariante (en la mayoría de los casos) en el tiempo.

4.3.3 Estudio de los métodos de reconocimiento utilizando la red vascular

¿Es posible utilizar los métodos de reconocimiento de rostros obtenidos del Capítulo 3 para realizar un sistema de reconocimiento robusto de rostros que use la información de la red vascular? Para resolver esta pregunta, este apartado analiza cada uno de los métodos seleccionados del Capítulo 3 utilizando la metodología *VascularNetwork*.

Se realiza un análisis visual del punto de vista de cómo operan los métodos de reconocimiento de rostros utilizados anteriormente con el fin de obtener una decisión acerca de que métodos podrían resultar buenos para trabajar con la metodología *VascularNetwork*.

Histogramas LBP: El histograma LBP convierte los valores de intensidad de los píxeles, comparando el pixel central con sus vecinos de una región de 3x3. Esto quiere decir que al utilizar una imagen binaria de la red vascular, el método LBP crearía una nueva representación de la red vascular, idéntica a la red vascular, la cual podría ser utilizada para realizar el reconocimiento de rostros. Esta representación es mostrada en la Figura 43. Luego la etapa de histogramas ayuda al método LBP a crear una representación que tenga la información completa de la red vascular del rostro. A continuación, la etapa de calces puede ser realizada de la misma manera que la efectuada en el Capítulo 3. Por lo tanto, es posible decir que el método histograma LBP puede ser utilizado para el reconocimiento de la red vascular.

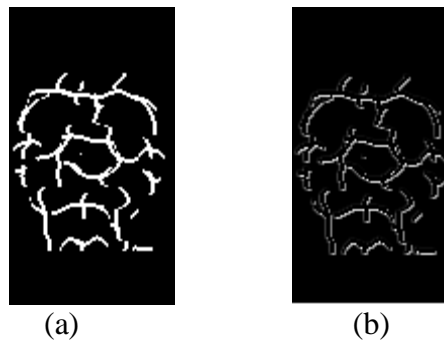


Figura 43. Representación del método LBP para la red vascular. (a) Red vascular. (b) representación LBP para la Figura 43(a).

Descriptores Jets de Gabor: El método de descriptores jets de Gabor, calcula los jets de Gabor en puntos de la imagen espaciados en una longitud de onda establecida. Si se utiliza una imagen de la red vascular, la cual es binaria, y se aplica el método de Gabor, los jets pueden ser calculados en puntos que no contienen información relevante de la imagen de venas o en puntos donde no existe información (píxeles con valor 0), debido a que la imagen es binaria y a que el método opera calculando los jets en cada punto de una grilla dada. Por este motivo, el método de descriptores jets de Gabor no sería una buena herramienta para realizar el reconocimiento de rostros con imágenes de red vascular.

Histogramas de WLD: El método WLD calcula la información de excitación de los píxeles de la imagen y la orientación. Si se utiliza una imagen de red vascular, el método WLD convierte la imagen de red vascular en una nueva representación de excitación de la red vascular y calcula la orientación de los píxeles de la imagen de la red vascular. La representación de excitación es similar a la representación de intensidades del método LBP. Luego el método opera combinando ambas características de excitación y orientación, generando un histograma para realizar la representación total de la red vascular del rostro. Luego el proceso de calces es realizado utilizando alguna distancia de disimilitud. De este análisis es posible ver que el método histogramas de WLD puede ser utilizado para el reconocimiento de la red vascular.

Descriptores SIFT: El método SIFT permite encontrar descriptores invariantes a cambios de rotación, traslación y escala basados en la obtención de puntos de interés. Un ejemplo de los puntos de interés obtenidos para el caso de la red vascular de un sujeto es mostrado en la Figura 44. Se observa la gran cantidad de puntos de interés producidos por el método SIFT en la Figura 44. La Figura 45 y la Figura 46 muestran algunos calces realizados por SIFT para diferentes imágenes de red vascular. Es posible ver que en la Figura 45, el método SIFT realiza una buena etapa de calces con la imagen de test, mostrando que la imagen de red vascular puede ser reconocida por medio de su estructura. Sin embargo, en la Figura 46 se observa que la etapa de calces es realizada con demasiados calces falsos, lo cual hace que el sistema no permita el reconocimiento de la estructura del sujeto de test. En base al análisis de estas imágenes es posible concluir que el buen desempeño de método SIFT depende exclusivamente de la etapa de calces, debido a que la red vascular genera muchos descriptores, los cuales pueden ser calzados equívocamente.

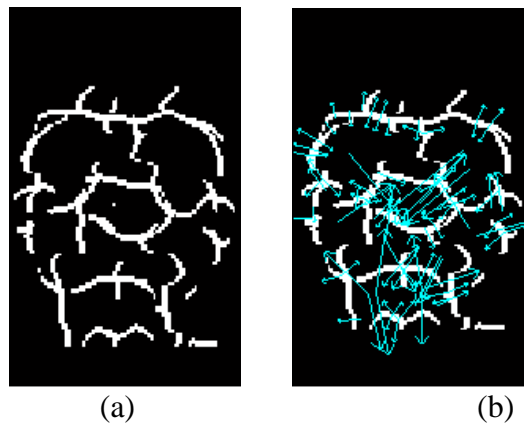


Figura 44. Descriptores obtenidos con el método SIFT. (a) Imagen de red vascular. (b) Descriptores red vascular de la Figura 44(a) por medio del método SIFT.

Descriptores SURF: Al igual que los descriptores SIFT, los descriptores SURF obtienen descriptores locales para representar la imagen de la red vascular por medio de sus puntos de interés adquiridos en base a la metodología SURF. Es por esto, que se puede concluir que al igual que SIFT, este método debería funcionar adecuadamente cuando no existan problemas en la etapa de calces producto de calces falsos. La Figura 47 muestra los calces obtenidos por el método SURF para los mismos casos mostrados en la Figura 46, realizados por el método SIFT. Se puede ver de la Figura 47 que los calces realizados con el método SURF funcionan mejor que los calces mostrados en la Figura 46, y que son realizados con SIFT, obteniendo mejor resultado en la localización de la estructura más parecida. Sin embargo en la Figura 47(c), se observan calces falsos, lo cual hace suponer que el comportamiento de los calces de SIFT también está presente para los calces SURF, sólo que en este ejemplo el resultado es mejor para SURF.

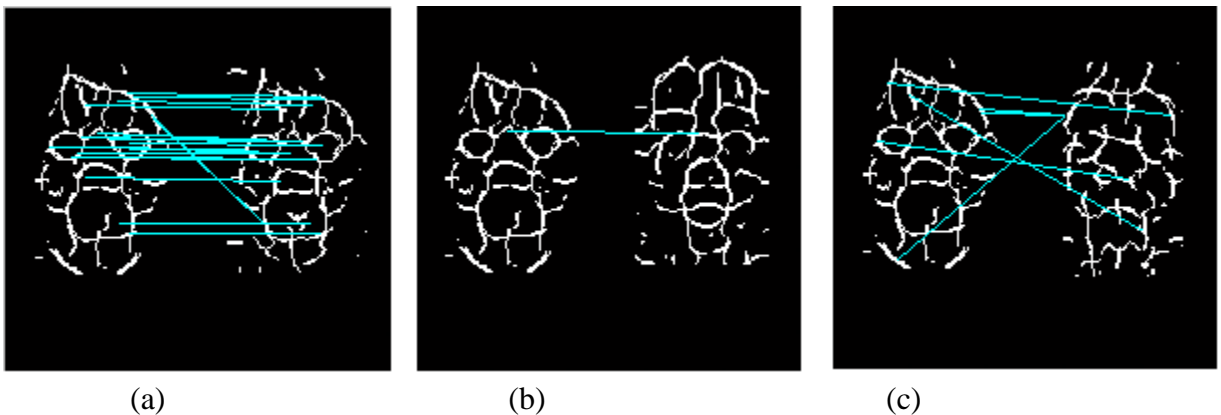


Figura 45. Calces obtenidos con el método SIFT. (a) calces obtenidos para el mismo sujeto de test y de galería. (b) y (c) calces obtenidos para distintos sujetos de test y de galería.

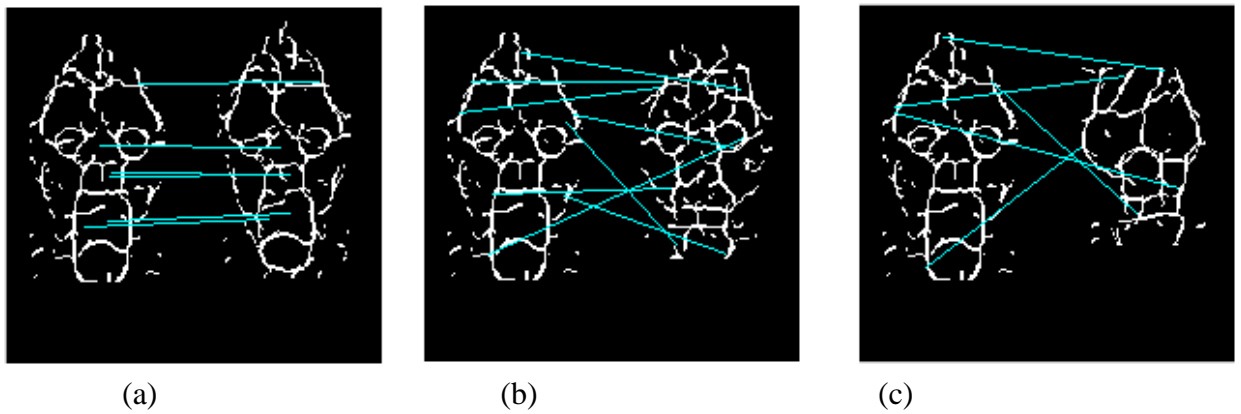


Figura 46. Calces obtenidos con el método SIFT. (a) calces obtenidos para el mismo sujeto de test y de galería. (b) y (c) calces obtenidos para distintos sujetos de test y de galería. Se observan muchos calces falsos.

Como conclusión del análisis de los métodos de reconocimiento de rostros y el uso de la imagen con la información de la red vascular, es notorio que la mayoría de los métodos pueden trabajar adecuadamente en conjunto con la metodología de *VascularNetwork*, excepto el método de descriptores jets de Gabor. Los métodos basados en calces locales, tales como LBP y WLD convierten la imagen de red vascular a una nueva representación basada en intensidad de píxeles (método LBP), o utilizando su información de intensidad-excitación y orientación (método WLD), además de usar un esquema basado en histogramas de regiones, los que son concatenados para obtener una representación completa del rostro. Esto facilitaría la etapa de reconocimiento, producto que se buscan pocos patrones en cada región, debido a que es sólo la información de la red vascular. Los métodos basados en calces globales, como SIFT y SURF, obtienen muchos descriptores de las imágenes de la red vascular del rostro, lo cual es bueno si se quiere representar una imagen a través de descriptores. Sin embargo, sin una etapa adecuada de calces, muchos descriptores permiten que se realicen calces errados o falsos, y esto provocaría que el desempeño de esta clase de métodos disminuiría.

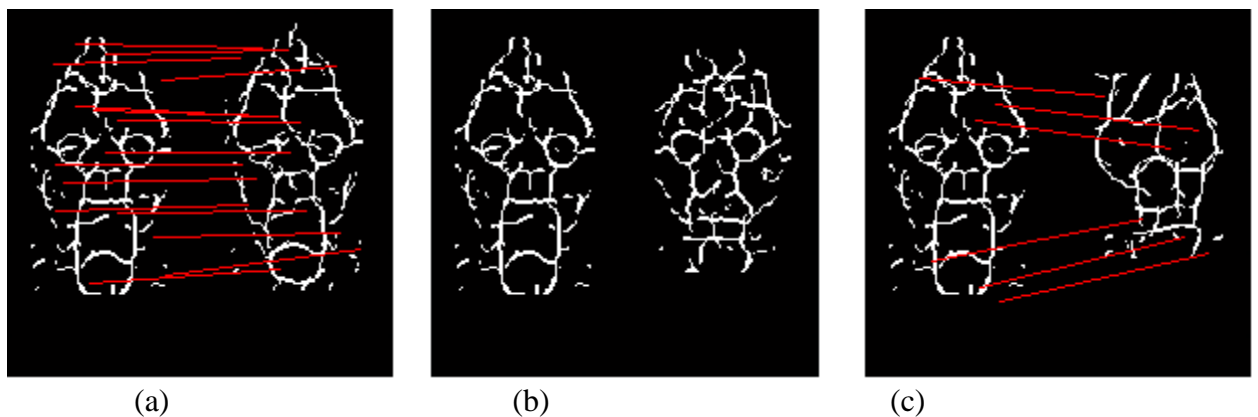


Figura 47. Calces obtenidos con el método SURF. (a) calces obtenidos para el mismo sujeto de test y de galería. (b) y (c) calces obtenidos para distintos sujetos de test y de galería. Se observan calces incorrectos en la imagen (c).

4.3.4 Factores que limitan la metodología VascularNetwork

Diferentes factores afectan la captura de la imagen térmica y por ende afectan el comportamiento de la metodología *VascularNetwork*. Algunos factores que deterioran el comportamiento de esta metodología han sido estudiados con anterioridad en algunas publicaciones [8] [9][14][53], tales como el uso de lentes, casos de transpiración y arduo ejercicio, la diferencia de tiempo en la adquisición de imágenes térmicas, factores que limitan el comportamiento de los sistemas de reconocimiento de rostros. Sin embargo, factores como la configuración de la cámara térmica y rotaciones en las imágenes no han sido puestas en consideración cuando se ha estudiado la metodología *VascularNetwork*.

Por este motivo, este apartado tiene por finalidad mostrar algunas limitantes y soluciones en la captura de la imagen térmica, como también mostrar los efectos que se obtienen ante mala calibración de la cámara térmica, efecto del uso de lentes, rotaciones y otros casos, donde es posible observar factores negativos para el uso de la metodología *VascularNetwork*.

Configuración de la cámara térmica: La forma de captura de la cámara térmica depende de diferentes parámetros que posee cada cámara térmica como el control de brillo y contraste de las imágenes térmicas. Además, las cámaras térmicas cuentan con un proceso de calibración interno, el cual depende de la cantidad de frames o imágenes con los cuales se quiera realizar este proceso de calibración. El proceso de calibración trata de corregir la respuesta de cada uno de los detectores del arreglo térmico (*Focal Plane Array*), con el fin de obtener una respuesta uniforme en cada uno de los detectores térmicos. Además el punto de operación de las cámaras térmicas varía producto del calentamiento interno de sus detectores, lo cual hace que la respuesta de la cámara térmica varíe en el tiempo. Entonces ¿Qué configuración de la cámara térmica usar? Esta es una pregunta difícil cuya respuesta depende de la aplicación y del tipo de cámara térmica a usar. Sin embargo, basándonos en la respuesta y sensibilidad térmica de la cámara térmica usada (FLIR TAU 320) y la aplicación que deseamos realizar, en este caso de reconocimiento de rostros, es necesario realizar un análisis para el control de parámetros a usar en este trabajo.

Se estudiaron varios parámetros y metodologías que posee la cámara térmica con la finalidad de mantener la imagen térmica lo más estable posible, es decir que la visualización de la imagen sea constante y no varíe demasiado sus niveles de brillo y contraste. Como es posible de recordar, en el Capítulo 3, se comentó que la configuración de la cámara usada para la obtención de la base térmica UCHThermalFace se realizó con parámetros fijos de brillo y contraste, obtenidos para maximizar la información del rostro. Esta elección trajo algunos problemas de saturación en las imágenes adquiridas en la sesión exterior (outdoor), producto del calor y el cambio del punto de operación de la cámara térmica. Es por este motivo, que se realiza un nuevo análisis para obtener la mejor representación en base a los requerimientos de visualización constante de la imagen térmica. Se estudiaron los parámetros de brillo y contraste de 2 maneras: usando una configuración de parámetros fijos, dados por el autor del trabajo, y parámetros automáticos, dados por la cámara térmica en base a un algoritmo eficiente de eualización del histograma de la escena.

- Parámetros de brillo y contraste fijos: Esta representación permite seleccionan los parámetros de brillo y contraste de manera manual con el fin de obtener una visualización adecuada de la escena. Tal como se uso en el Capítulo 3, se trata de maximizar la

información del rostro térmico con el fin de poder analizar en forma más detallada las características que posee éste. Sin embargo, la imagen con parámetros fijos funciona adecuadamente para sujetos que tengan una temperatura similar, pero si la condición de calor corporal varía, la representación de la imagen térmica varía, por ende la representación cambia. Además, si la cámara térmica varía su punto de operación por motivos ambientales o motivos de calentamiento de sus propios sensores, la imagen térmica también se ve afectada. Ejemplos de imágenes térmicas de 2 sujetos obtenidos con parámetros fijos en sesiones de interior y exterior son mostrados en la Figura 48. Se aprecia que para parámetros fijos de brillo y contraste, la imagen térmica obtenida tiene una representación distinta.

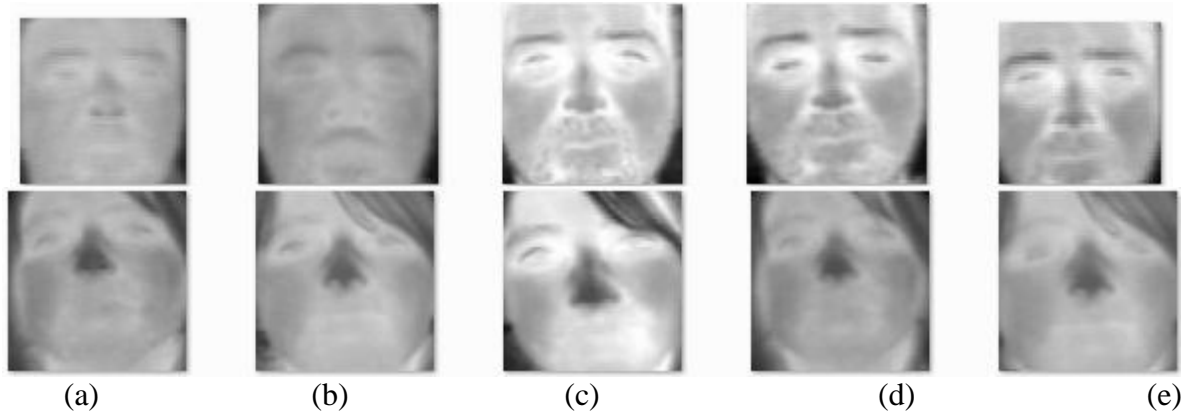


Figura 48. Ejemplos de imágenes térmicas obtenidas con la configuración de parámetros fijos. (a)-(b)-(c) muestran imágenes térmicas obtenidas en sesiones de interior; (d)-(e) corresponden a imágenes obtenidas en sesiones de exterior. Se observa como varía su representación entre las imágenes de interior-exterior.

- **Parámetros de brillo y contraste automáticos:** Para este caso, la cámara térmica realiza el control de los parámetros de brillo y contraste por medio de un algoritmo de ecualización de histograma de la escena térmica, con el fin de mantener la visualización de la escena estable. Si la escena varía, los parámetros de brillo y contraste también lo hacen, pero se adaptan al ajuste de histograma manteniendo la escena de manera uniforme. Sin embargo, si el fondo es variable, esto hace que los valores también varíen, generando resultados no deseados. En cambio, si se usa un fondo fijo, se puede obtener la misma representación de la imagen térmica. Ejemplos de imágenes térmicas obtenidas con parámetros automáticos, ajuste de histograma y fondo fijo para sesiones de interior y exterior, son mostradas en la Figura 49. Es posible observar de esta representación que las imágenes tienden a ser más estables, manteniendo sus niveles de brillo y contraste. En comparación con las imágenes observadas en la Figura 48, en las cuales varían mucho sus niveles de tonos de gris.

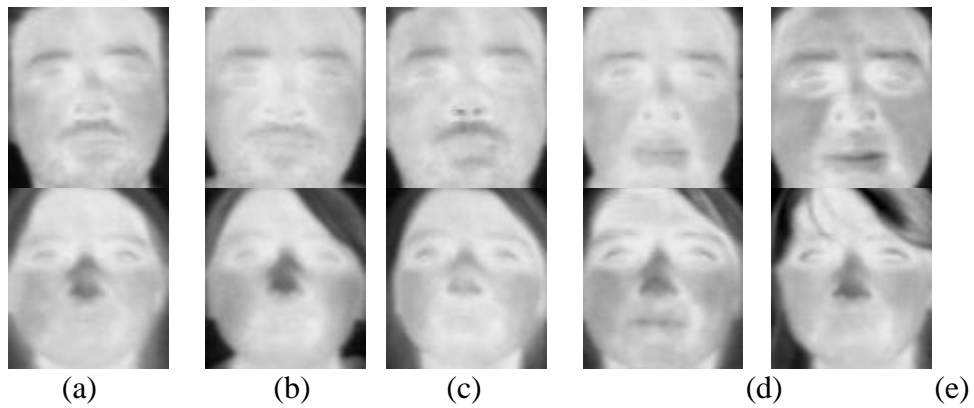


Figura 49. Ejemplos de imágenes térmicas obtenidas con la configuración de parámetros automáticos. (a)-(b)-(c) muestran imágenes térmicas obtenidas en sesiones de interior; (d)-(e) corresponden a imágenes obtenidas en sesiones de exterior. Se observa que su representación de interior-exterior varía menos en comparación con la Figura 48.

Es importante comentar que además de este análisis de parámetros fijos y automáticos se trabajó en ideas del control automático de la cámara térmica y métodos basados en conversión de intensidad de píxeles a temperatura, con el fin de mantener la cámara térmica estable y obtener de esta forma la misma representación constante para las imágenes térmicas adquiridas. Además se usaron técnicas basadas en corrección de intensidad de píxeles de las imágenes, como análisis de histogramas, normalización de píxeles y post procesamiento de la imagen térmica basada en modelo de “*Blood Perfusion*” [73]. Estas ideas finalmente no se consideraron debido a que los resultados experimentales no dieron resultados relevantes.

Al analizar si la red vascular extraída cambiaba producto de la representación obtenida con los parámetros fijos o automáticos, es fácil ver que si la imagen térmica cambiaba con los parámetros fijos, la red vascular también variaba producto del cambio de contraste y brillo que realizaba la escena. Por este motivo se decidió usar la configuración de parámetros automáticos, manteniendo la escena con un fondo fijo para tener uniformidad en la captura de las imágenes. Si bien las imágenes no se mantienen del todo uniformes producto de cambios corporales del individuo y del ambiente, es una buena opción para mantener la imagen térmica uniforme. En trabajos a futuro se espera seguir trabajando en el control de la cámara térmica, para poder visualizar sin variaciones de brillo ni contraste la escena y solucionar el problema del cambio de punto de operación de la cámara térmica.

Efectos del uso de lentes: Los lentes o anteojos son frecuentemente usados por los humanos en el diario vivir. Si se observa la representación que se obtiene de un sujeto utilizando lentes con una imagen térmica, es posible visualizar la pérdida de información en la zona donde se encuentra el lente o anteojos debido a que los lentes no permiten el paso de la radiación infrarroja, opacando la zona donde se utiliza este material. La Figura 50 muestra este efecto, donde se observa a un sujeto utilizando lentes y su correspondiente imagen de red vascular. Es posible observar en la Figura 50, que el uso de lentes limita el desempeño de la metodología *VascularNetwork* ya que no identifica zonas de red vascular, donde se utilizaban los lentes. Sin embargo, el reconocimiento de rostros dependerá exclusivamente del algoritmo de reconocimiento de rostros a usar, debido a que esta pérdida de información puede ser considerada

como una oclusión. Observando los resultados con respecto del Capítulo 3, los métodos estudiados se comportan adecuadamente cuando existen zonas del rostro ocluidas.

Efecto de rotaciones: Considerando el efecto de la rotación en el rostro térmico, es posible mencionar que afecta cuando se quiere reconocer rostros utilizando el algoritmo *VascularNetwork*. Como se observó en el Capítulo 3, en el reconocimiento de rostros utilizando la base UCHThermalFace, que posee imágenes de rotaciones, algunos métodos de reconocimiento son afectados más por las rotaciones que otros. Si utilizamos una representación basada en el algoritmo *VascularNetwork*, la cual corresponde a una estructura de la red vascular, esta estructura cuando es rotada tiende a perder la forma original de la estructura de la red vascular. Un ejemplo de esta situación puede ser visto en la Figura 51, la cual muestra un conjunto de 5 rotaciones para un sujeto y corresponde a un ejemplo del conjunto de rotación de la base UCHThermalFace.

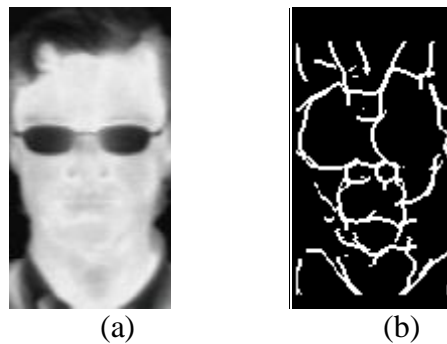


Figura 50. Imagen térmica de un sujeto con lentes. (a) Imagen térmica. (b) Imagen de red vascular. Se observa la pérdida de información producto de la utilización del lente.

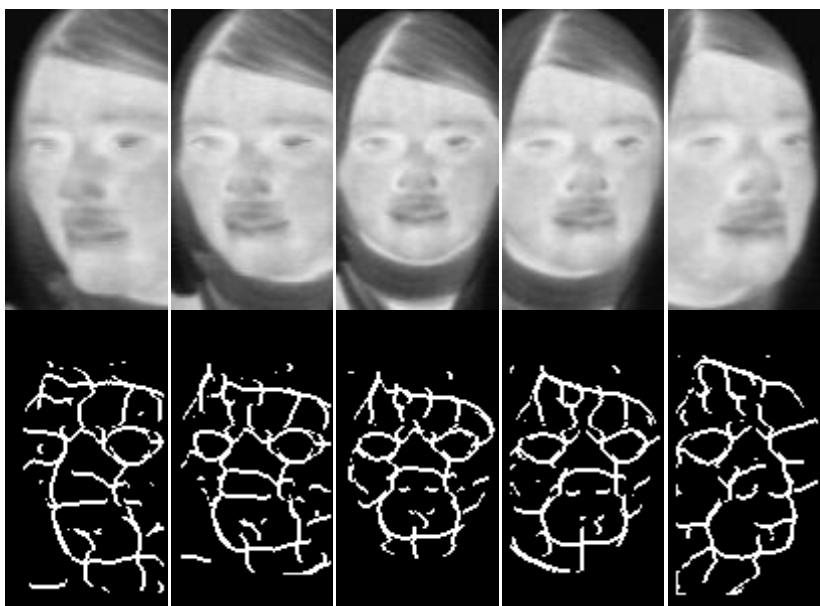


Figura 51. Imágenes térmicas y de red vascular para un sujeto del conjunto de rotación de la base UCHThermalFace. Arriba: Imágenes térmicas. Abajo: Imágenes de red vascular.

Otros casos: Algunos factores como la transpiración [53] y la lluvia en el rostro humano, son factores que limitan el desempeño de sistemas de reconocimiento de rostros térmicos y más aún cuando se utiliza un enfoque basado en el algoritmo *VascularNetwork*. Estas condiciones son adversas y no son solucionables por la cámara térmica, obteniendo resultados como el mostrado en la Figura 52, donde se puede apreciar un ejemplo de un sujeto con el rostro húmedo. Se observa que las características vasculares no aparecen totalmente, producto que la imagen térmica se encuentra con tonos de contrastes muy oscuros. Estas limitaciones son extremas y externas, que afectan el desempeño de los sistemas de reconocimiento debido a que los algoritmos de reconocimiento son incapaces de reconocer sujetos con tanta variabilidad en su rostro térmico.

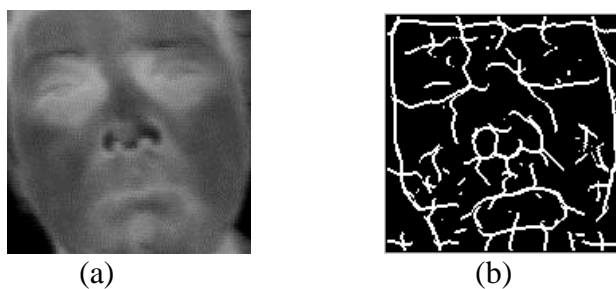


Figura 52. Sujeto con el rostro húmedo. (a) Imagen térmica. (b) Imagen de red vascular

4.4 Metodología propuesta: *VascularThermalFace*

Este apartado presenta una nueva metodología llamada *VascularThermalFace*. Esta nueva metodología permite el uso combinado de características fisiológicas de la red vascular del rostro humano obtenidas del algoritmo *VascularNetwork*, explicado en el Apartado 4.2, y la intensidad de los píxeles del rostro térmico, obtenidos de una cámara térmica. La finalidad de esta metodología consiste en mantener la información proveniente de los píxeles del rostro térmico y

de incorporar la información de la red vascular del rostro humano para poder obtener una nueva representación del rostro térmico. Al utilizar esta metodología *VascularThermalFace* con métodos de reconocimiento de rostros, como los usados en el Capítulo 3, es posible elevar las tasas de desempeño que obtienen los métodos de reconocimiento de rostros operando sólo con la imagen térmica. Además, el uso de esta metodología propuesta permitiría solucionar los problemas que aparecen al utilizar el algoritmo *VascularNetwork*, observadas en el Apartado 4.3, como rotaciones, producto que superpone las características de la red vascular y el rostro térmico, obteniendo una nueva representación que posee más información que el rostro térmico y la red vascular por si solas.

La nueva metodología *VascularThermalFace*, usa el procedimiento de *VascularNetwork* para extraer la ubicación de la red vascular, el cual es sencillo y requiere sólo de 2 pasos importantes: segmentación del rostro térmico y extracción de la red vascular por operadores morfológicos. Luego, a través de una etapa de superposición de los píxeles de la red vascular en el rostro térmico, se genera la nueva imagen térmica (*VascularThermalFace*). La superposición de la red vascular consiste en generar un píxel blanco en la imagen de rostro térmico en la misma posición donde se encontraba la red vascular. El pseudo código del algoritmo *VascularThermalFace* es presentado en la Figura 53.

La principal idea de esta nueva metodología es usar la información de la red vascular en el rostro térmico con el fin de obtener mejoras en métodos de reconocimiento de rostros donde el desempeño de reconocimiento es bajo usando sólo el rostro térmico. Esto se basa en que algunos algoritmos de reconocimiento bajan el rendimiento y tasas de reconocimiento producto de diversos factores que alteran la imagen térmica, cómo cambios corporales y cambios en condiciones de captura de la muestra. Sin embargo, como la red vascular obtenida por el algoritmo *VascularNetwork* es única y repetible [53], en conjunto con el rostro térmico, podría solventar estas limitaciones de captura del espectro térmico y ser una nueva representación del rostro que sea robusta para la aplicación de métodos de reconocimiento de rostros. Este enfoque permitiría ayudar a métodos basados en calces globales como por ejemplo SIFT o SURF, debido a su similitud con las huellas dactilares impresas en el rostro térmico. Un ejemplo de la representación de la metodología *VascularThermalFace* es mostrado en la Figura 54.

Algoritmo: $VascularThermalFace(I_{VASCULAR_NETWORK}, I_{TERMICA})$

Entrada: $I_{VASCULAR_NETWORK}$: Imagen de Red Vascolar

Entrada: $I_{TERMICA}$: Imagen Térmica del Rostro

```
1:  $I_{VASCULAR\_THERMAL\_FACE} \leftarrow I_{TERMICA}$ 
2: for  $i, j \in I_{VASCULAR\_THERMAL\_FACE}$  do
3:   if  $I_{VASCULAR\_NETWORK}(i, j) \neq 0$  then
4:      $I_{VASCULAR\_THERMAL\_FACE}(i, j) \leftarrow 255$ 
5:   end if
6: end for
```

Salida: $I_{VASCULAR_THERMAL_FACE}$: Imagen superposición de rostro térmico y red vascular

Figura 53. Pseudo código del algoritmo $VascularThermalFace$.

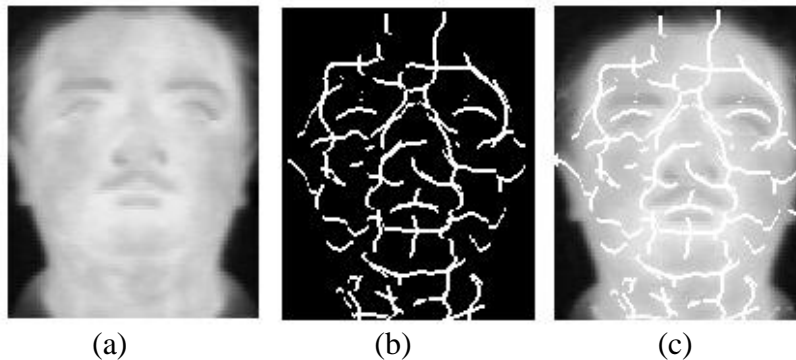


Figura 54. Imágenes de la metodología $VascularThermalFace$. (a) Imagen térmica, (b) Imagen resultante de aplicar el algoritmo $VascularNetwork$, (c) Imagen resultante de aplicar la metodología $VascularThermalFace$.

4.5 Análisis de reconocimiento usando la base UCHThermalFace

Para validar la metodología propuesta $VascularThermalFace$ se utiliza las bases de datos UCHThermalFace. Como se vio en el Capítulo 3, aún falta mejorar en algunos aspectos los resultados de los métodos de reconocimiento de rostros en ambientes controlados y no controlados. Uno de los factores no analizados anteriormente es el caso de la utilización de la red vascular como una característica del rostro térmico, la cual podría ayudar a solucionar problemas de robustez debido a que es única y repetible [8][53].

Se desea analizar la base de datos térmica UCHThermalFace desde 2 enfoques basados en la red vascular: utilizando la metodología $VascularNetwork$ que extrae la red vascular y

utilizando la metodología *VascularThermalFace*, que combina las características de la red vascular y los píxeles del rostro térmico. Se quiere revisar los siguientes experimentos realizados en el capítulo anterior: (i) tamaño de la ventana, (ii) oclusión parcial del rostro, (iii) precisión en la detección de ojos, (iv) galería de interior versus exterior, (v) expresiones faciales; con el fin de analizar si existe una mejora en los resultados de reconocimiento de los métodos seleccionados en el capítulo anterior y con esto validar la metodología propuesta *VascularThermalFace*. La descripción de la base de datos y de los experimentos se encuentra totalmente detallada en el Capítulo 3.

4.5.1 Resultados de reconocimiento

Los resultados obtenidos de la aplicación de los métodos de reconocimiento de rostros son dibujados en gráficos de barras, y muestran el desempeño de todos los métodos de reconocimiento seleccionados en el Capítulo 3 para cada uno de los 3 enfoques utilizados: rostro térmico, metodología *VascularNetwork* y la metodología *VascularThermalFace*. Los resultados mostrados en los gráficos de barra para el caso del rostro térmico fueron los obtenidos en el Capítulo 3, y corresponden al promedio de los resultados obtenidos de los experimentos de reconocimiento de rostros para los conjuntos de Rotación y Palabras mostrados en las Tablas 4-10 y las Figuras 31 y 32. Se visualizan los resultados obtenidos para los casos de las metodologías *VascularNetwork* y *VascularThermalFace* en gráficos de barras, con el fin de realizar una comparación y análisis de los resultados obtenidos para cada método de reconocimiento. Se muestran los resultados para cada categoría de experimentos:

Tamaños de la ventana: La Figura 55 muestra ejemplos de imágenes de distintos tamaños usados para la evaluación de los métodos de reconocimiento de rostros para el caso de la metodología *VascularThermalFace*. Los resultados obtenidos por los métodos de reconocimiento de rostros escogidos en el Capítulo 3 son presentados en la Figura 56 mediante gráficos de barras para el caso de interior y en la Figura 57 para el caso de exterior.

Se puede concluir, de las Figuras 56 y 57, que la utilización de la metodología *VascularNetwork* no mejora el desempeño de los métodos obtenidos en el Capítulo 3, obteniendo resultados de desempeño más bajos que los obtenidos anteriormente (Tabla 4 y Tabla 5). Esto se debe principalmente a los resultados de reconocimiento para el conjunto de rotación de la base UCHThermalFace, explicado en el Apartado 4.3.4, donde se señala que el reconocimiento de rostros utilizando la metodología *VascularNetwork*, no obtiene buenos resultados debido al cambio en la estructura de la red vascular producto de la rotación. En cambio se observa que al aplicar la metodología *VascularThermalFace*, existe una mejora significativa para algunos métodos de reconocimiento, tanto en el caso de interior como en el caso de exterior, mejorando el rendimiento de los métodos de reconocimiento porque la metodología *VascularThermalFace* incorpora la mezcla de información del rostro térmico y su red vascular. La mejora más significativa en el caso de interior es de un 5.39% en la tasa de reconocimiento para el método WLD-HI-80, llegando a subir desde un 91.5% hasta sobre un 96% en la tasa de reconocimiento, seguido de GJD-BC con un 4.68% de mejora (mejora desde un 91.37% a un 95.36%) y SURF-M con un 2.79% (de un 94.4% hasta un 96.6%). En el caso de exterior, sólo se observa una mejora en el caso de GJD-BC de 10.26% en la tasa de reconocimiento, lo que se representa en un aumento desde un 80.45% a un 88.71% en el desempeño. En ambos casos de interior y exterior, las variantes de SIFT no obtienen mejora al utilizar un enfoque basado en el método

VascularNetwork o la metodología *VascularThermalFace*, pero es posible visualizar de las Figuras 56 y 57 que las tasas de rendimiento no varían demasiado al aplicar el esquema de *VascularThermalFace*. Esto, concuerda con el análisis realizado en 4.3.3, donde se señala el problema de SIFT al realizar la etapa de calces, producto de la cantidad de descriptores obtenidos por las imágenes de red vascular.

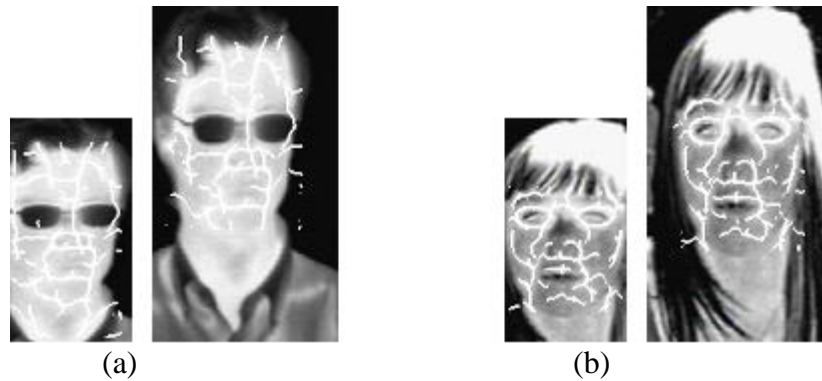


Figura 55. Ejemplos de rostros con diferentes tamaños de cortes para la metodología *VascularThermalFace* (base de datos UCHThermalFace). (a) Sesión interior. Tamaños de ventana (en píxeles): 81x150, 125x225. (b) Sesión exterior. Tamaños de ventana (en píxeles): 81x150, 125x225.

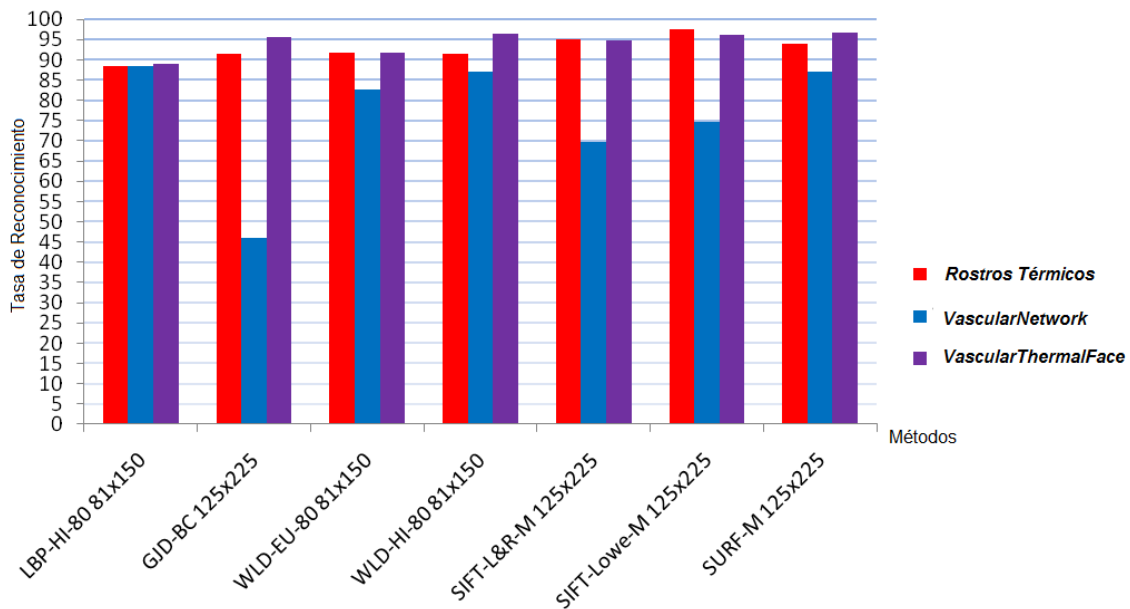


Figura 56. Resultados de los experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos de Rotación y Palabras.

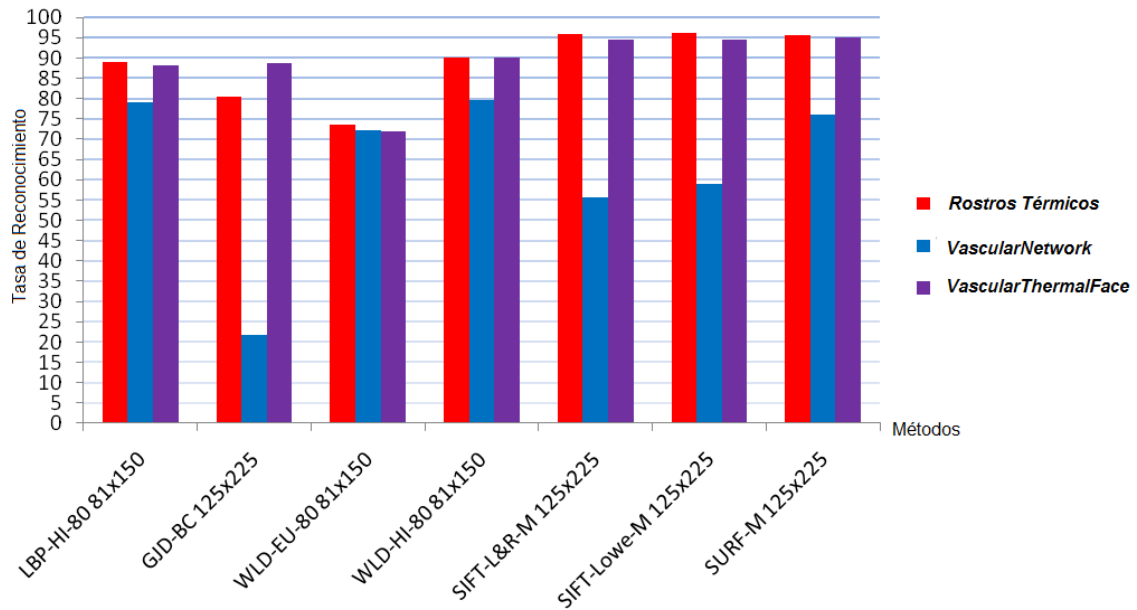


Figura 57. Resultados de los experimentos de distintos tamaños para la ventana de la imagen, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos de Rotación y Palabras.

Oclusión parcial del rostro: La Figura 58 muestra ejemplos de imágenes de 2 sujetos con oclusión parcial del rostro utilizando la metodología *VascularThermalFace* usados para la evaluación de los métodos de reconocimiento de rostros. Las Figuras 59 y 60 muestran los resultados de los métodos de reconocimiento ante una oclusión parcial del rostro para el caso de interior y exterior respectivamente, mediante gráficos de barras.

Se puede concluir, observando las Figuras 59 y 60 que para el caso de interior, todos los métodos, excepto SIFT-Lowe-M, tienen un aumento en las tasas de reconocimiento al aplicar el esquema de *VascularThermalFace*. Esto es debido a que las imágenes del rostro térmico, obtenidas al aplicar la metodología de *VascularThermalFace*, incorporan la red vascular obteniendo más información, lo que hace que los métodos de reconocimiento operen de manera robusta a oclusiones, alcanzando mejores resultados. El método SIFT-Lowe-M tiene un rendimiento similar al mostrado en el Capítulo 3, donde sólo se utilizó la información del rostro térmico. Este aumento, permite obtener mejoras considerables con respecto a la utilización de reconocimiento con el rostro térmico de los individuos, llegando a mejorar el desempeño, para el caso de GJD-BC, en 11.50% la tasa de reconocimiento de rostros. Otras mejoras significativas en las tasas de reconocimiento son para SURF-M (6.05%) y para WLD-HI-80 (5.48%). Para el caso de exterior, el método de GJD-BC logra mejorar en 12.08% su tasa de reconocimiento obtenida en el Capítulo 3 (ver Tabla 7), obteniendo un aumento de un 72.8% hasta un 81.62% en el desempeño del método. El método WLD-HI-80 mejora un 5.48% en su desempeño. Particularmente, para el caso de exterior y con oclusión parcial del rostro, los demás métodos de reconocimiento mantienen un desempeño similar al obtenido en el Capítulo 3. Esto es debido a la saturación en los píxeles que poseen las imágenes de exterior de la base UCHThermalFace, y en conjunto con la oclusión parcial del rostro hacen que los métodos de reconocimiento no mejoren tanto como en el caso de interior, donde no existe la saturación. Para el caso de la metodología

VascularNetwork, ninguno de los métodos seleccionados obtiene mejora alguna, esto es debido a la dificultad de realizar reconocimiento de venas con partes ocluidas y además de la dificultad de los métodos de reconocimiento cuando se usa el conjunto de rotación en el caso de la metodología *VascularNetwork*, en conjunto con la saturación de las imágenes térmicas en la sesión de exterior que hace que disminuya la cantidad de información de la red vascular para algunos sujetos. Un ejemplo de reconocimiento con oclusión parcial del rostro para el método SIFT, es presentado en la Figura 61. En esta Figura 61, se aprecia el reconocimiento de rostros utilizando los 3 enfoques de: rostro térmico, *VascularNetwork* y *VascularThermalFace*. Se aprecia que cuando se usa el enfoque de rostros térmicos, producto de la oclusión y rotación, aparecen algunos calces en zonas alejadas de la oclusión. Para el caso de *VascularNetwork* se puede apreciar en el sujeto superior, esta metodología ayuda a generar muchos descriptores producto de la semejanza entre la estructura de la red vascular, sin embargo para el sujeto inferior, se generan pocos calces, producto que la oclusión no permite calzar la forma de la red vascular. Para el caso *VascularThermalFace*, se aprecia que la información conjunta de rostro térmico y red vascular ayudan al método SIFT a obtener muchos calces correctos para cada uno de los sujetos, inclusive cuando se tienen imágenes rotadas y ocluidas.

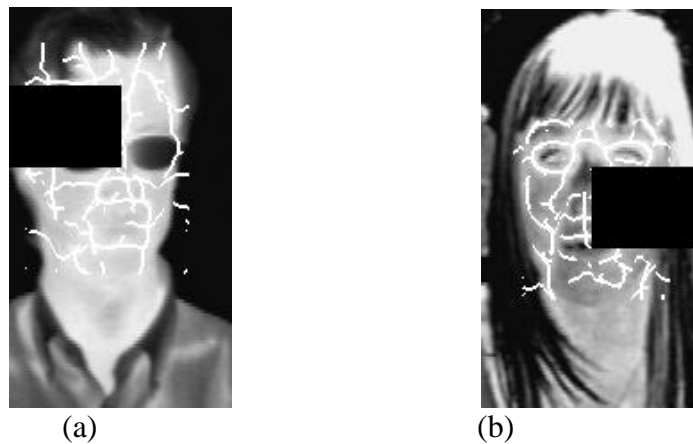


Figura 58. Ejemplos de imágenes de oclusión parcial de rostros para la metodología *VascularThermalFace* (base de datos UCHThermalFace). (a) Sesión interior. Tamaños de ventana (en pixeles): 125x225. (b) Sesión exterior. Tamaños de ventana (en pixeles): 125x225.

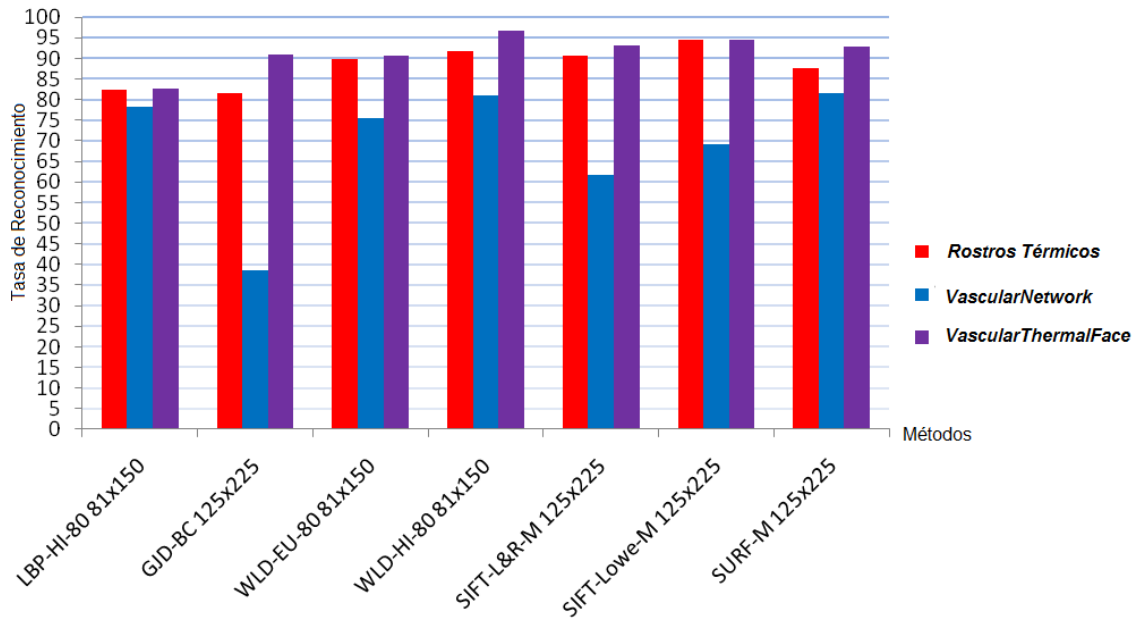


Figura 59. Resultados de experimentos de oclusión parcial, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos de Rotación y Palabras.

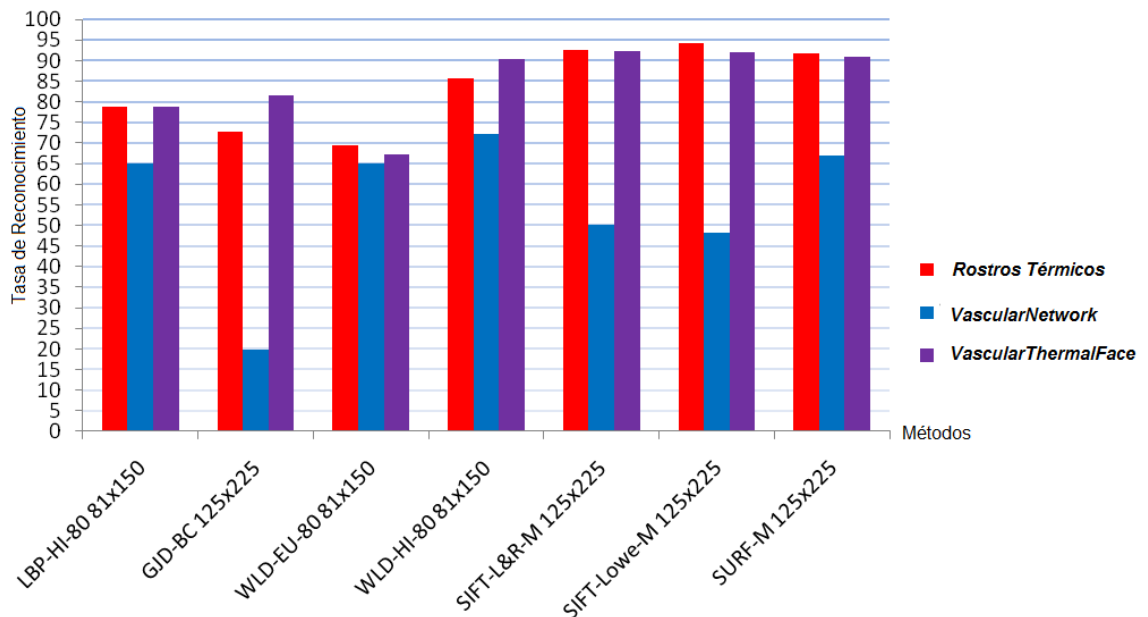


Figura 60. Resultados de experimentos de oclusión parcial, base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos de Rotación y Palabras.

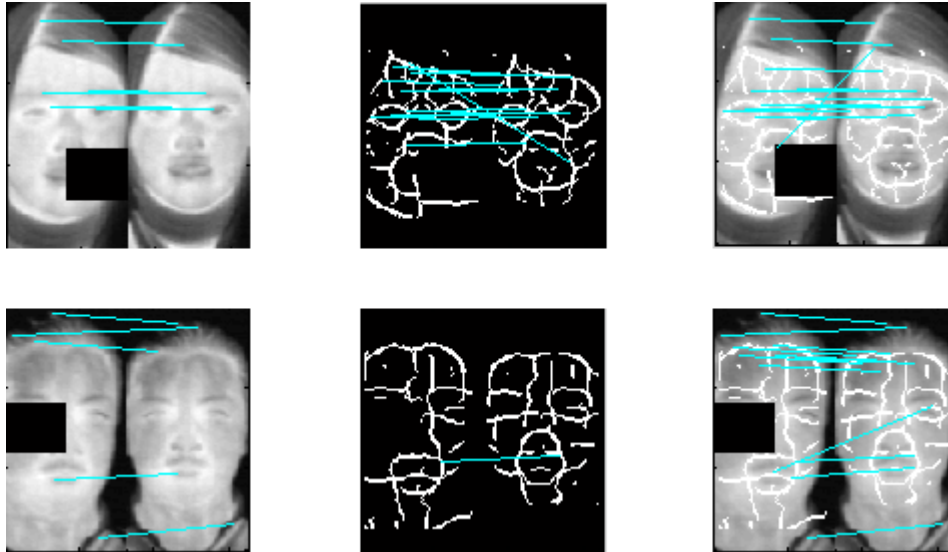


Figura 61. Ejemplos de calces correctos para algunos sujetos de la base de datos UCHThermalFace para la oclusión parcial del rostro térmico. En todos los casos la imagen de la derecha corresponde a la imagen de galería (R6) y en la imagen de la izquierda a la imagen de test. (Ver párrafo para la explicación).

Precisión en la detección de los ojos: Las Figuras 62, 63, 64 muestran la sensibilidad de cada método respecto a la aplicación de ruido en la posición de los ojos para el caso de interior y las Figuras 65, 66, 67 muestran los resultados de reconocimiento ante la aplicación de ruido en la posición de los ojos para el caso de exterior.

Cuando se utiliza la metodología *VascularNetwork*, solamente 2 métodos obtienen mejor desempeño que los obtenidos en el Capítulo 3 (Figura 31 y 32), para los casos de interior y exterior. Para el caso de 5% de ruido en la posición de los ojos, caso de exterior, el método WLD-HI-80 mejora el desempeño en 1.83%. Para el caso 10% de ruido en la posición de los ojos, sesión de interior, el método LBP-HI-80 mejora en 0.68%, siendo únicamente estos dos casos donde se obtiene una mejora para los métodos de reconocimiento cuando se usa ruido en la posición de los ojos. Los demás métodos de reconocimiento no mejoran al utilizar la metodología *VascularNetwork* debido a los factores mencionados anteriormente, como las rotaciones en las imágenes térmicas. El comportamiento de los métodos de reconocimiento ante ruido en la posición de los ojos para el caso de la red vascular es de la misma manera que en el caso de rostro térmico, obtenido en el Capítulo 3, al aumentar el ruido en la posición de los ojos, el desempeño de los métodos de reconocimiento tiende a decaer. Ejemplos de calces realizados por el método SIFT son mostrados en la Figura 68. Se observa como el rendimiento de SIFT decae debido a que la red vascular no es calzada adecuadamente cuando existe una variación considerable en su estructura producto del ruido en la posición de los ojos y debido a la cantidad de calces falsos obtenidos por SIFT.

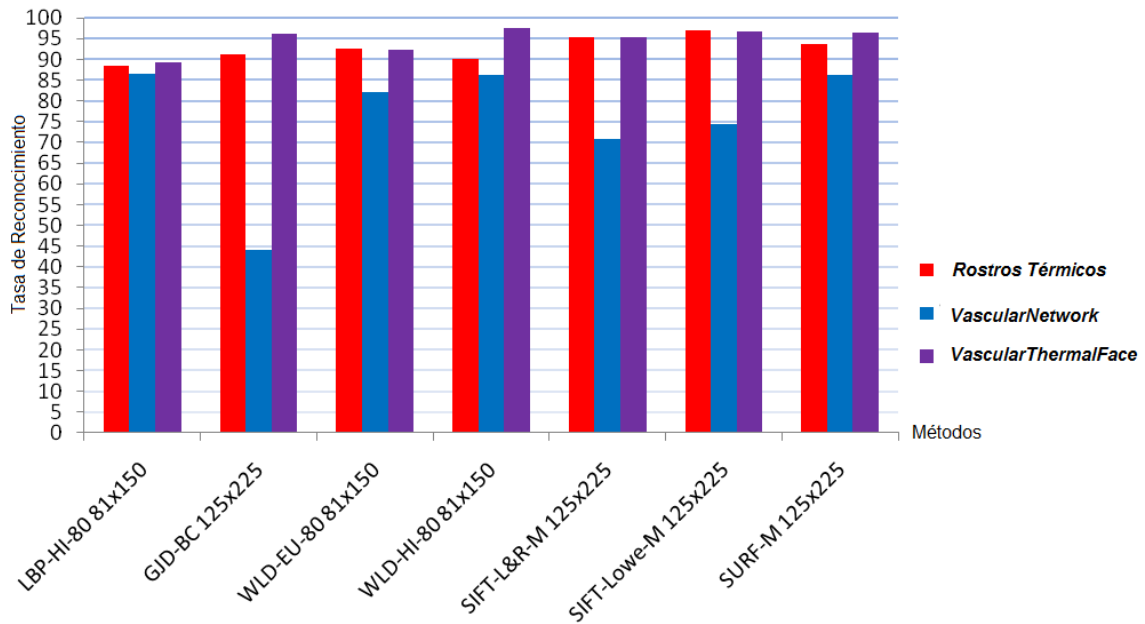


Figura 62. Resultados de experimentos de ruido en la posición de los ojos (2.5%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos Rotación y Palabras.

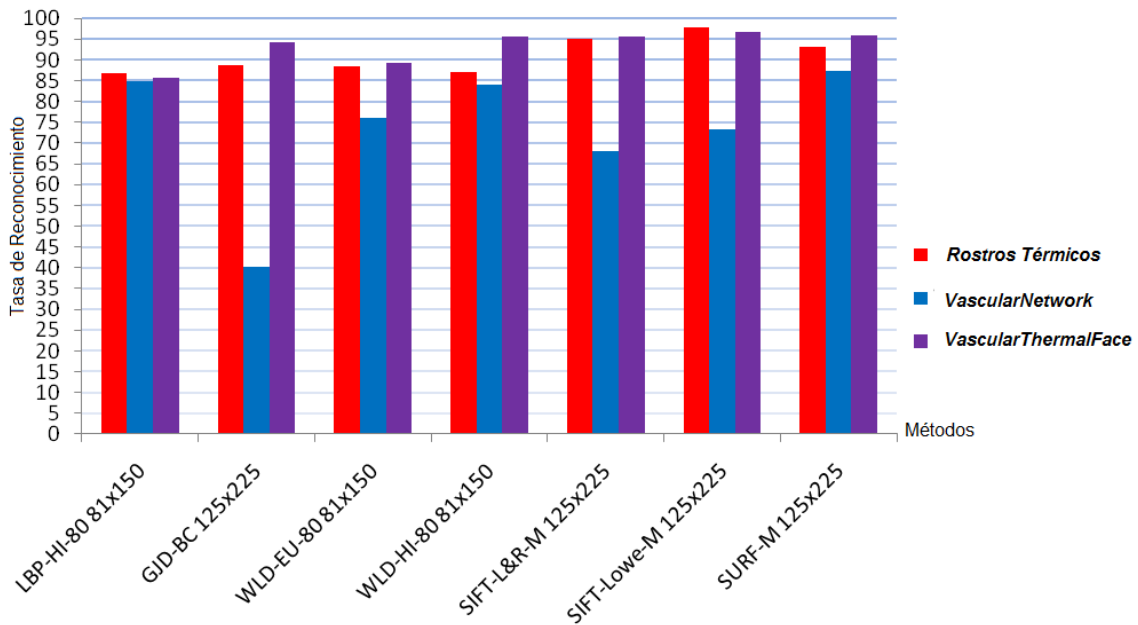


Figura 63. Resultados de experimentos de ruido en la posición de los ojos (5%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos Rotación y Palabras.

Para el caso de *VascularThermalFace*, cuando se usa 2.5% de ruido en la posición de los ojos, es posible observar de las Figuras 62 y 65, que existe un mejora para 4 métodos de

reconocimiento LBP-HI-80, GJD-BC, WLD-HI-80 y SURF-M, donde el aumento en la tasa de reconocimiento más importante es de 8.44% para el método WLD-HI-80, aumentando desde un 90% a un 97.6% en la tasa de reconocimiento. Para el caso de exterior, 2 métodos mejoran su desempeño GJD-BC con un 11.01% y WLD-HI-80 con un 6.74%. Cuando se utiliza 5% y 10% de ruido en la posición de los ojos, se observa de las Figuras 63, 64, 66, 67, que la mayoría de los métodos tienden a mejorar el desempeño con la metodología *VascularThermalFace*, debido a la inserción de mayor información producto que la red vascular esta sobrepuesta a la imagen que soluciona los problemas del rostro térmico y red vascular. Los resultados de las mejoras en el desempeño de los métodos de reconocimiento para este caso pueden ser vistos en la Tabla 15. Cabe destacar que para el caso de 10% de ruido, al utilizar este enfoque basado en *VascularThermalFace*, el método WLD-HI-80, en el caso de interior, mejora aproximadamente un 15% del valor que poseía solo al utilizar el rostro (de un 78% hasta ~90% en el desempeño de reconocimiento).

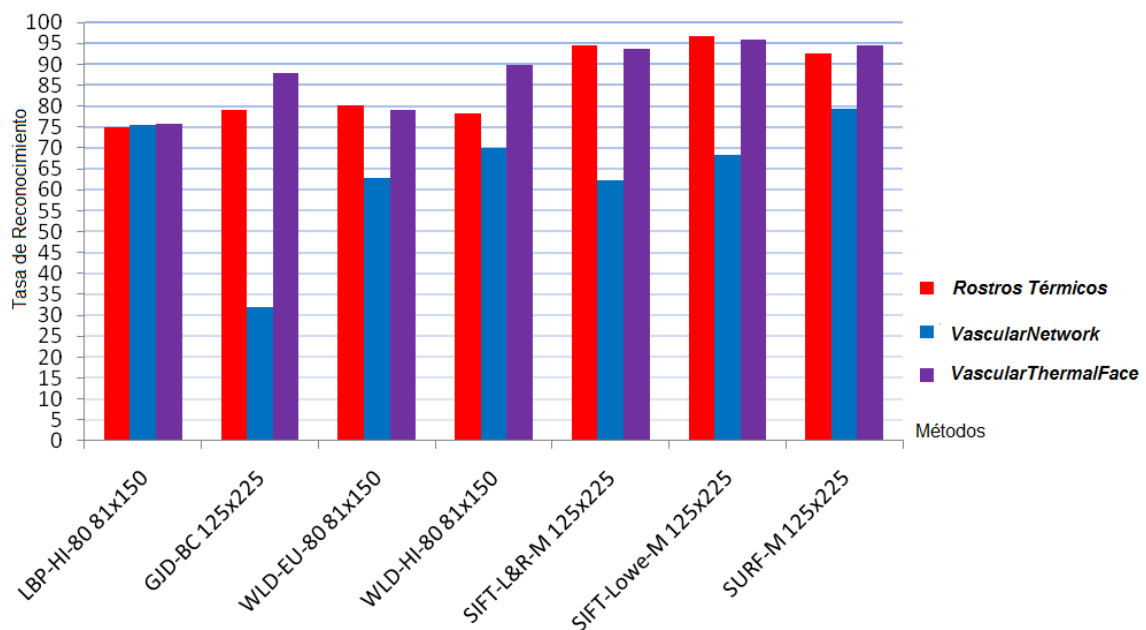


Figura 64. Resultados de experimentos de ruido en la posición de los ojos (10%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión interior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos Rotación y Palabras.

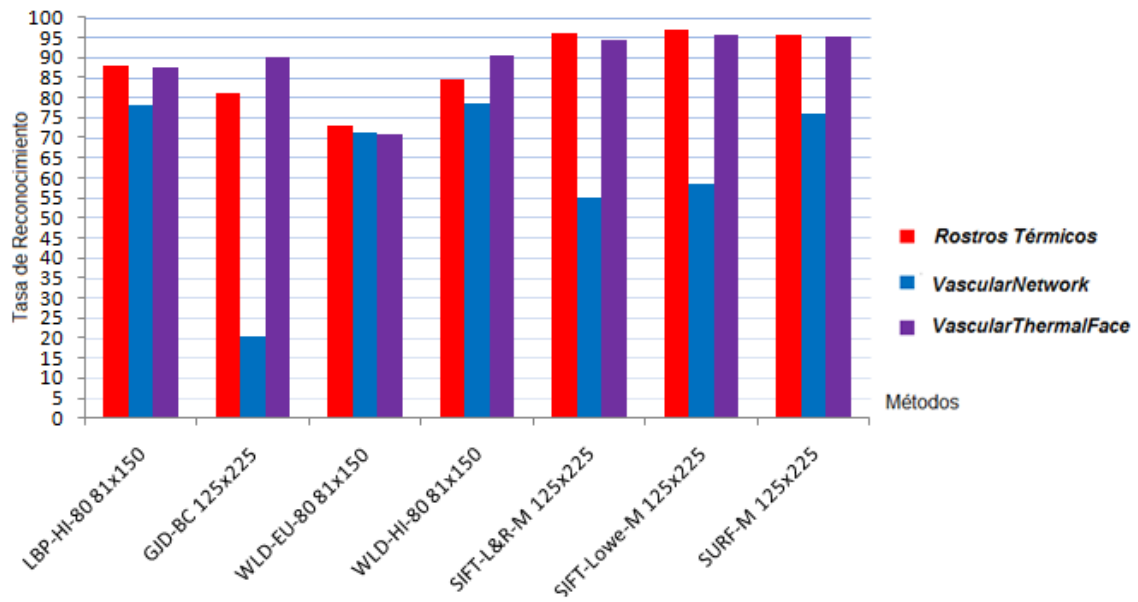


Figura 65. Resultados de experimentos de ruido en la posición de los ojos (2.5%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos Rotación y Palabras.

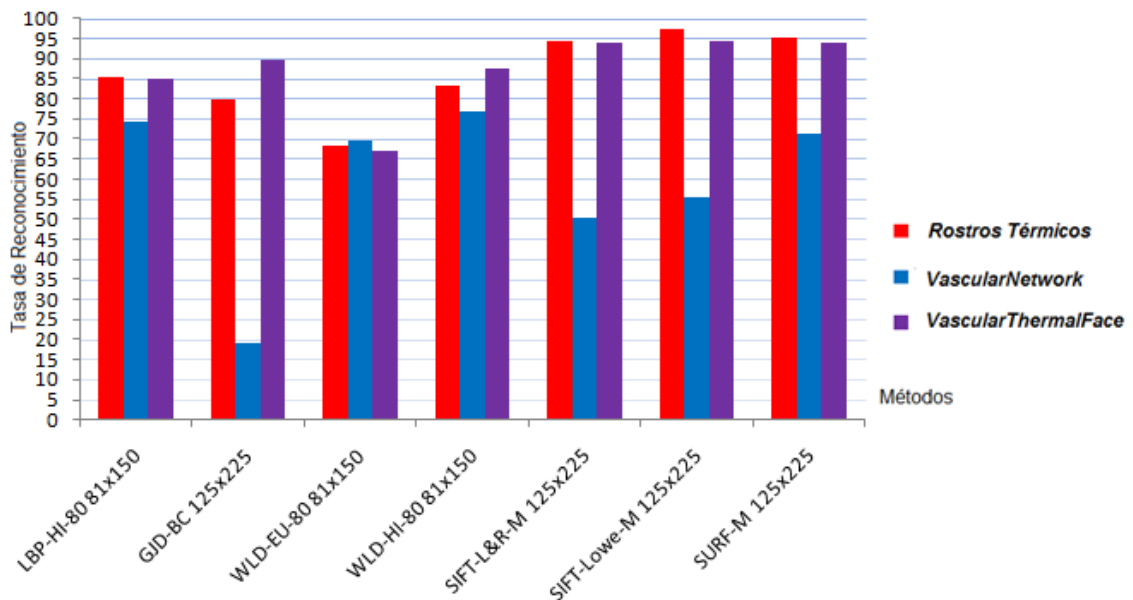


Figura 66. Resultados de experimentos de ruido en la posición de los ojos (5%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos Rotación y Palabras.

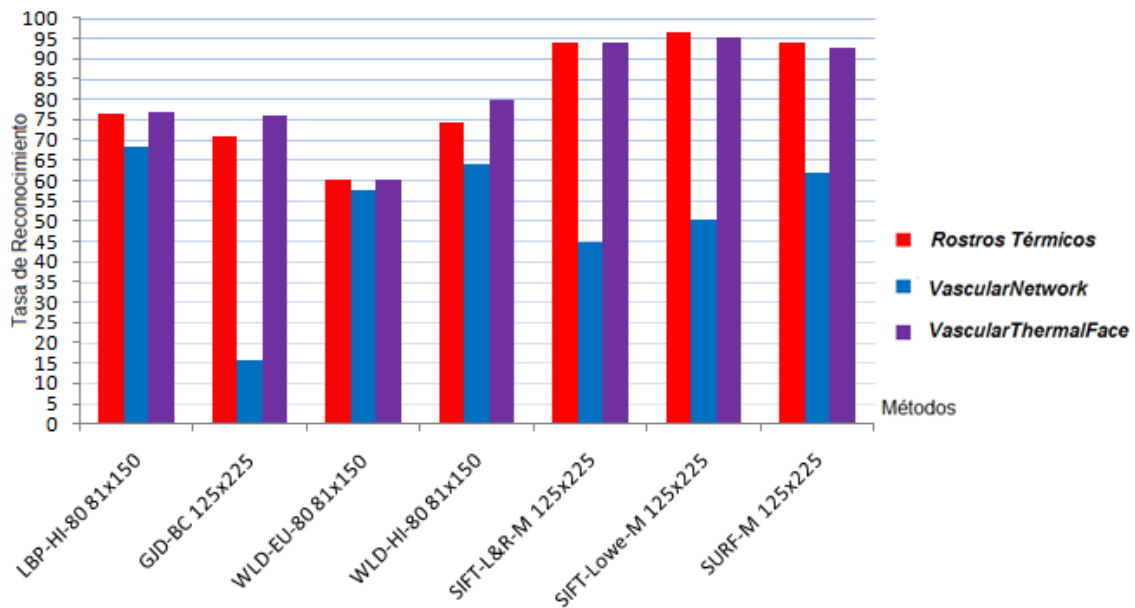


Figura 67. Resultados de experimentos de ruido en la posición de los ojos (10%), base de datos UCHThermalFace. Mejores variantes de los métodos. Sesión exterior. Tasa de reconocimiento para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*. Conjuntos Rotación y Palabras.

Galería de interior vs exterior: En la Figura 69 se muestran los resultados de los métodos de reconocimiento de rostros obtenidos al aplicar una galería de imágenes capturadas en una sesión de interior con un test de imágenes de exterior, y en la Figura 70 se muestra el resultado los métodos de reconocimiento de rostros del caso contrario, para una galería de imágenes de exterior versus un test de imágenes de interior para los 3 enfoques: rostros térmico, *VascularNetwork* y *VascularThermalFace*.

Tanto para el caso de la *VascularNetwork* y *VascularThermalFace*, existe un aumento considerable en las tasas de desempeño de los métodos de reconocimiento de rostros obtenidas en el Capítulo 3. Es posible observar de la Figura 69, que para el caso de *VascularNetwork* que cuando se usan galería de interior o de exterior, los métodos LBP-HI-80 y las variantes de WLD-X-80 obtienen mejoras significativas de hasta un 43% en el caso de galería interior para LBP-HI-80 (de un 21.8% a un 31.2%) y un 35% en el caso de galería de exterior para el mismo método (de un 18.1% a un 24.8%). Si bien este aumento parece ser muy bueno, los valores obtenidos para este experimento son bajos para realizar un reconocimiento de rostros adecuado y robusto en el caso de galerías de interior versus exterior y viceversa. En la Figura 70 se pueden observar las mejoras obtenidas al utilizar el enfoque de *VascularThermalFace*. Se observa que varios métodos aumentan su desempeño de reconocimiento en comparación con los resultados del Capítulo 3. Resultados relevantes para el caso de galerías de interior son LBP-HI-80 con una mejora de 15.6% y para el caso de galerías de exterior para el método SIFT-Lowe-M con un 12.79% de mejora. Este aumento en las tasas de reconocimiento versus los resultados de reconocimiento obtenidos en el Capítulo 3 parecen ser muy buenos pero cómo resultados de reconocimiento son muy bajos, por lo que es difícil realizar un sistema de reconocimiento robusto para el caso de galerías de interior versus exterior y viceversa. La explicación de los bajos resultados de los métodos de reconocimiento cuando se usan galerías de interior versus imágenes de test de

exterior y viceversa es debido a los factores nombrados en el Capítulo 3, como la saturación de la cámara térmica producto del sol, en el caso de exterior, lo cual hace que los pixeles de las imágenes térmicas se encuentran en zonas saturadas, o muy blancos o muy negros, también a la configuración de la cámara térmica usada y al propio calentamiento de los detectores térmicos. Como se explico en la Sección 4.3.4, una mala configuración de la cámara, modifica mucho la visualización de la imagen térmica y por ende la aplicación de algoritmos de reconocimiento ante condiciones adversas trae como resultados el bajo desempeño de los métodos de reconocimiento. Es por este motivo que en el siguiente capítulo se analizará nuevamente este tema desde una perspectiva temporal y con más sesiones de interior versus exterior.

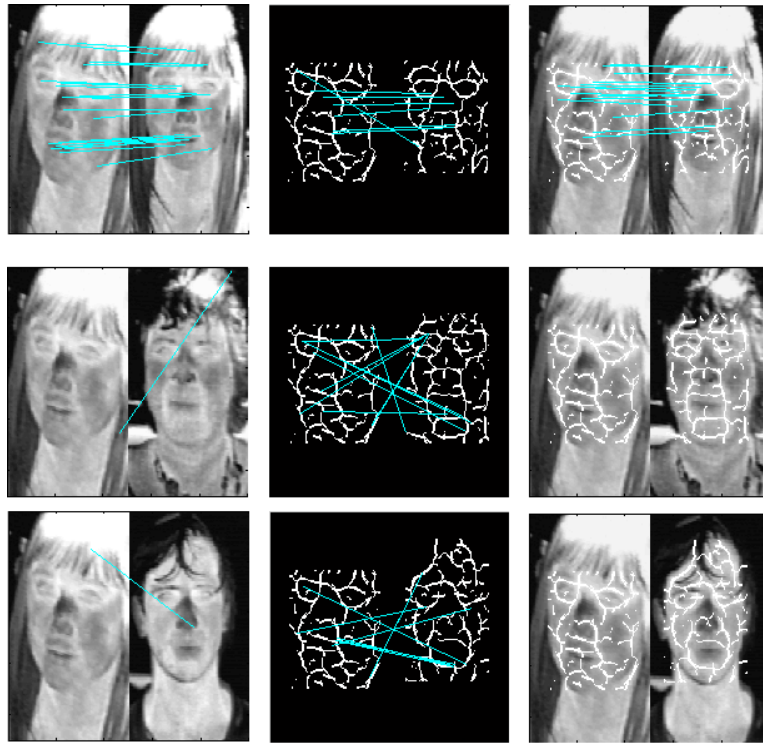


Figura 68. Ejemplos de calces SIFT. Ejemplos de calces realizados por el método SIFT para un sujeto de prueba de la base de datos UCHThermalFace para el caso de 10% de ruido en la posición de ojos. La imagen de la derecha corresponde a la imagen de galería (R6) y en la imagen de la izquierda a la imagen de test. Se observan muchos calces falsos en el caso de red vascular. Para *VascularThermalFace* se observa un buen desempeño de SIFT.

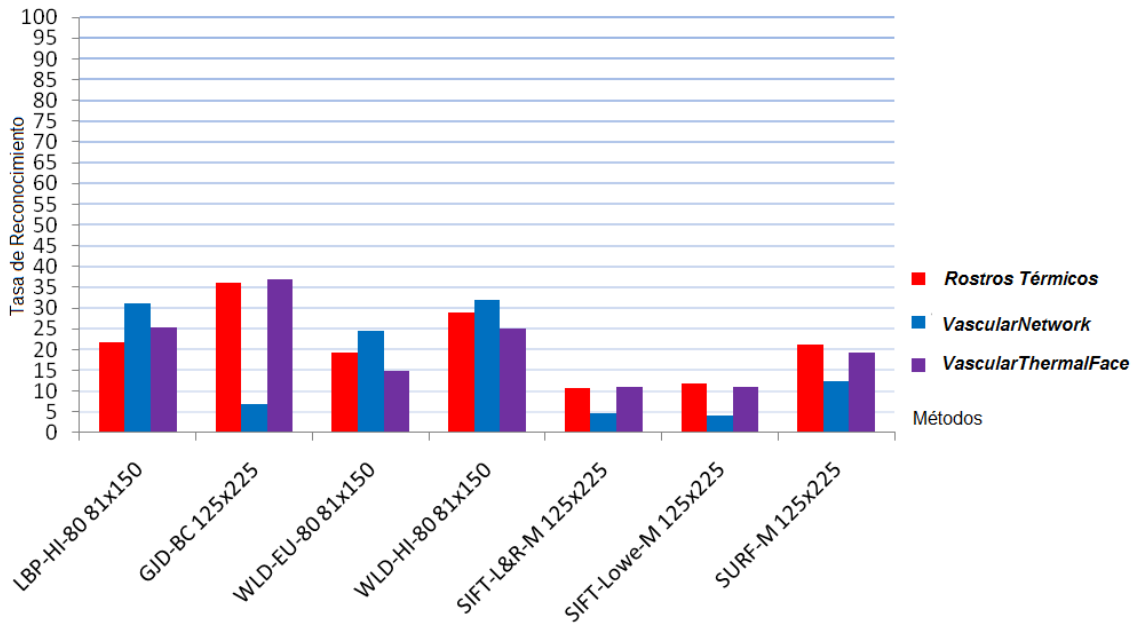


Figura 69. Resultados de experimentos de galería de interior versus exterior, base de datos UCHThermalFace. Mejores variantes de los métodos. Imágenes de interior para el conjunto de galería. Tasa de reconocimiento. Conjuntos Rotación y Palabras.

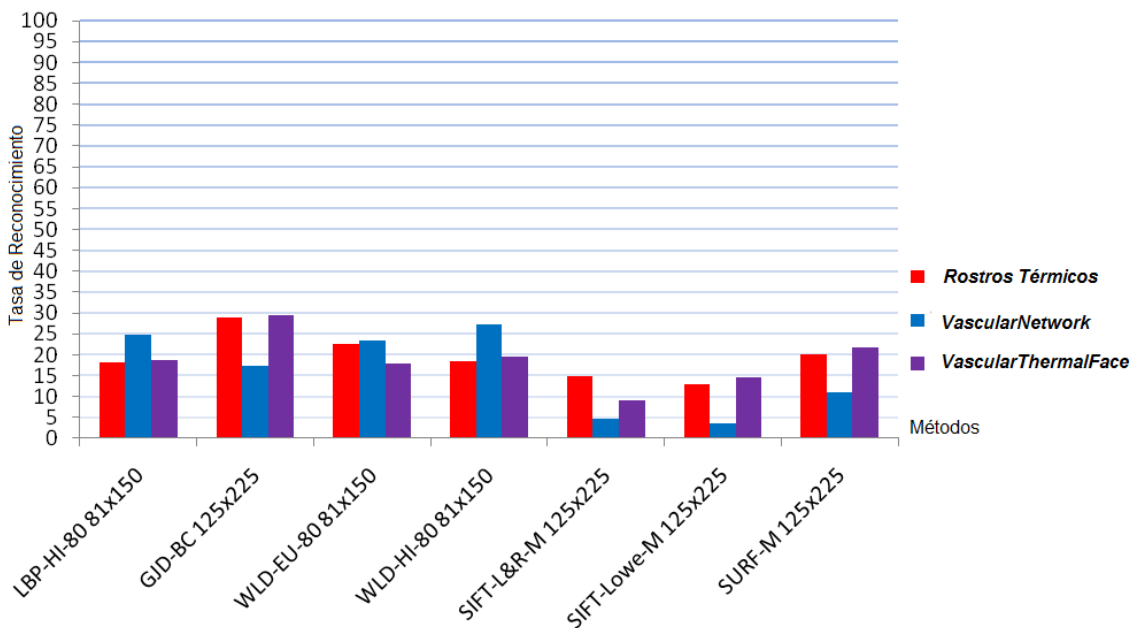


Figura 70. Resultados de experimentos de galería de exterior versus interior, base de datos UCHThermalFace. Mejores variantes de los métodos. Imágenes de exterior para el conjunto de galería. Tasa de reconocimiento. Conjuntos Rotación y Palabras.

Expresiones faciales: En la Tabla 13 son reportados los resultados del experimento de reconocimiento de rostros que usan el conjunto de Expresión. Recordar que las imágenes de rostros del conjunto de expresión fueron obtenidas con un detector de rostro térmico, por lo tanto

no se encuentran alineadas. La Tabla 13 muestra el promedio de la tasa de reconocimiento para 6 experimentos para las variantes de rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Se puede observar que el mejor resultado es obtenido por las variantes SIFT-X-X, con una alta tasa de reconocimiento (~99.8%) para el caso de *VascularThermalFace*, seguido del método SURF que aumenta su desempeño en el resultado de reconocimiento de rostros obtenido en el Capítulo 3, para ambas variantes *VascularNetwork* y *VascularThermalFace*. En general se observa de la Tabla 13, que todas las variantes obtienen mejoras en sus resultados al aplicar la metodología *VascularThermalFace*, donde las imágenes de la base de datos no se encuentran alineadas, poseen diversas expresiones y son frontales. Al utilizar la metodología de *VascularNetwork* todos los métodos bajan su desempeño de reconocimiento de rostros, excepto SURF. Esto se debe a que las imágenes no están alineadas y la red vascular aparece en diversas posiciones, lo cual hace que los métodos de reconocimiento basados en calces locales no operen de manera adecuada, las variantes de SIFT poseen el problema de muchos calces falsos para la red vascular y el método SURF opera bien para este caso.

Para resumir los resultados obtenidos por las metodologías de *VascularNetwork* y *VascularThermalFace* se muestran 2 Tablas. Las Tablas 14 y 15 contienen las mejoras producidas en los resultados de reconocimiento de rostros obtenidos por cada uno de los métodos de reconocimiento usados en esta tesis. Las mejoras de los métodos de reconocimiento se calculan sobre el desempeño obtenido en el Capítulo 3 y que corresponde al resultado de reconocimiento cuando se usa el rostro térmico. La Tabla 14 muestra las mejoras producidas al aplicar el enfoque *VascularNetwork*, mientras que la Tabla 15 muestra las mejoras obtenidas al aplicar la metodología de *VascularThermalFace*.

Tabla 13. Resultados de experimentos de expresiones faciales, base de datos UCHThermalFace. Mejores variantes de los métodos. Tasa de reconocimiento. Conjuntos de Expresiones (E1-E3, V1-V3).

Métodos	Rostro Térmico Promedio [%]	<i>VascularNetwork</i> Promedio[%]	<i>VascularThermalFace</i> Promedio[%]
LBP-HI-80	94.9	85.2	95.0
GJD-BC	94.4	93.1	95.4
WLD-EU-80	93.9	78.5	94.1
WLD-HI-80	94.1	82.6	94.4
SIFT-L&R-M	98.9	89.0	99.8
SIFT-Lowe-M	99.6	91.0	99.8
SURF-M	87.2	94.5	97.6

Tabla 14. Mejoras obtenidas en los resultados de reconocimiento al aplicar la metodología *VascularNetwork*. La Tabla muestra las mejoras obtenidas y la variación porcentual alcanzada en base al resultado de reconocimiento de rostros obtenido en el Capítulo 3 para cada experimento. La mejora se mide usando $\left(\frac{RC4}{RC3} - 1\right) * 100$, y la variación porcentual se calcula usando $(RC4 - RC3)$, con RC4: porcentaje de reconocimiento obtenido en el Capítulo 4 y RC3: porcentaje de reconocimiento obtenido en el Capítulo 3.

Métodos	Tamaño Ventana Interior	Tamaño Ventana Exterior	Oclusión Parcial Interior	Oclusión Parcial Exterior	2.5% Ruido Interior	2.5% Ruido Exterior	5% Ruido Interior	5% Ruido Exterior	10% Ruido Interior	10% Ruido Exterior	Galería Interior-Exterior	Galería Exterior-Interior	Expresiones
LBP-HI-80	-	-	-	-	-	-	-	-	0.68% 0.51%	-	43.00% 9.38%	35.85% 6.68%	-
GJD-BC	-	-	-	-	-	-	-	-	-	-	-	-	-
WLD-EU-80	-	-	-	-	-	-	-	1.83% 1.25%	-	-	26.90% 5.20%	3.69% 0.83%	-
WLD-HI-80	-	-	-	-	-	-	-	-	-	-	10.65% 3.09%	47.48% 8.77%	-
SIFT-L&R-M	-	-	-	-	-	-	-	-	-	-	-	-	-
SIFT-Lowe-M	-	-	-	-	-	-	-	-	-	-	-	-	-
SURF-M	-	-	-	-	-	-	-	-	-	-	-	-	8.37% 7.30%

Tabla 15. Mejoras obtenidas en los resultados de reconocimiento al aplicar la metodología *VascularThermalFace*. La Tabla muestra las mejoras obtenidas y la variación porcentual alcanzada en base al resultado de reconocimiento de rostros obtenido en el Capítulo 3 para cada experimento. La mejora se mide usando $\left(\frac{RC4}{RC3} - 1\right) * 100$, y la variación porcentual se calcula usando $(RC4 - RC3)$, con RC4: porcentaje de reconocimiento obtenido en el Capítulo 4 y RC3: porcentaje de reconocimiento obtenido en el Capítulo 3.

Métodos	Tamaño Ventana Interior	Tamaño Ventana Exterior	Oclusión Parcial Interior	Oclusión Parcial Exterior	2.5% Ruido Interior	2.5% Ruido Exterior	5% Ruido Interior	5% Ruido Exterior	10% Ruido Interior	10% Ruido Exterior	Galería Interior-Exterior	Galería Exterior-Interior	Expresiones
LBP-HI-80	0.67% 0.60%	-	0.56% 0.46%	0.08% 0.06%	0.71% 0.63%	-	-	-	1.26% 0.94%	0.38% 0.29%	15.58% 3.40%	3.92% 0.71%	0.11% 0.10%
GJD-BC	4.68% 4.28%	10.26% 8.26%	11.50% 9.38%	12.08% 8.80%	5.55% 5.06%	11.01% 8.92%	5.98% 5.31%	12.26% 9.78%	11.31% 8.94%	7.06% 5.00%	2.83% 1.20%	2.46% 0.71%	1.01% 1.00%
WLD-EU-80	-	-	0.79% 0.71%	-	-	-	0.80% 0.71%	-	-	0.33% 0.20%	-	-	0.21% 0.20%
WLD-HI-80	5.39% 4.94%	-	5.48% 5.03%	5.48% 4.69%	8.44% 7.60%	6.74% 5.71%	9.61% 8.37%	5.14% 4.28%	14.86% 11.64%	7.76% 5.75%	-	5.74% 1.06%	0.32% 0.30%
SIFT-L&R-M	-	-	2.94% 2.66%	-	-	-	0.45% 0.43%	-	-	0.02% 0.02%	2.89% 0.31%	-	0.91% 0.90%
SIFT-Lowe-M	-	-	-	-	-	-	-	-	-	-	-	12.79% 1.56%	0.20% 0.20%
SURF-M	2.79% 2.63%	-	6.05% 5.29%	-	2.95% 2.77%	-	3.14% 2.98%	-	1.94% 1.80%	-	-	8.20% 1.65%	11.93% 10.40%

Es posible observar de la Tabla 14 que los métodos de reconocimiento evaluados utilizando el enfoque de *VascularNetwork* no obtienen mejoras significativas en el desempeño de reconocimiento de rostros comparado con el resultado obtenido en el Capítulo 3, donde sólo se utilizaba el rostro térmico para la base de datos UCHThermalFace. Esto se debe a varios factores,

como la rotación de las imágenes de red vascular, la configuración de la cámara térmica, la saturación de las imágenes térmicas y al propio desempeño de los métodos, debido a que no están diseñados para trabajar con la red vascular. Sin embargo, al observar la Tabla 15, es posible ver que al utilizar la metodología de *VascularThermalFace*, los métodos de reconocimiento de rostros aumentan su desempeño (en la mayoría de los casos) debido a la incorporación de la red vascular en conjunto con el rostro térmico, lo cual aumenta la información y permite a los métodos de reconocimiento aumentar su desempeño. Se observa que los casos en que mejoran más los resultados de reconocimiento obtenidos del Capítulo 3 son GJD-BC y WLD-HI-80. Esta mejora en los resultados obtenidos para el caso de GJD-BC es producto que los jets ahora tienen información complementaria en algunos puntos de la grilla de la imagen producto de la información vascular y de la orientación obtenida de las líneas de la red vascular, lo cual hace que al realizar la cuenta de Borda sea más fácil de clasificar las imágenes. Para el caso de WLD-HI-80, la información de excitación y de orientación se ve beneficiada con la red vascular en la imagen térmica, la cual hace que al comparar los histogramas de las imágenes sea más fácil de encontrar el más parecido.

4.6 Discusión y conclusiones del capítulo

En este capítulo se presentó una nueva metodología llamada *VascularThermalFace* que combina las características vasculares del rostro humano y la información del rostro térmico. Esta nueva metodología utiliza un algoritmo de extracción de la red vascular del rostro térmico llamado *VascularNetwork*, con la cual extrae la información de la red vascular para combinarla con la información del rostro térmico.

Se presentó el algoritmo *VascularNetwork*, se estudiaron sus características, donde se demostró que éste permite extraer la red vascular superficial del rostro humano y que ésta es una característica del rostro que puede ser usada a lo largo del tiempo. También se estudiaron las limitaciones que posee el uso de esta metodología y la obtención de la red vascular desde varios puntos de vista, como por ejemplo: las rotaciones, el uso de lentes en los individuos, la saturación de la cámara térmica, entre otros factores.

Esta nueva metodología *VascularThermalFace* fue evaluada utilizando la base de datos UCHThermalFace, con la cual se realizaron experimentos basado en el análisis de distintos métodos de reconocimiento de rostros en el espectro térmico. La base de datos UCHThermalFace incluye aspectos importantes para analizar la nueva metodología tales como rotaciones con diferentes ángulos de orientación e inclinación, variaciones en la condición del ambiente (interior/exterior), expresiones faciales y oclusión. Los métodos usados para validar esta metodología fueron: histogramas de LBP, descriptores jets de Gabor, descriptores SIFT, histogramas de WLD y SURF. Se utilizaron 3 enfoques diferentes para evaluar los métodos de reconocimiento de rostros: rostro térmico, *VascularNetwork* y *VascularThermalFace*. Comentarios de los principales resultados de este estudio y algunas conclusiones son mostradas a continuación:

Comentarios respecto al tamaño de la ventana: Al utilizar la metodología de *VascularNetwork* no existe mejora en los resultados obtenidos en el Capítulo 3, donde se usaban imágenes del rostro térmico para los métodos de reconocimiento. Al usar la metodología propuesta *VascularThermalFace*, el método GJD-BC obtiene una mejora considerable en los casos de

interior y exterior. Además la metodología *VascularThermalFace* mejora los resultados de reconocimiento de los métodos WLD-HI-80, SURF-M y LBP-HI-80 en el caso de interior. Las variantes de SIFT no se ven afectadas al utilizar ninguno de estos dos algoritmos debido a la etapa de generación de calces, obteniendo muchos calces falsos, lo cual trae consigo una disminución en las tasas de reconocimiento. Mejores resultados son obtenidos con la aplicación directa del rostro térmico (Capítulo 3).

Comentarios acerca del alineamiento, oclusiones y expresiones: Al analizar estos experimentos, se puede concluir que la mayoría de los métodos mejora en gran medida al utilizar la metodología propuesta *VascularThermalFace*. Esto hace que el nuevo algoritmo mejore la robustez de cada método al utilizar la combinación de la red vascular y el rostro térmico. En la mayoría de los casos los métodos que mejoran más son GJD-BC y WLD-HI-80. En el caso del método *VascularNetwork* no existe una mejora significativa en cada uno de los métodos al utilizar esta técnica. Esto es debido a las rotaciones que posee la base UCHThermalFace, lo cual hace que la red vascular cambie la forma de la estructura, más aún cuando ésta no se encuentra alineada.

Comentarios con respecto a las condiciones interior/exterior: Al analizar el comportamiento de los métodos de reconocimiento de rostros cuando las imágenes de test son adquiridas en una configuración de exterior y las imágenes de galería son adquiridas en interior, o viceversa, el desempeño de todos los métodos es muy bajo, lo cual fue observado en el Capítulo 3. Sin embargo la utilización de los métodos de *VascularNetwork* y *VascularThermalFace* mejora en varios puntos porcentuales los resultados obtenidos anteriormente. Aunque existen mejoras significativas en los resultados, todavía el desempeño de los métodos es bajo para realizar un reconocimiento robusto de rostros en condiciones de interior y exterior. Este aspecto es analizado en el Capítulo 5 por medio de un estudio de galerías de interior y exterior capturadas en varias sesiones en el tiempo y modificando el problema de saturación de las imágenes de exterior mostrado en el Capítulo 3.

Conclusiones acerca de los enfoques propuestos. La aplicación del método *VascularNetwork* permite mejorar sólo resultados de galerías de interior versus exterior, donde al utilizar la red vascular, esta característica se mantiene en las condiciones de interior y exterior. En el caso del enfoque de *VascularThermalFace*, al utilizar la base UCHThermalFace, es posible observar que existe una mejora considerable para algunos métodos de reconocimiento de rostros, en especial para los métodos GJD-BC y WLD-HI-80 que son los métodos que mejoran más cantidad de veces, con un promedio de mejoras para el desempeño de reconocimiento de ~8% para el caso de GJD-BC y ~7.5% para el caso de WLD-HI-80 en todos los experimentos mostrados anteriormente. Las variantes que menos mejoran los resultados de reconocimiento son las variantes de SIFT, debido al problema en la etapa de calces, generando gran cantidad de calces falsos, lo cual hace que el desempeño de reconocimiento disminuya. Sin embargo, los resultados obtenidos con la metodología *VascularThermalFace* son comparables a los resultados obtenidos en el Capítulo 3, en cambio los resultados obtenidos al aplicar *VascularNetwork* son bastante bajos por el problema de calces. El método SURF obtiene algunas mejoras en los resultados producto de poseer descriptores más selectivos que SIFT, lo cual hace que los calces puedan ser realizados más efectivamente.

5 Estudio Comparativo de Reconocimiento de Rostros en el Tiempo

5.1 Introducción

Muchos estudios comparativos de métodos de reconocimiento de rostros en el dominio térmico basan sus análisis en la obtención de resultados que provienen de imágenes térmicas capturadas durante una sesión de interior o exterior [5][46][67][68], de la misma manera como fue mostrado en los Capítulos 3 y 4. Sin embargo, otras investigaciones [12][14][42][74] muestran que el uso de imágenes térmicas capturadas en múltiples sesiones dan como resultados bajas tasas de reconocimiento de rostros producto de la variación temporal que sufren éstas cuando son capturadas con un lapso de tiempo muy grande. En otras palabras, si se adquiere una imagen térmica de un sujeto en un día determinado y otra imagen del mismo sujeto en otro día, al tratar de reconocer la identidad del sujeto, el desempeño de los métodos de reconocimiento se verá afectado por las variaciones ambientales y cambios en el metabolismo del sujeto que se produjeron entre la adquisición de ambas imágenes.

Las publicaciones [12][14][42][74] señalan que el resultado del reconocimiento de rostros térmico disminuye debido al lapso de tiempo que transcurre en la adquisición de las imágenes térmicas utilizadas. Esto quiere decir, que las imágenes capturadas entre periodos extensos de tiempo sufren alteraciones debido a factores como: (i) el efecto del metabolismo, debido al cambio y redistribución del flujo de sangre y las venas, que hace que se realice un cambio abrupto en la temperatura local de la piel, (ii) efecto del cambio en las condiciones ambientales, y (iii) cambios de temperatura producto del exceso de calor o frío, o situaciones cotidianas como mucho ejercicio físico.

Estudios de reconocimiento de rostros térmicos [14][42] basados en métodos de apariencia como PCA muestran bajos resultados de reconocimiento cuando existe este lapso temporal en la captura de las imágenes térmicas. En [14][42] se realiza un estudio de reconocimiento de rostros usando imágenes del espectro visible e imágenes térmicas, obtenidas en múltiples sesiones de tiempo. Los resultados de este estudio señalan que el desempeño del método PCA es bajo cuando utiliza imágenes térmicas adquiridas en varias sesiones de tiempo. Además señalan que utilizando un esquema de fusión entre las imágenes visibles y las térmicas es posible mejorar considerablemente el desempeño de reconocimiento y con ello el problema del lapso temporal. Entre sus conclusiones, ambos estudios concuerdan que cuando existe un lapso temporal en la adquisición de las imágenes térmicas, el desempeño es bajo y que los sistemas basados en la fusión de imágenes visibles e infrarrojos deberían ser la solución a seguir. Sin embargo, estos estudios ([14][42]) no solucionan el problema del reconocimiento de rostros térmico cuando

existe lapso temporal, sólo muestran que existe el problema. Además no utilizan otros métodos de reconocimiento de rostros en el dominio térmico para solucionar el problema.

En [13] se reproducen los estudios de [14][42] y se muestra que el método de reconocimiento de rostros “*Algoritmo Equinox*” obtiene un buen rendimiento para el reconocimiento de rostros en el espectro térmico y el visible, cuando existe esta variabilidad temporal. También se muestra que un esquema basado en la fusión de imágenes visibles y térmicas, es la mejor opción para realizar el reconocimiento de rostros en diferentes periodos de tiempo. No obstante, aún se requiere estudiar el fenómeno del lapso temporal con algoritmos de reconocimiento de rostros más sofisticados que los métodos basados en apariencia.

Al observar los resultados obtenidos de los Capítulos 3 y 4, específicamente las Tablas 8-9 y Figuras 69-70, cuando se usaron imágenes térmicas con variación temporal en la adquisición de datos, obtenidas bajo diferentes sesiones de captura en ambientes de interior y exterior, es posible ver que el desempeño en las tasas de reconocimiento de los métodos disminuyen bastante en comparación con los resultados obtenidos en una misma sesión. Debido a estos resultados, es que se hace necesario analizar en detalle el efecto de la variabilidad temporal en las imágenes térmicas.

Este capítulo analiza el desempeño de los métodos de reconocimiento de rostros cuando existe un lapso temporal en la adquisición de las imágenes térmicas. Para estudiar este problema se crearon dos bases de datos térmicas, la base de datos UCHThermalTemporalFace y la base UCHThermalFaceIndoorOutdoor. Se usan los enfoques utilizados en los Capítulos 3 y 4, de rostro térmico, *VascularNetwork* y *VascularThermalFace* para establecer qué métodos de reconocimiento de rostros son los más robustos ante este problema.

5.2 Análisis de reconocimiento usando imágenes térmicas con variabilidad temporal

Se analizará la efectividad y robustez de los métodos de reconocimiento de rostros cuando son usadas imágenes térmicas adquiridas en diferentes instantes de tiempo. Para realizar esta tarea, se utilizarán los métodos seleccionados en el Capítulo 3, histogramas LBP, histogramas de WLD, descriptores jets de Gabor, y las variantes de SIFT y el método SURF. Se utilizarán los 3 enfoques vistos en el Capítulo 4, de rostro térmico, *VascularNetwork* y *VascularThermalFace*, con el fin de ver si es posible obtener altas tasas de reconocimiento cuando estos enfoques son utilizados en conjunto con los métodos de reconocimiento de rostros. Además se buscará validar la metodología *VascularThermalFace* propuesta en el Capítulo 4 a través del análisis de sus resultados.

5.2.1 Descripción de la base de datos UCHThermalTemporalFace

Se construyó una base de datos llamada UCHThermalTemporalFace con imágenes térmicas obtenidas durante 69 días a un conjunto de 7 personas. Las imágenes fueron adquiridas en distintas horas del día, durante 50 sesiones, desde la hora inicial 0 hasta la hora final 1663, lo cual significa más de 69 días en la adquisición de datos para realizar este trabajo. Se usaron 7 sujetos para realizar este experimento, los cuales fueron cambiando su apariencia en forma natural durante el proceso de captura (diferentes cortes de pelo, barba, bigote, etc.). El experimento fue

realizado en un recinto de laboratorio, en un ambiente interior, donde en algunas ocasiones se utilizaba aire acondicionado para mantener constante la temperatura del ambiente. Las imágenes fueron obtenidas con una cámara térmica FLIR TAU 320 (ver Sección 2.3), ubicada aproximadamente a 1.1 m de los individuos de prueba, donde se capturó una imagen frontal por individuo. Para realizar el reconocimiento, se utilizaron las mejores variantes de cada método obtenidas en el Capítulo 3. Las variantes seleccionadas de cada método corresponden a LBP-HI-80, GJD-BC, WLD-EU-80, WLD-HI-80, SIFT-L&R-M, SIFT-Lowe-M y SURF-M. Se utilizaron los 3 enfoques analizados en el Capítulo 4 para evaluar las metodologías seleccionadas del Capítulo 3, el rostro térmico, *VascularNetwork* y *VascularThermalFace*. La explicación detallada de los métodos y sus parámetros se encuentran detallados en los Capítulos 2 y 3, y la explicación de las metodologías *VascularNetwork* y *VascularThermalFace*, se encuentran explicados en el Capítulo 4. La configuración de la cámara térmica fue realizada de acuerdo al Apartado 4.3.4 del Capítulo 4, es decir con parámetros automáticos de brillo y contraste, y un fondo fijo. La Figura 71 muestra imágenes del rostro térmico de un individuo capturado en varias sesiones de tiempo junto con los enfoques de *VascularNetwork* y *VascularThermalFace*.

Tal como se utilizó en el Capítulo 3 y 4, las imágenes fueron cortadas de acuerdo al tamaño seleccionado para cada método de reconocimiento obtenido en el Capítulo 3. Se utilizaron imágenes de tamaño 81x150 píxeles para el caso de LBP-HI-80 y WLD-X-80, y se utilizaron imágenes de tamaño 125x225 píxeles para el método GJD-BC, las variantes de SIFT y el método SURF.

Se crearon 50 conjuntos correspondientes a cada una de las sesiones, desde el conjunto C1 hasta el conjunto C50. Los conjuntos fueron creados tomando una imagen térmica de cada sujeto (en total 7) en cada una de las 50 sesiones. Se realiza el reconocimiento de rostros para cada método seleccionado utilizando como conjunto de galería el conjunto C1, obtenido de la primera sesión, versus todos los conjuntos creados para la base de datos UCHThermalTemporalFace, desde el C1 hasta el C50. Un esquema representativo de los conjuntos C1-C50 versus el número de horas en que fueron capturados, es presentado en la Figura 72.

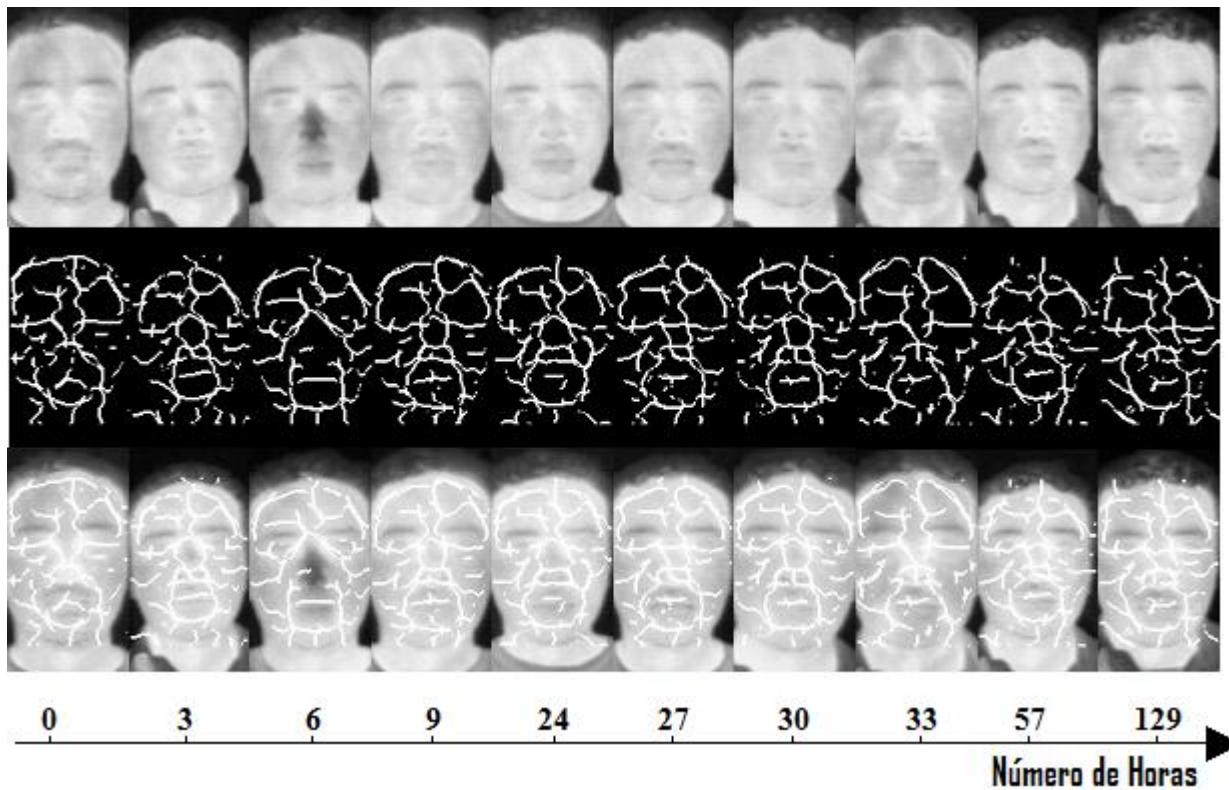


Figura 71. Ejemplo de imágenes para la base de datos UCHThermalTemporalFace, tamaño 81x150 píxeles. Imágenes de rostro térmico, *VascularNetwork* y *VascularThermalFace* para 10 sesiones distintas de tiempo. Los números representan la hora en la cual fue obtenida la imagen térmica con respecto a la primera sesión (diferencia temporal en adquisición de datos). Es posible observar la variabilidad temporal en las imágenes y los cambios temporales producto de la temperatura corporal del sujeto.

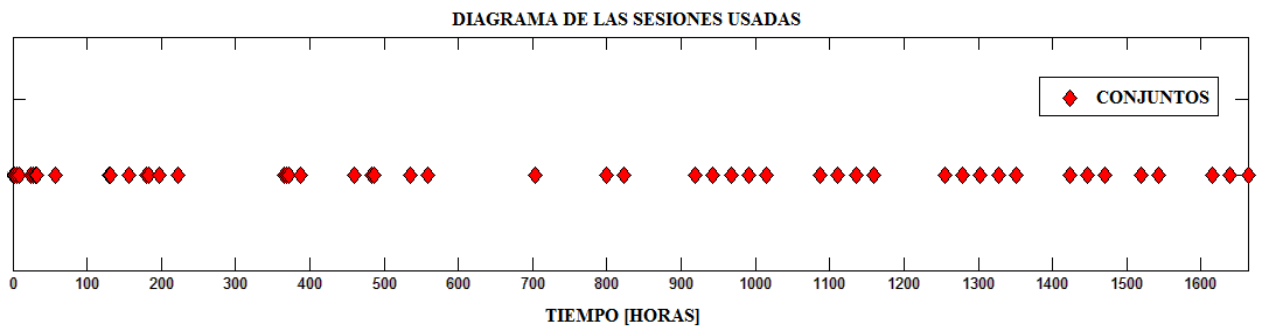


Figura 72. Diagrama representativo de las sesiones de captura realizadas en el laboratorio. Cada rombo representa un conjunto de imágenes térmicas de 7 sujetos en el periodo de tiempo de 69 días (1663 horas).

5.2.2 Descripción de los experimentos

Para evaluar y analizar el desempeño los métodos de reconocimiento de rostros, se realizaron 4 experimentos: i) reconocimiento de rostros en el tiempo, ii) reconocimiento de rostros en el tiempo con oclusión parcial, iii) reconocimiento de rostros en el tiempo con ruido en la posición de los ojos, y iv) reconocimiento de rostros en el tiempo con oclusión parcial del rostro y ruido de 10% en la posición de los ojos. La descripción de los experimentos es mostrada a continuación:

Reconocimiento de rostros en el tiempo: Este experimento consiste en realizar el reconocimiento de rostros en el tiempo, es decir, calcular las tasas de reconocimiento obtenidas para cada algoritmo de reconocimiento de rostros. Para esto se utiliza el conjunto C1, de la primera sesión como galería, versus los conjuntos de test, que corresponden a todas las sesiones que van desde el conjunto C1 (inclusive) hasta el C50. Este experimento analiza el desempeño de los métodos de reconocimiento cuando existe diferente periodo en la adquisición de las imágenes de rostros térmicos.

Reconocimiento de rostros en el tiempo con oclusión parcial: Este experimento trata de ver el comportamiento de los métodos de reconocimiento de rostros cuando existe una oclusión parcial del rostro. La oclusión de las imágenes térmicas se realiza de la misma manera que en el Capítulo 3, la oclusión es aleatoria para cada uno de los conjuntos C1-C50, esto quiere decir que cada imagen es dividida en 10 secciones y una de éstas es llevada a valor de intensidad 0 (pintada de color negro) aleatoriamente. Se usa como galería el conjunto C1 sin tener algún tipo de oclusión versus las demás sesiones de tiempo con oclusión parcial del rostro. Ejemplos de imágenes con oclusión parcial del rostro para la base UCHThermalTemporalFace son mostrados en la Figura 73.

Reconocimiento de rostros en el tiempo con ruido en la posición de los ojos: Se analiza el comportamiento de los métodos de reconocimiento de rostros bajo el efecto de una mala alineación de éste cuando es agregado ruido en la posición de la anotación de los ojos. Se utiliza ruido independientemente en la posición x y posición y de los ojos, de valores 2,5%, 5% y 10% de la distancia entre ojos para cada uno de los conjuntos C1-C50 de la base de datos UCHThermalTemporalFace. Se usa como galería el conjunto C1 sin tener algún tipo de ruido en la posición de los ojos versus las 50 sesiones (C1-C50), que poseen ruido en la posición de los ojos. Ejemplo de imágenes con ruido en posición de los ojos es mostrado en la Figura 74.

Reconocimiento de rostros en el tiempo con oclusión parcial del rostro y ruido de 10% en la posición de los ojos: Este experimento reúne los dos casos anteriores de oclusión parcial del rostro y ruido en la posición de los ojos. Primero se introduce ruido de 10% en la posición de los ojos y luego se realiza la oclusión parcial. Con este experimento se pretende encontrar un método robusto ante ambas perturbaciones.

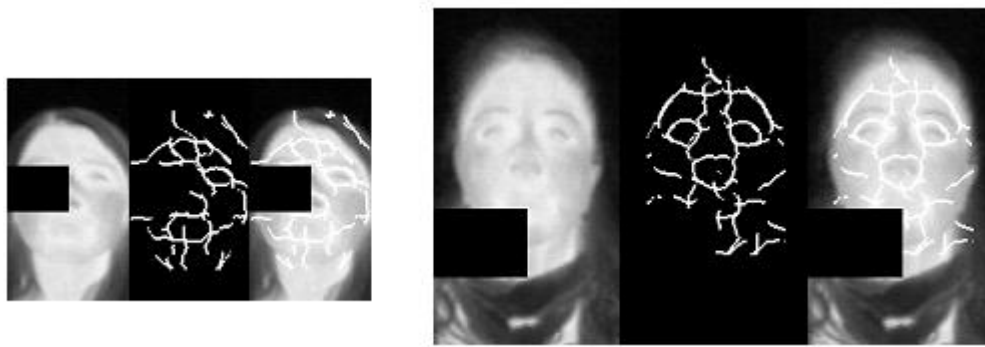


Figura 73. Ejemplos de imágenes con oclusión parcial (base de datos UCHThermalTemporalFace). (Izquierda) Tamaño de la ventana 81x150 píxeles. (Derecha) Tamaño de la ventana 125x225. Se observan los 3 enfoques usados: rostro térmico, *VascularNetwork* y *VascularThermalFace*.

5.2.3 Resultados de reconocimiento

El desempeño de los diferentes métodos es evaluado con la tasa de reconocimiento utilizando el criterio del Top-1, que consiste en encontrar el individuo que posea la distancia mínima entre un conjunto de galería y un sujeto de prueba (test), que comparta el mismo identificador que el sujeto de prueba. En todos los experimentos, el conjunto C1, obtenido en la primera sesión, sin incluir ninguna oclusión y sin incluir ruido en la posición de los ojos, es utilizado como conjunto de galería debido a que se quiere realizar el reconocimiento en el tiempo de la primera sesión versus todas las sesiones (C1-C50). Note que para el conjunto de test se incluye C1 debido a que se analiza el efecto cuando C1 tiene oclusión parcial del rostro y ruido en la posición de los ojos.

En las Tablas 16-21 se observa la tasa de reconocimiento promedio obtenida por cada método y su desviación estándar. Se utilizaron 3 enfoques para realizar el reconocimiento de rostros: i) rostro térmico, ii) *VascularNetwork*, y iii) *VascularThermalFace*. Para cada sesión, se calculó la tasa de reconocimiento obtenida por cada método de reconocimiento, luego se obtuvo el promedio de la tasa de reconocimiento de las 50 sesiones y la desviación estándar.

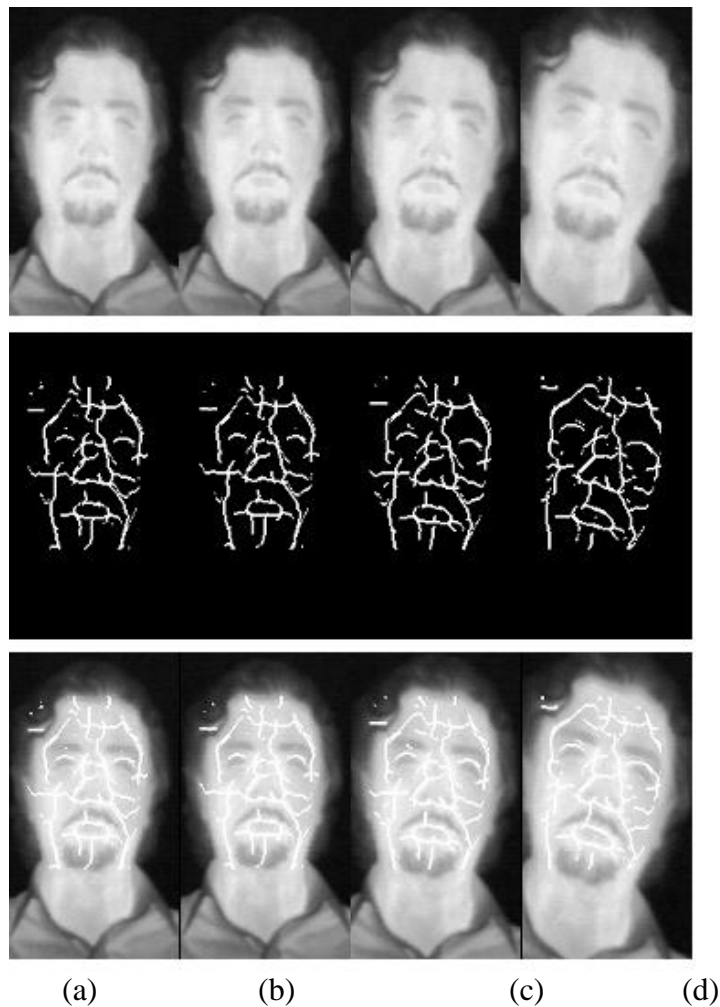


Figura 74. Ruido en la posición de los ojos. Ejemplos de imágenes alineadas después de aplicar ruido en la posición de los ojos (base de datos UCHThermalTemporalFace). Tamaño de la ventana 125x225 píxeles. (a) Sin ruido; (b) 2.5% ruido; (c) 5% ruido; (d) 10% ruido. Se observan los 3 enfoques usados: rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Los resultados son resumidos para cada uno de los experimentos:

Reconocimiento de rostros en el tiempo: La Tabla 16 muestra el desempeño de los diferentes métodos cuando es realizado el reconocimiento de rostros en el tiempo. Las Figuras 75-77 muestran los diagramas temporales del desempeño de los métodos de reconocimiento para los enfoques de rostro térmico, *VascularNetwork* y *VascularThermalFace*. Se observa de las Figuras 75-76 que es difícil visualizar cuales métodos obtienen una buena tasa de reconocimiento durante las 50 sesiones, sin embargo los métodos SIFT y SURF son los que visualmente obtienen el desempeño más bajo de reconocimiento. La Figura 77 muestra claramente que los métodos obtienen una mejora significativa en la tasa de desempeño durante el experimento de reconocimiento de rostros en el tiempo, debido a que se puede visualizar un aumento en la tasa de desempeño de todos los métodos.

Se puede concluir de las Figuras 75-77 y la Tabla 16 que cuando se utiliza el rostro térmico con métodos basados en calces locales como LBP-HI-80, WLD-X-80, GJD-BC, éstos obtienen buenos resultados (sobre ~90% en la tasa de reconocimiento), donde el mejor desempeño es obtenido con el método GJD-BC 125x225 con 96.6 %, seguido por el método WLD-HI-80 81x150 con un 94.9%. En cambio los métodos basados en calces globales, como las variantes de SIFT y el método SURF obtienen bajo rendimiento, representado en bajas tasas de reconocimiento en comparación a los métodos basados en calces locales. Las bajas tasas de reconocimiento obtenidas con los métodos basados en calces globales, se deben a los cambios que sufren los individuos anatómicamente durante las 50 sesiones, debido al cambio natural que sufren los individuos a lo largo del tiempo transcurrido en la adquisición de las imágenes térmicas, lo que se ve reflejado en las variaciones del contraste de las imágenes térmicas, y que dificulta la tarea de encontrar descriptores parecidos para esta clase de métodos.

Cuando se utiliza la metodología *VascularNetwork*, se puede concluir que los métodos de reconocimiento de rostros LBP-HI-80 y las variantes de WLD obtienen buenos resultados al utilizar este tipo de información, seguido del método SURF. Esto se logra, y como se explicó en el Capítulo 4, Sección 4.3.3, los métodos LBP y WLD pueden generar una buena representación de la estructura de la red vascular que es ordenada en distintos niveles de localidad, y puede ser reconocida de manera eficiente usando alguna medida de disimilitud. Los resultados de reconocimiento para LBP y las variantes de WLD son sobre ~90% de reconocimiento, con un resultado de 96.3% de reconocimiento para el método LBP-HI-80. Para el método SURF, se concluye que funciona adecuadamente, debido a que la etapa de calces SURF calza de mejor manera los descriptores obtenidos de la imagen de red vascular. Las variantes de SIFT y el método GJD obtienen bajo desempeño de reconocimiento utilizando esta clase de información, debido a que SIFT obtiene muchos descriptores que son calzados inadecuadamente en la etapa de calces, como se mostró en el Capítulo 4, y el método GJD opera mal debido a que no está diseñado para trabajar con imágenes binarias, debido a que al utilizar los jets, estos pueden estar en zonas negras (con valor de intensidad 0) y no genera información de reconocimiento.

Al utilizar la metodología propuesta *VascularThermalFace*, se obtiene un buen desempeño de reconocimiento para los métodos evaluados, obteniendo mejoras significativas en los resultados para los métodos basados en calces globales y manteniendo el desempeño de reconocimiento en los basados en calces locales. Se puede destacar el rendimiento de GJD-BC con un 96.6% en la tasa de reconocimiento (similar resultado obtenido utilizando la información del rostro térmico), y las mejoras en las tasas de reconocimiento de SIFT-Lowe-M 125x225 y SURF-M 125x225 desde un 73.7% a un 89.1%, y desde un 76.6% a un 92.6%, respectivamente. Esta metodología mejora los métodos basados en calces globales debido a que incorpora la información de la red vascular en el rostro térmico lo cual facilita el cálculo y matching de descriptores.

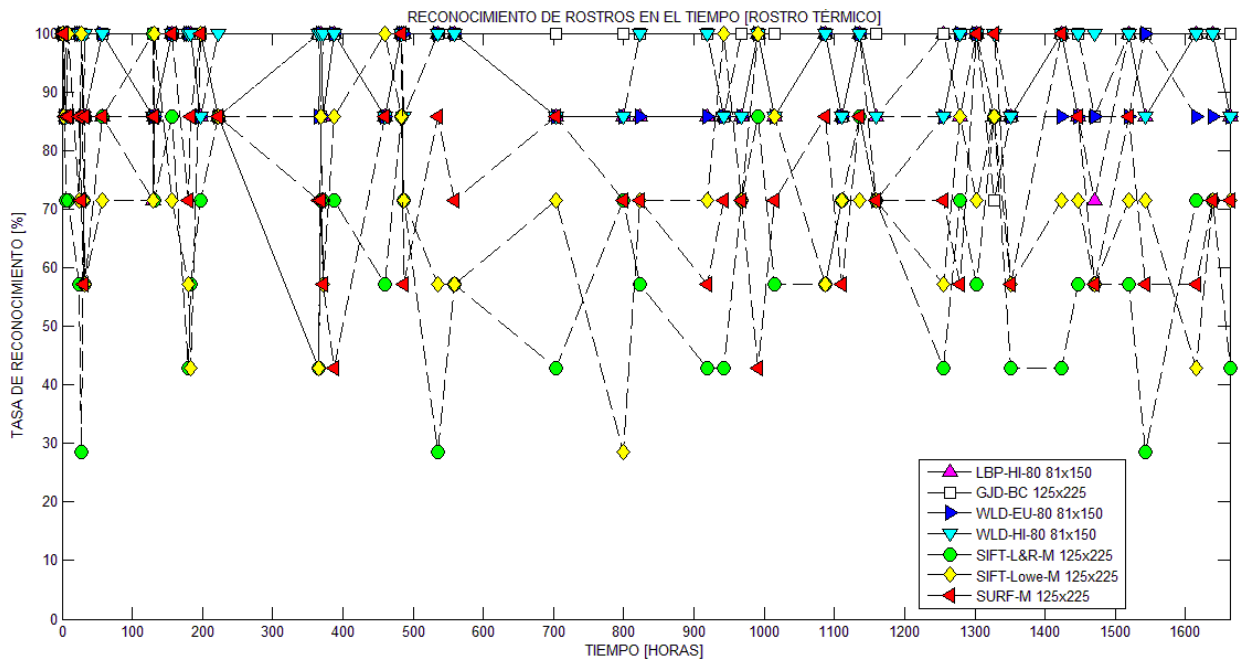


Figura 75. Reconocimiento de rostros en el tiempo utilizando el enfoque de rostro térmico. Se observa el desempeño de los métodos de reconocimiento versus el número de horas correspondiente al periodo de adquisición de las imágenes térmicas en cada sesión.

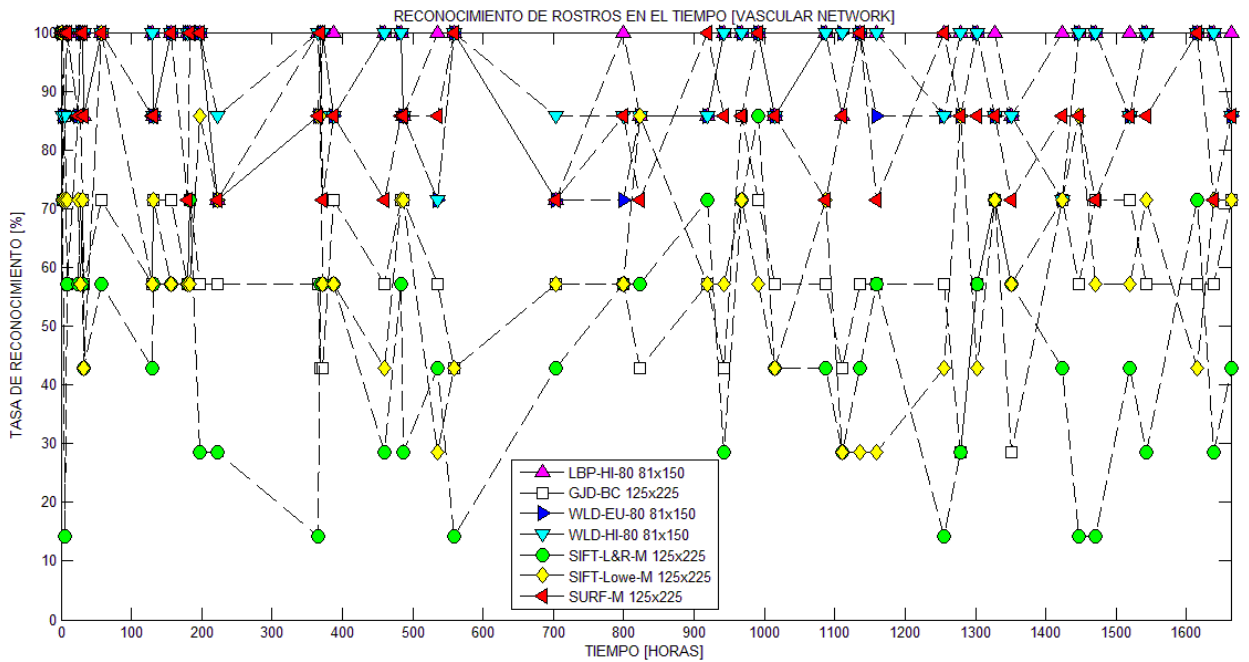


Figura 76. Reconocimiento de rostros en el tiempo utilizando el enfoque de *VascularNetwork*. Se observa el desempeño de los métodos de reconocimiento versus el número de horas correspondiente al periodo de adquisición de las imágenes térmicas en cada sesión.

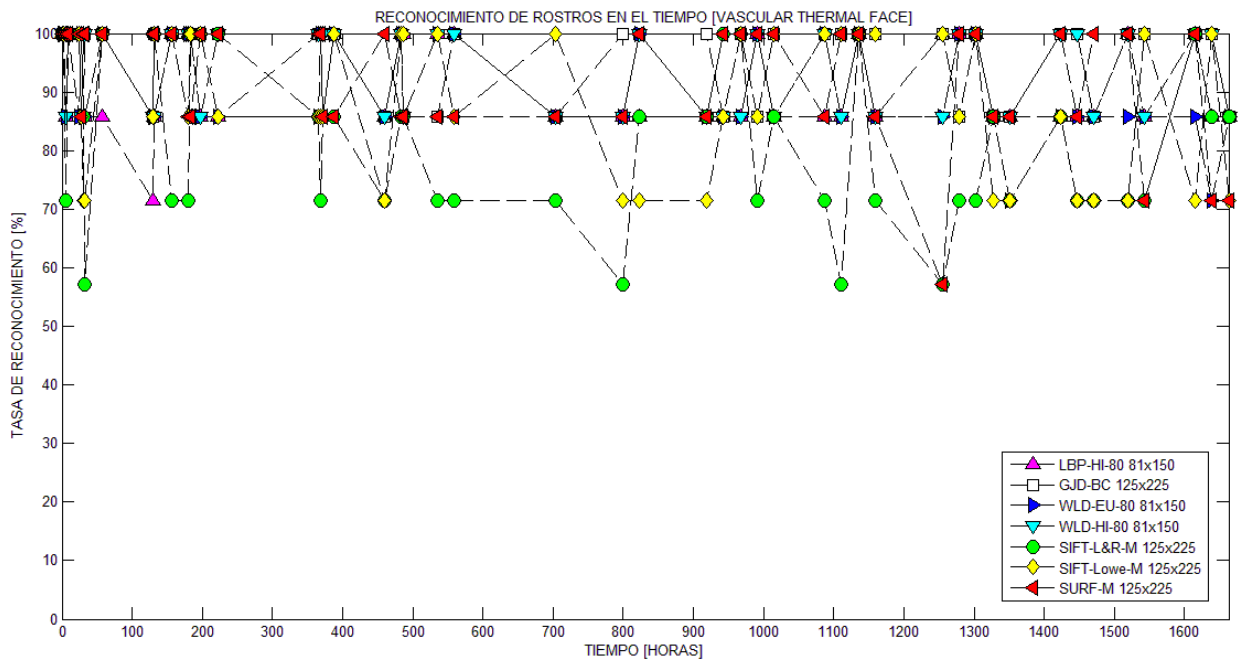


Figura 77. Reconocimiento de rostros en el tiempo utilizando el enfoque de *VascularThermalFace*. Se observa el desempeño de los métodos de reconocimiento versus el número de horas correspondiente al periodo de adquisición de las imágenes térmicas en cada sesión.

Como resumen, para cada uno de los métodos, el mejor resultado obtenido para cada enfoque de imágenes: rostro térmico, *VascularNetwork* y *VascularThermalFace*, considerando los diferentes parámetros y tamaños de ventana usados (obtenidos del Capítulo 3), es dado por las siguientes variantes:

- GJD-BC, 125x225 pixeles, *VascularThermalFace*: 96.6%
- GJD-BC, 125x225 pixeles, Rostro Térmico: 96.6%
- LBP-HI-80 81x150 pixeles, *VascularNetwork*: 96.3%
- WLD-HI-80 81x150 pixeles, Rostro Térmico: 94.9%
- SURF-M 125x225 pixeles, *VascularThermalFace*: 92,6%
- WLD-EU-80 81x150 pixeles, *VascularNetwork*: 92.0%
- SIFT-Lowe-M 125x225 pixeles, *VascularThermalFace*: 89.1%
- SIFT-L&R-M 125x225 pixeles, *VascularThermalFace*: 82.6%

Tabla 16. Resultados de los experimentos de reconocimiento de rostros en el tiempo, base de datos UCHThermalTemporalFace. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	92.0	8.2	96.3	7.5	91.7	7.7
GJD-BC 125x225	96.6	6.8	61.7	14.3	96.6	6.2
WLD-EU-80 81x150	91.1	7.6	92.0	9.6	91.1	8.1
WLD-HI-80 81x150	94.9	6.9	93.7	8.2	93.7	7.2
SIFT-L&R-M 125x225	64.0	18.3	47.4	20.1	82.6	13.9
SIFT-Lowe-M 125x225	73.7	17.4	62.6	17.5	89.1	12.1
SURF-M 125x225	76.6	15.7	87.4	10.6	92.6	10.1

Reconocimiento de rostros en el tiempo con oclusión parcial: La Tabla 17 muestra la sensibilidad de los métodos respecto a la oclusión parcial de una región de la cara, utilizando los 3 enfoques de la imagen: rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Se observa de la Tabla 17, para el caso de rostro térmico, que el mejor desempeño es alcanzado por el método WLD-HI-80 81x150 con un resultado de 94.3%, seguido de LBP-HI-80 81x150 con un resultado de 91.7%. Se observa que para este caso, los métodos GJD-BC, las variantes de SIFT y SURF son más afectados por la oclusión, que los métodos LBP-X-X y WLD-X-X, lo que concuerda con lo obtenido en el Capítulo 3 para el caso de oclusión.

Para el caso de *VascularNetwork*, se observa que la oclusión parcial del rostro hace que todos los métodos disminuyan su desempeño debido a que son más afectados por la oclusión, excepto en el caso de las variantes de WLD donde existe una disminución pero no es tan significativa como los demás métodos.

Utilizando la metodología propuesta *VascularThermalFace*, los resultados obtenidos para la oclusión son buenos, debido a que los métodos más afectados por la oclusión como GJD-BC, y las variantes SIFT, aumentan su desempeño de reconocimiento al usar este enfoque. Esto es debido a que la combinación de la red vascular con el rostro térmico hace que estos métodos mejoren su desempeño cuando existe una oclusión parcial del rostro. Las demás metodologías obtienen buenos resultados para este enfoque manteniendo los resultados de cuando no existía oclusión parcial del rostro térmico.

Tabla 17. Resultados de los experimentos de oclusión parcial del rostro térmico, base de datos UCHThermalTemporalFace. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	91.7	8.7	89.5	10.3	91.1	8.6
GJD-BC 125x225	79.1	14.8	52.3	13.7	87.7	14.1
WLD-EU-80 81x150	90.3	8.4	87.1	12.9	89.7	8.2
WLD-HI-80 81x150	94.3	7.6	90.0	10.9	93.7	7.2
SIFT-L&R-M 125x225	54.0	17.5	43.4	17.5	71.1	16.5
SIFT-Lowe-M 125x225	57.4	18.8	54.2	17.1	78.3	15.3
SURF-M 125x225	60.3	20.7	79.7	14.7	84.3	13.6

Reconocimiento de rostros en el tiempo con ruido en la posición de los ojos: Las Tablas 18-20 muestran la sensibilidad de cada método con respecto a la precisión en la detección de ojos para los casos de 2.5%, 5% y 10% de ruido respectivamente. Se puede observar de las Tablas 18-20 que los métodos basados en calces locales como LBP, GJD-BC y las variantes de WLD son muy sensibles a variaciones en la posición de los ojos cuando el ruido es significativo, en este caso para 5% y 10% de ruido en la posición de los ojos. En cambio se observa que los resultados de reconocimiento obtenidos por las variantes de SIFT y el método SURF son muy robustos a variaciones en la posición de los ojos, lo que concuerda con el análisis obtenido en el Capítulo 3 para el caso de precisión en la detección de ojos.

Para todos los enfoques de las imágenes, rostro térmico, *VascularNetwork* y *VascularThermalFace*, el comportamiento de los métodos es similar, al aumentar el ruido en la posición de los ojos el desempeño de los métodos de reconocimiento disminuye. Sin embargo, los métodos SIFT y SURF, mantienen su desempeño, ya que son robustos al ruido en la posición de los ojos producto de su etapa de calces, la cual fue explicada en el Capítulo 3.

Al utilizar un ruido muy grande para desalinearse las imágenes, en este caso de 10% de ruido en la posición de los ojos, los resultados de la Tabla 20 muestran que el desempeño de los métodos empeora para el caso del rostro térmico y *VascularNetwork*. Sin embargo, al analizar el desempeño de los métodos de reconocimiento de rostros ante la metodología propuesta *VascularThermalFace*, es posible observar los buenos resultados de este enfoque, obteniendo mejoras en todos los métodos de reconocimiento de rostros. Además al usar este enfoque con el método SIFT y el método SURF, se logra un desempeño alto en comparación con los demás métodos estudiados.

Tabla 18. Resultados de los experimentos de 2.5% de ruido en posición de los ojos del rostro térmico, base de datos UCHThermalTemporalFace. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	92.0	8.2	96.0	7.6	92.6	7.8
GJD-BC 125x225	98.0	6.5	62.6	13.5	95.4	8.4
WLD-EU-80 81x150	91.7	7.8	92.6	10.5	91.1	8.1
WLD-HI-80 81x150	94.3	7.1	94.3	8.7	94.3	7.1
SIFT-L&R-M 125x225	63.7	18.1	45.7	19.9	85.1	12.9
SIFT-Lowe-M 125x225	72.5	19.7	60.6	18.4	87.1	12.9
SURF-M 125x225	73.1	15.9	84.2	12.6	90.0	10.9

Reconocimiento de rostros en el tiempo con oclusión parcial del rostro y ruido de 10% en la posición de los ojos: La Tabla 21 refleja la sensibilidad de cada método con respecto a la precisión en la detección de ojos con 10% de ruido en conjunto con la oclusión parcial del rostro.

Del análisis anterior, obtenido al observar las Tablas 18-19-20, es posible determinar que el ruido en la posición de los ojos afecta más a los métodos basados en calces locales como histogramas LBP y WLD, y GJD-BC en comparación con las variantes de SIFT y SURF que son más robustos ante esta clase de perturbación. De la Tabla 17, se puede decir que los métodos basados en calces globales (SIFT y SURF) son más sensibles al efecto de la oclusión parcial del rostro que los métodos LBP, WLD y GJD-BC. En base a esto, al observar la Tabla 21, es posible concluir que el conjunto de perturbaciones de ruido y oclusión afecta en un rango superior al desempeño de los métodos de reconocimiento analizados. Esto se ve reflejado en el desempeño de todos los métodos de reconocimiento, en especial cuando se usan los enfoques de rostro térmico y *VascularNetwork*, debido a que las tasas de desempeño bajan considerablemente en comparación con los casos de oclusión y ruido cuando se utilizan de forma separada. Sin embargo para el caso *VascularThermalFace*, las tasas de reconocimiento de los métodos disminuyen en menor cantidad que los enfoques de rostro térmico y *VascularNetwork*. Para este caso, todos los métodos obtienen similares tasas de reconocimiento del orden del ~74%, exceptuando al método WLD que obtiene una tasa de reconocimiento de 83.1%, es decir que disminuye su tasa original de reconocimiento (cuando no había ruido ni oclusión) en aproximadamente ~12% producto de las perturbaciones, logrando así ser uno de los métodos más eficientes y robustos ante esta clase de alteraciones.

Tabla 19. Resultados de los experimentos de 5% de ruido en posición de los ojos del rostro térmico, base de datos UCHThermalTemporalFace. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	89.7	10.8	93.4	10.1	90.0	9.7
GJD-BC 125x225	94.3	8.2	58.0	15.9	92.3	10.5
WLD-EU-80 81x150	90.0	8.8	86.9	11.1	90.0	8.8
WLD-HI-80 81x150	93.4	9.3	89.1	9.8	92.6	9.7
SIFT-L&R-M 125x225	64.3	16.1	45.4	15.4	81.1	17.2
SIFT-Lowe-M 125x225	74.0	16.7	58.6	18.9	84.6	13.5
SURF-M 125x225	72.3	17.2	82.6	12.3	88.6	11.2

Tabla 20. Resultados de los experimentos de 10% de ruido en posición de los ojos del rostro térmico, base de datos UCHThermalTemporalFace. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	78.6	12.9	82.0	13.2	78.9	12.3
GJD-BC 125x225	83.7	14.4	50.0	15.3	84.9	13.9
WLD-EU-80 81x150	80.0	12.9	78.3	15.3	80.0	13.5
WLD-HI-80 81x150	80.6	14.1	78.0	15.9	83.7	12.2
SIFT-L&R-M 125x225	62.0	18.6	46.5	21.7	79.1	18.7
SIFT-Lowe-M 125x225	72.6	17.9	55.1	17.3	87.4	13.7
SURF-M 125x225	71.7	16.7	79.7	16.3	88.0	12.0

Tabla 21. Resultados de los experimentos de 10% de ruido en posición de los ojos con oclusión parcial del rostro térmico, base de datos UCHThermalTemporalFace. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	77.7	12.3	72.0	16.1	76.0	12.4
GJD-BC 125x225	66.9	15.9	43.4	12.8	75.4	19.2
WLD-EU-80 81x150	78.6	13.6	71.7	18.4	75.4	14.2
WLD-HI-80 81x150	79.7	13.3	75.1	16.8	83.1	10.7
SIFT-L&R-M 125x225	53.7	18.8	42.6	19.9	68.9	18.9
SIFT-Lowe-M 125x225	54.6	23.7	47.7	20.3	72.0	17.1
SURF-M 125x225	55.7	17.9	69.1	16.9	75.1	18.4

5.2.4 Discusión

Como resultado para este apartado, se observa que ante cambios de temperatura corporal, variaciones en el ambiente, cambios de aspecto de los individuos, variaciones en la pose y cambios de expresión, es posible generar un sistema de reconocimiento de rostros térmico en el tiempo que tenga una tasa elevada de reconocimiento utilizando la metodología propuesta *VascularThermalFace*. Esto permite validar el uso de esta metodología para el desarrollo de sistemas de reconocimiento de rostro que utilicen imágenes térmicas.

Como se observó, se obtienen buenos resultados de reconocimiento de rostros para el caso del método GJD-BC, las variantes de WLD y el caso de LBP-HI-80, sobre un ~90% de reconocimiento en promedio, para los distintos enfoques estudiados en este capítulo. Sin embargo, se puede concluir que al utilizar una variante de calces globales como SIFT o SURF, éstas poseen un bajo desempeño para el caso de rostro térmico y *VascularNetwork*, cuando se realiza el reconocimiento de rostros en el tiempo. Esto es debido a cambios de intensidad de los píxeles, cambios de aspecto y variaciones en la pose, que hacen que estos métodos obtengan bajas tasas de reconocimiento en comparación con métodos basados en calces locales. Este desempeño puede ser mejorado utilizando la metodología *VascularThermalFace*, la cual permite a estos métodos aumentar sus tasas de reconocimiento debido a que incorpora información de la red vascular en conjunto con la información del rostro, la cual sirve como ayuda a métodos basados en calces globales para realizar el reconocimiento de rostros en el tiempo.

En gran medida sólo algunos de los métodos estudiados son robustos a oclusión y al ruido en la posición de los ojos en el tiempo. De los métodos estudiados los métodos más robustos a oclusión corresponden a las variantes de WLD y LBP-X-X, mientras que los más robustos a alineamiento y ruido de los ojos son las variantes SIFT y SURF. El método GJD-BC es el algoritmo que resulta más afectado ante perturbaciones de oclusión y alineamiento, lo que concuerda con lo analizado en el Capítulo 3 y 4.

Acerca de que método de reconocimiento obtiene los mejores resultados, observando y analizando las Tablas 16-21, es posible ver que sin oclusión ni ruido en la posición de los ojos, los métodos que operan de mejor manera, obteniendo los mejores resultados son GJD-BC, para el caso de rostro térmico y el enfoque *VascularThermalFace*; LBP-HI-80 para el caso de *VascularNetwork*; WLD-HI-80 para el caso de rostro térmico y *VascularThermalFace*; y SURF-M para *VascularThermalFace*. Analizando las perturbaciones como oclusión y ruido en la posición de los ojos, los métodos que obtienen los mejores resultados son las variantes de WLD-X-80, para el caso de rostro térmico y *VascularThermalFace*; y SURF-M para el caso *VascularThermalFace*. Por estos motivos, se puede decir que los métodos basados en histogramas de WLD son una excelente elección si es necesario un método robusto y estable, que opere en tiempo real (resultados del Capítulo 3), y que obtenga tasas de reconocimiento altas (resultados Capítulos 3, 4 y 5), en combinación con la metodología propuesta *VascularThermalFace*, que permite la robustez del método cuando existe mala alineación de las imágenes térmicas.

5.3 Análisis de reconocimiento usando diferentes galerías con variabilidad temporal

Este estudio tiene por finalidad realizar un análisis de los métodos de reconocimiento cuando son utilizadas imágenes de rostros térmicas obtenidas en diferentes sesiones de interior versus imágenes de rostros de test obtenidos en sesiones en exterior y viceversa, con distinto tiempo en la adquisición de datos (variabilidad temporal). La idea principal de este experimento es comparar bajo distintas variables (variabilidad temporal, condiciones ambientales diferentes, cambios corporales de los individuos) el desempeño de cada uno de los métodos seleccionados y verificar que es posible obtener tasas de reconocimientos altas, bajo condiciones de interior y exterior.

Este estudio se presenta debido a que en el Capítulo 3 y Capítulo 4, se obtuvieron bajos resultados al analizar el experimento de diferentes galerías de interior versus exterior para los métodos de reconocimiento de rostros. Debido a esto, es preciso profundizar qué métodos obtienen altas tasas de reconocimiento de reconocimiento considerando un análisis exhaustivo de imágenes térmicas adquiridas en diferentes sesiones de interior y exterior.

Los bajos resultados mostrados por los algoritmos de reconocimiento en los Capítulos 3 y 4, se deben a que las imágenes térmicas obtenidas para la base UCHThermalFace se encontraban con píxeles saturados producto que el experimento fue realizado en un ambiente exterior, en verano, con temperaturas sobre 33°C en donde la cámara recibía directamente la radiación solar. Debido a esto, las imágenes térmicas de un sujeto obtenidas en las sesiones de interior y exterior no eran similares producto de la saturación, lo cual llevaba a que los resultados entregados por los métodos de reconocimiento fueran bajos.

Como se mostro en el Sección 4.3.4, se realizó una nueva configuración de la cámara térmica para realizar el Capítulo 4 y 5. Se usó una configuración de la cámara térmica con parámetros automáticos, manteniendo la escena con un fondo fijo para mantener la uniformidad en la captura de la imagen. Con estos parámetros, la escena se mantiene constante y varía levemente producto de los cambios corporales del individuo y el ambiente. El análisis de la configuración de la cámara se encuentra explicado en el Capítulo 4, Sección 4.3.4.

5.3.1 Descripción de la base de datos UCHThermalFaceIndoorOutdoor

Se construyó una base de datos llamada UCHThermalFaceIndoorOutdoor con imágenes térmicas de 20 personas durante 4 días (no seguidos). El día 1 y 3 se realizaron las sesiones de interior y el día 2 y 4 las sesiones de exterior. Las sesiones de interior fueron realizadas bajo situaciones diferentes, una con aire acondicionado y otra sin éste. Mientras que en el experimento de exterior, al igual que en el caso interior, fue realizado bajo 2 situaciones diferentes, una con la cámara no expuesta al sol y otra con la cámara expuesta directamente a éste, donde la temperatura ambiente máxima alcanzada para este experimento fue de 31°C.

Las imágenes fueron obtenidas con una cámara térmica FLIR TAU 320 (ver Sección 2.3) ubicada aproximadamente a 1.1 m de los individuos de prueba, donde se capturó una imagen frontal por individuo. Para el caso de exterior, se usó una cubierta para proteger la cámara de los rayos del sol con el fin de no saturar tanto la cámara como en el experimento de galerías del Capítulo 3. Todas las sesiones fueron capturadas con una diferencia temporal de entre uno a dos días aproximadamente.

Al igual que en el experimento de análisis temporal de la base UCHThermalTemporalFace, para realizar el reconocimiento, se utilizan las mejores variantes de cada método obtenidas en el Capítulo 3. Las variantes seleccionadas de cada método corresponden a LBP-HI-80, GJD-BC, WLD-EU-80, WLD-HI-80, SIFT-L&R-M, SIFT-Lowe-M y SURF-M. Se utilizan los 3 enfoques utilizados en el Capítulo 4 para evaluar las metodologías seleccionadas del Capítulo 3, el rostro térmico, *VascularNetwork* y *VascularThermalFace*. La configuración de la cámara térmica es realizada de acuerdo a lo descrito en la Sección 4.3.4 del Capítulo 4, es decir con parámetros automáticos de brillo y contraste. Tal como se utilizó en el Capítulo 3 y 4, las imágenes fueron cortadas de acuerdo al tamaño seleccionado para cada método de reconocimiento obtenido en el Capítulo 3. Se utilizaron imágenes de tamaño 81x150 píxeles para el caso de LBP-HI-80 y WLD-X-80, y se utilizaron imágenes de tamaño 125x225 píxeles para el método GJD-BC, las variantes de SIFT y el método SURF.

Se construyeron 4 conjuntos llamados I1, I2, O1, O2, donde I representa la sesión interior y la O la sesión exterior. Cada uno de estos conjuntos está constituido por 20 imágenes de rostros térmicos obtenidas en cada una de las sesiones. El conjunto I1 representa al conjunto de interior sin aire acondicionado, el conjunto I2 representa al conjunto con aire acondicionado, el conjunto O1 representa el conjunto exterior con la cámara aislada al contacto solar, y el conjunto O2 representa el conjunto con la cámara expuesta al sol. Un ejemplo de imágenes de rostro de un sujeto para cada una de las sesiones y utilizando los 3 enfoques de rostro térmico, *VascularNetwork* y *VascularThermalFace*, es mostrado en la Figura 78.

5.3.2 Descripción del experimento

Para evaluar y analizar los métodos de reconocimiento para este apartado, se utiliza un experimento de galerías de interior versus test de exterior en el tiempo y viceversa. El experimento consiste en realizar el reconocimiento de rostros utilizando cada conjunto como galería y como test con el fin de realizar una Tabla de datos cruzada, que muestre los desempeños de cada método de reconocimiento de rostros. Cada uno de los 4 conjuntos {I1, I2, O1, O2} es utilizado como galería y test, es decir, se forman las siguientes 16 combinaciones de pares de

conjuntos: {I1/I1; I1/I2; I1/O1; I1/O2; I2/I1; I2/I2; I2/O1; I2/O2; O1/I1; O1/I2; O1/O1; O1/O2; O2/I1; O2/I2; O2/O1; O2/O2}. Es importante mencionar que al usar el mismo conjunto de test versus galería ({I1/I1; I2/I2; O1/O1; O2/O2}), el resultado es 100% de reconocimiento. Se evalúa el desempeño de los métodos de reconocimiento usando el reconocimiento del más parecido (Top-1) para generar la Tabla de datos cruzada de reconocimiento. Las Tablas de validación cruzada son calculadas para cada uno de los enfoques de imágenes usados: rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Los resultados obtenidos por los métodos de reconocimiento de rostros son mostrados resumiendo las 16 combinaciones de pares de conjuntos en sólo 6 combinaciones que entregan todo el comportamiento del experimento. Se eliminan los conjuntos de galería y test cuando son el mismo, y para los demás conjuntos se calculan los promedios entre los pares de conjuntos XY e YX. Finalmente, se obtienen las siguientes combinaciones para analizar: {I1/I2; I1/O1; I1/O2; I2/O1; I2/O2; O1/O2}.

También es calculada la media y desviación estándar sobre cada una de las Tablas de validación cruzada para generar una Tabla resumen que muestre los resultados promedios para cada método analizado y por cada uno de los enfoques usados.

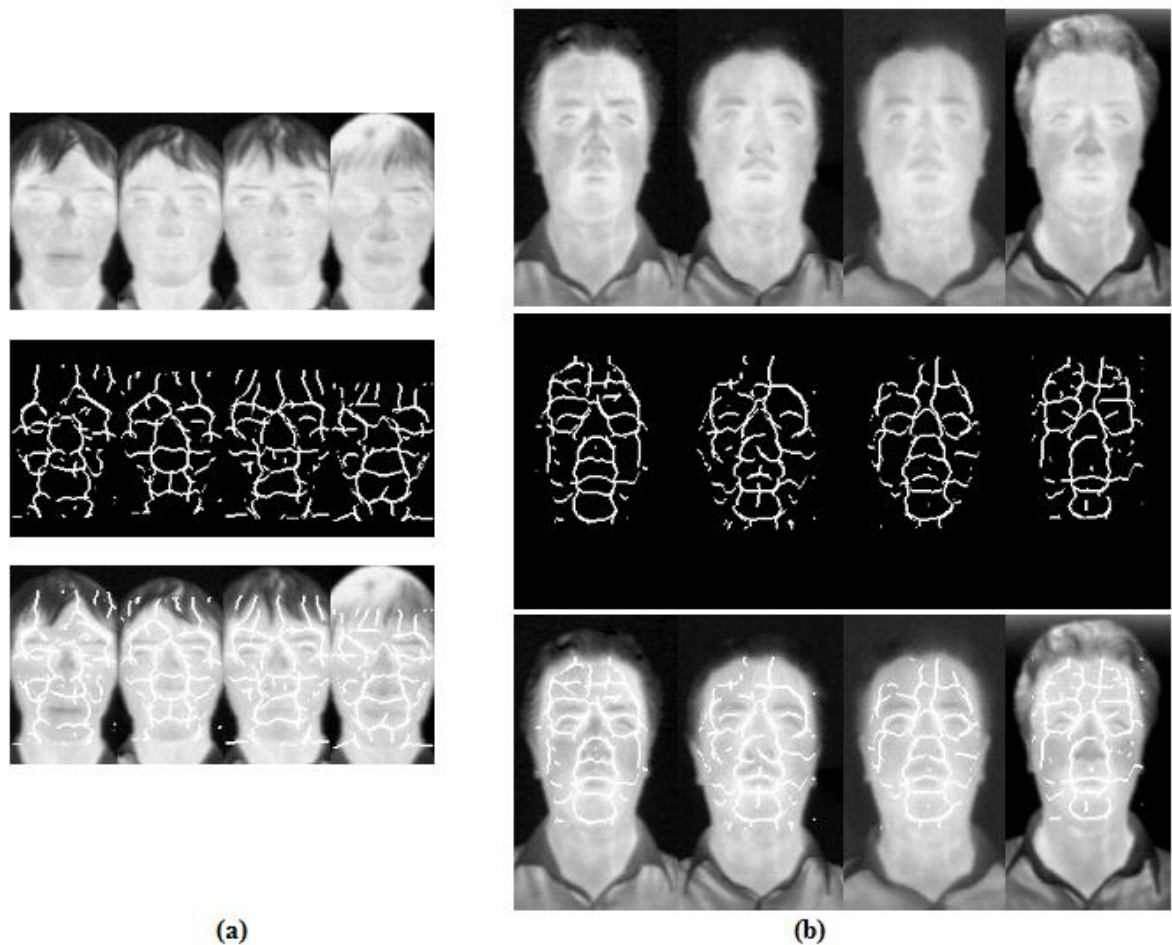


Figura 78. Ejemplos de imágenes de la base de datos UCHThermalFaceIndoorOutdoor. (a) Tamaño de la ventana 81x150 píxeles. De izquierda a derecha: Imagen de interior I1, Imagen de interior I2, imagen de exterior O1, imagen de exterior O2. (b) Tamaño de la ventana 125x225 píxeles. De izquierda a derecha: Imagen de interior I1, Imagen de interior I2, imagen de exterior O1, imagen de exterior O2. Se observan los 3 enfoques usados: rostro térmico, *VascularNetwork* y *VascularThermalFace*.

5.3.3 Resultados de reconocimiento

El desempeño de los diferentes métodos es evaluado con la tasa de reconocimiento utilizando el criterio del Top- 1. En todos los análisis es usado un conjunto como galería y el resto como test, es decir se escoge un conjunto de galería {I1, I2, O1, O2}, y el resto se ocupa como test. Luego se repite el proceso para todas las combinaciones. No se considera la evaluación del mismo conjunto como galería y como test, debido a que su resultado es 100%.

Las Tablas 22-23-24 muestran el desempeño de los métodos de reconocimiento de rostros utilizando los 3 enfoques de rostro térmico, *VascularNetwork* y *VascularThermalFace* para las diferentes combinaciones {I1/I2; I1/O1; I1/O2; I2/O1; I2/O2; O1/O2} de las sesiones de interior y exterior.

Se puede observar de la Tabla 22 que el desempeño de los métodos de reconocimiento al utilizar el rostro térmico es mejor para casos de interior-interior o exterior-exterior que cuando se usan casos de interior-exterior y viceversa. Sin embargo, los métodos basados en calces locales, LBP, WLD y GJD-BC, obtienen buenos resultados para experimentos de interior-interior, exterior-exterior, interior-exterior y viceversa. Esto es debido a que ante variaciones en la imagen térmica producto del lapso temporal en la adquisición de éstas, trae como consecuencia que los métodos basados en calces globales como SIFT y SURF obtengan tasas más bajas de reconocimiento producto del cambio en la visualización del rostro térmico (diferentes representación en contraste para imágenes del mismo rostro de un sujeto), lo que se ve reflejado en este experimento para los casos de interior versus exterior y viceversa. Se destaca el método LBP-HI-80 debido a sus buenos resultados de reconocimiento para todas las combinaciones de interior y exterior, seguido de WLD-HI-80.

De la Tabla 23 se puede observar el desempeño de los métodos cuando se utiliza un enfoque basado en la red vascular. Se visualiza que los métodos disminuyen su rendimiento en comparación con el caso de rostro térmico (Tabla 22). Solamente los métodos LBP y WLD obtienen resultados sobresalientes.

La Tabla 24 muestra el desempeño de los métodos de reconocimiento de rostros cuando se utiliza la metodología propuesta *VascularThermalFace*. Se observa una mejora considerable en las tasas de reconocimiento cuando se encuentra en condiciones de interior versus interior, exterior versus exterior y casos de interior-exterior y viceversa. Para este experimento todos los métodos obtienen buenos resultados de reconocimiento, inclusive es posible ver el aumento en los métodos basados en calces globales como SIFT y SURF cuando se utiliza el enfoque *VascularThermalFace*. Al utilizar esta metodología, se incorpora la información de la red vascular al rostro térmico, la cual es usada como una característica biométrica que permite mejorar el desempeño de los métodos ante captura de imágenes térmicas de interior y exterior. Los métodos que presentan mejores resultados son LBP, las variantes WLD, SURF y GJD-BC.

En la Tabla 25 se observa la tasa de reconocimiento promedio obtenida por cada método y su desviación estándar. Para cada conjunto de galería versus test, se calculó la tasa de reconocimiento obtenida por cada método de reconocimiento, luego se obtuvo el promedio de la tasa de reconocimiento para las sesiones de interior versus exterior y viceversa, con el fin de obtener una métrica que nos indique el desempeño de los métodos de reconocimiento en las sesiones de interior y exterior, debido que poseen diferencia temporal. La idea de obtener la media total de reconocimiento es analizar la robustez de cada método ante cambios de la captura de la imagen térmica en sesiones de interior y exterior.

Tabla 22. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor. Mejores variantes de los métodos. Tasa de reconocimiento para las siguientes combinaciones {I1/I2; I1/O1; I1/O2; I2/O1; I2/O2; O1/O2}. Enfoque: Rostro térmico.

Métodos	Rostro Térmico					
	$\overline{I1I2}$ [%]	$\overline{I1O1}$ [%]	$\overline{I1O2}$ [%]	$\overline{I2O1}$ [%]	$\overline{I2O2}$ [%]	$\overline{O1O2}$ [%]
LBP-HI-80 81x150	97.5	67.5	70.0	77.5	92.5	85.0
GJD-BC 125x225	87.5	70.0	65.0	77.5	80.0	85.0
WLD-EU-80 81x150	87.5	67.5	75.0	85.0	85.0	90.0
WLD-HI-80 81x150	90.0	72.5	67.5	82.5	82.5	85.0
SIFT-L&R-M 125x225	65.0	47.5	22.5	52.5	47.5	47.5
SIFT-Lowe-M 125x225	60.0	45.0	27.5	70.0	37.5	60.0
SURF-M 125x225	75.0	65.0	42.5	65.0	67.5	82.5

Tabla 23. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor. Mejores variantes de los métodos. Tasa de reconocimiento para las siguientes combinaciones {I1/I2; I1/O1; I1/O2; I2/O1; I2/O2; O1/O2}. Enfoque: *VascularNetwork*.

Métodos	<i>VascularNetwork</i>					
	$\overline{I1I2}$ [%]	$\overline{I1O1}$ [%]	$\overline{I1O2}$ [%]	$\overline{I2O1}$ [%]	$\overline{I2O2}$ [%]	$\overline{O1O2}$ [%]
LBP-HI-80 81x150	72.5	55.0	67.5	72.5	70.0	82.5
GJD-BC 125x225	35.0	25.0	30.0	17.5	25.0	45.0
WLD-EU-80 81x150	67.5	52.5	70.0	67.5	67.5	77.5
WLD-HI-80 81x150	77.5	60.0	67.5	70.0	75.0	77.5
SIFT-L&R-M 125x225	45.0	12.5	25.0	25.0	15.0	27.5
SIFT-Lowe-M 125x225	42.5	35.0	27.5	35.0	20.0	32.5
SURF-M 125x225	70.0	50.0	35.0	62.5	57.5	62.5

Tabla 24. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor. Mejores variantes de los métodos. Tasa de reconocimiento para las siguientes combinaciones {I1/I2; I1/O1; I1/O2; I2/O1; I2/O2; O1/O2}. Enfoque: *VascularThermalFace*.

Métodos	<i>VascularThermalFace</i>					
	$\overline{I1I2}$ [%]	$\overline{I1O1}$ [%]	$\overline{I1O2}$ [%]	$\overline{I2O1}$ [%]	$\overline{I2O2}$ [%]	$\overline{O1O2}$ [%]
LBP-HI-80 81x150	97.5	70.0	75.0	77.5	87.5	87.5
GJD-BC 125x225	92.5	65.0	75.0	72.5	75.0	82.5
WLD-EU-80 81x150	92.5	65.0	60.0	77.5	82.5	80.0
WLD-HI-80 81x150	95.0	67.5	67.5	75.0	82.5	80.0
SIFT-L&R-M 125x225	67.5	52.5	47.5	72.5	50.0	70.0
SIFT-Lowe-M 125x225	80.0	57.5	50.0	72.5	57.5	77.5
SURF-M 125x225	87.5	70.0	65.0	82.5	77.5	80.0

Tabla 25. Resultados de los experimentos de galería de interior versus exterior y viceversa, base de datos UCHThermalFaceIndoorOutdoor. Mejores variantes de los métodos. Tasa de reconocimiento (\bar{X} [%]) y desviación estándar (σ) para 3 enfoques: Rostro térmico, *VascularNetwork* y *VascularThermalFace*.

Métodos	Rostro Térmico		<i>VascularNetwork</i>		<i>VascularThermalFace</i>	
	\bar{X} [%]	σ	\bar{X} [%]	σ	\bar{X} [%]	σ
LBP-HI-80 81x150	81.7	12.5	70.0	10.4	82.5	10.8
GJD-BC 125x225	77.5	9.4	29.6	10.8	77.1	9.4
WLD-EU-80 81x150	81.7	10.1	67.1	8.6	76.3	13.8
WLD-HI-80 81x150	80.0	10.2	71.3	8.0	77.9	12.5
SIFT-L&R-M 125x225	47.1	13.8	25.0	11.5	60.0	11.2
SIFT-Lowe-M 125x225	50.0	16.0	32.1	7.7	65.8	12.4
SURF-M 125x225	66.3	13.4	56.3	12.3	77.1	8.3

Se puede concluir de la Tabla 25 que el experimento de galerías de interior versus exterior, con variabilidad temporal, mejora los resultados obtenidos en los Capítulos 3 y 4.

Cuando se utiliza el rostro térmico, los métodos basados en calces locales como LBP-HI-80, WLD-X-80, GJD-BC, son los métodos que obtienen los mejores resultados para este análisis, concordando con lo obtenido con los resultados de la base de datos UCHThermalTemporalFace. Las tasas de reconocimiento de estos métodos de calces locales para este experimento son del orden de ~80, aumentando considerablemente los resultados vistos para experimentos similares de galerías de interior versus exterior del Capítulo 3 y 4. Como se observa de la Tabla 25, los

métodos basados en calces globales como las variantes de SIFT y el método SURF, obtienen pobre desempeño, lo cual es debido al cambio de los niveles de contraste en las imágenes térmicas adquiridas en exterior, como puede ser observado en la Figura 79. Estos cambios de contraste son generados por varias razones como: la exposición de la cámara térmica a los rayos del sol, los cambios anatómicos que sufren los individuos, que son perturbaciones no controlables para estos sistemas.

Al observar los resultados obtenidos por la metodología *VascularNetwork*, se aprecia que también obtiene un comportamiento similar a los resultados obtenidos para la base *UCHThermalTemporalFace*, donde se puede concluir que los métodos LBP y WLD obtienen buenos resultados utilizando esta información, seguido del método SURF. Como se explico en el Capítulo 4, los demás métodos no operan adecuadamente antes las imágenes de red vascular debido a que no están hechos para trabajar con esta información, como el caso de GJD-BC, y no es adecuado para SIFT debido a que los descriptores generados por este tipo de imágenes obtienen muchos calces falsos. En promedio los resultados para este enfoque con los métodos LBP y WLD es ~70% en el desempeño de reconocimiento.

Al usar la metodología propuesta *VascularThermalFace*, se obtienen desempeños similares a los obtenidos cuando se usa el rostro térmico, mejorando el caso de LBP, las variantes de SIFT y SURF. Coincidiendo con los resultados obtenidos de la base *UCHThermalTemporalFace*, la metodología propuesta mejora los métodos basados en calces globales como SIFT y SURF en varios puntos porcentuales, debido a la información que agrega de la red vascular al rostro térmico, para el caso de SURF aumenta en un ~13% y las variantes de SIFT en ~25% y ~30%, para SIFT-L&R y SIFT-Lowe, respectivamente .

En resumen, para cada uno de los métodos, el mejor resultado de las tasas de reconocimiento promedio obtenido para cada enfoque de imágenes: rostro térmico, *VascularNetwork* y *VascularThermalFace*, considerando los diferentes parametros y tamaños de ventana usados (obtenidos del Capítulo 3), es dado por las siguientes variantes:

- LBP-HI-80 81x150 pixeles, *VascularThermalFace*: 82.5%
- WLD-EU-80 81x150 pixeles, Rostro Térmico: 81.7%
- WLD-HI-80 81x150 pixeles, Rostro Térmico: 80.0%
- GJD-BC, 125x225 pixeles, Rostro Térmico: 77.5%
- SURF-M 125x225 pixeles, *VascularThermalFace*: 77.1%
- SIFT-Lowe-M 125x225 pixeles, *VascularThermalFace*: 65.4%
- SIFT-L&R-M 125x225 pixeles, *VascularThermalFace*: 60.4%

5.3.4 Discusión

Como resultado principal de este apartado, se observa que el uso de imágenes térmicas obtenidas de sesiones de interior y exterior, pueden ser correctamente reconocidas por los

métodos de reconocimiento de rostros usados en este trabajo, obteniendo tasas de sobre ~80% de reconocimiento de rostros y mejorando los resultados de reconocimiento alcanzadas para los experimentos de galerías de interior versus exterior realizados en los Capítulos 3 y 4. Además se observó que es posible alcanzar buenos resultados de reconocimiento utilizando la metodología propuesta *VascularThermalFace*.

Tal como se comentó en la Sección 5.2.4, para la base de datos UCHThermalTemporalFace, los métodos basados en calces locales obtienen buenos resultados cuando se realiza el reconocimiento de rostros con imágenes obtenidas en diferentes sesiones de interior y exterior, para los casos interior-interior, exterior-exterior y los casos interior-exterior y viceversa, donde existe una variabilidad temporal en la adquisición de los datos. Para la base de datos UCHThermalFaceIndoorOutdoor el desempeño de los métodos basados en calces locales, métodos LBP, WLD y GJD-BC, son adecuados para un reconocimiento de rostros que posean imágenes capturadas en distintas sesiones de interior y exterior. En cambio, métodos basados en calces globales como SIFT o SURF, poseen bajo desempeño utilizando imágenes térmicas que han sido adquiridas con diferencia temporal o en sesiones no controladas, como también en sesiones de interior-exterior o viceversa, debido a que las cámaras térmicas cambian la información adquirida, producto del calor corporal, la temperatura ambiente, factores que perturban la captura de la cámara térmica y hacen que la imagen posea diferente representación en sus píxeles. Esto nos hace pensar que estos métodos de reconocimiento de rostros, no son adecuados para realizar un sistema de reconocimiento de rostros que sea robusto debido a que no son capaces de reconocer rostros térmicos bajo las circunstancias de las sesiones de interior y exterior. En el Capítulo 3, se concluye que SIFT, es un excelente método para realizar reconocimiento de rostros, pero este análisis es sólo cuando se obtiene una representación térmica que es estable y no varía en el tiempo. Como es posible de visualizar al variar las imágenes en el tiempo, las variantes de SIFT no obtienen buenos resultados.

Al utilizar la metodología propuesta *VascularThermalFace*, se obtienen buenos resultados de reconocimiento para el caso de las sesiones interior y exterior, donde es posible de observar en la Tabla 24 que mejora el desempeño de los métodos de reconocimiento para los casos de interior-exterior y viceversa. Además esta metodología mejora en varios puntos porcentuales el desempeño de algunos algoritmos de reconocimiento como SIFT y SURF, debido a que incorpora información de la red vascular en conjunto con la información del rostro, la cual sirve como ayuda a métodos basados en calces globales para realizar el reconocimiento de rostros en el tiempo.

Se concluye además que sólo algunos métodos son robustos a imágenes adquiridas bajo diferentes sesiones de tiempo, en interior y exterior. Los métodos más robustos corresponden a las variantes de WLD-X-X, LBP-X-X y GJD-BC. Recordar que para esta base de datos se usó una sesión de exterior con mucha saturación en las imágenes, la cual hace que el desempeño de los métodos de reconocimiento disminuya. Sin embargo, estos métodos obtienen buenos resultados cuando se usa el enfoque de rostro térmico y la metodología propuesta *VascularThermalFace*, que valida su uso, sobre todo en casos de sesiones de interior-exterior y viceversa.

6 Conclusiones

En el presente trabajo de tesis se ha estudiado y analizado los sistemas de reconocimiento de rostros que utilizan imágenes térmicas, debido a que permiten solucionar las limitaciones que poseen las imágenes del espectro visible tales como la dependencia a las condiciones de iluminación y a la variación de expresiones faciales. Se realizaron 3 estudios de reconocimiento de rostros térmico: (i) un estudio comparativo de métodos de reconocimiento de rostros en ambientes no controlados, (ii) se presentó una metodología llamada *VascularThermalFace* para el reconocimiento de rostros que usa la información de la red vascular en conjunto con la información de la intensidad de los píxeles del rostro humano, y (iii) se realizó un estudio de reconocimiento de rostros con imágenes térmicas adquiridas en diferentes periodos de tiempo, donde además se analizó el reconocimiento de rostros bajo diferentes condiciones ambientales, utilizando imágenes capturadas en distintas sesiones de interior y en exterior durante varios días. Para realizar este estudio se crearon 3 bases de datos de imágenes de rostros térmicas llamadas UCHThermalFace, UCHThermalTemporalFace y UCHThermalFaceIndoorOutdoor.

En primer lugar, se presentó un estudio comparativo de sistemas de reconocimiento de rostros en el espectro térmico en ambientes controlados y no controlados. Se analizaron y seleccionaron métodos de reconocimiento de rostros considerando sus resultados de reconocimiento en base a requerimientos de operación, tales como: operación en tiempo real, sólo una imagen por persona, operación en línea (sin entrenamiento), un comportamiento robusto en condiciones de ambientes no controlados, y además de su desempeño otros estudios. El estudio comparativo fue realizado usando dos bases de datos: Equinox y UCHThermalFace, donde Equinox -una base de datos estándar de imágenes térmicas- fue usada como línea de comparación de los métodos de reconocimiento de rostros, y la base UCHThermalFace -creada en la Universidad de Chile para estudios de reconocimiento de rostros con imágenes térmicas- fue utilizada para la selección de los métodos de reconocimiento debido a que incorpora aspectos reales como rotaciones con distintos ángulos de orientación e inclinación, variaciones en las condiciones ambientales, sesiones de exterior y diferentes expresiones faciales. Los métodos seleccionados para realizar el estudio comparativo fueron: histogramas de LBP, descriptores jets de Gabor, descriptores SIFT, histogramas de WLD y SURF. Los experimentos realizados para analizar el estudio de reconocimiento corresponden a: estudios del tamaño de la ventana de la imagen, diferente alineamiento, oclusiones, diferentes expresiones faciales y distintas condiciones ambientales como captura de los individuos en sesiones de interior (en una pieza controlada), como en sesiones de exterior o al aire libre (ambientes no controlados).

Los resultados más importantes obtenidos de este estudio comparativo fueron: (i) la dependencia del tamaño de la ventana no es grande para los mejores métodos seleccionados (histogramas de WLD y SIFT), (ii) sólo algunos de los métodos analizados son robustos a alineamientos imprecisos, oclusión de rostros y a variaciones en la expresión, SIFT es el método más robusto a errores de alineación, a rotaciones y a cambios de expresiones faciales, y GJD-BC

el menos robusto ante estos factores, (iii) los resultados obtenidos de reconocimiento para los distintos métodos son bajos cuando se capturan imágenes térmicas en diferentes sesiones de interior y exterior, lo cual proviene de la saturación de los píxeles de las imágenes térmicas obtenidas en ambientes exteriores producto de la cámara térmica, y (iv) los métodos histogramas de LBP y WLD, son los métodos más rápidos para realizar el reconocimiento de rostros debido a que poseen el menor tiempo de procesamiento. Como conclusión del estudio comparativo, se seleccionan los métodos basados en histogramas de WLD debido a que poseen una operación en tiempo real, y un buen desempeño en los resultados de reconocimiento. Si el tiempo de procesamiento no fuera una consideración importante, las variantes de SIFT sería la elección a usar debido a la robustez a errores de alineación, rotaciones y cambios de expresiones faciales.

En segundo lugar, se presentó también una nueva metodología llamada *VascularThermalFace* que combina las características vasculares del rostro humano y la información del rostro térmico. Esta nueva metodología *VascularThermalFace* utiliza un algoritmo de extracción de la red vascular del rostro térmico llamado *VascularNetwork*, con la cual extrae la información de la red vascular para combinarla con la información del rostro térmico. La metodología propuesta fue evaluada y validada utilizando la base de datos UCHThermalFace en conjunto con los métodos seleccionados del estudio comparativo. Para validar la metodología se repitieron los experimentos del estudio comparativo anteriormente descrito, donde se analizó el desempeño de los métodos seleccionados utilizando la metodología *VascularNetwork* y la metodología propuesta *VascularThermalFace*. Los resultados más importantes obtenidos son: (i) no existen mejoras de los resultados para los métodos de reconocimiento de rostros al utilizar la metodología *VascularNetwork*, (ii) con la metodología propuesta *VascularThermalFace* se obtienen mejoras de los resultados obtenidos por los métodos GJD-BC, WLD, SURF y LBP, inclusive se mejora la robustez de los métodos ante malas alineaciones, rotaciones y expresiones faciales. Las variantes de SIFT no se mejoran con la metodología propuesta para este experimento.

En tercer lugar, se presentó además un estudio de reconocimiento de rostros utilizando imágenes térmicas adquiridas bajo distintas sesiones de tiempo, (un total de 69 días). Con estas imágenes térmicas se creó la base de datos UCHThermalTemporalFace. El estudio analizó el desempeño de los métodos de reconocimiento cuando existen cambios corporales, ambientales, variaciones en la pose, oclusión y cambios estéticos de los individuos durante el lapso de tiempo en la adquisición de las imágenes. Otro de los puntos importantes que cabe destacar, es que el análisis desarrollado en el trabajo, tiene por finalidad validar la metodología propuesta *VascularThermalFace*, observando el desempeño obtenido por los métodos de reconocimiento de rostros seleccionados. Como resultados de este estudio se obtiene que: (i) los métodos de reconocimiento basados en calces locales como GJD-BC, WLD y LBP, obtienen buenos resultados de reconocimiento de rostros en el tiempo, sin embargo para el caso de SIFT y SURF el desempeño es más bien bajo, cuando existe mucha variabilidad temporal, (ii) sólo algunos de los métodos estudiados son robustos a oclusión y al ruido en la posición de los ojos, como se observó en el estudio comparativo, WLD el más robusto ante ambas perturbaciones, en cambio los métodos SIFT y SURF son más robustos a alineamientos y ruido en la posición de los ojos, mientras que los descriptores jets de Gabor es el método menos robusto a factores no deseados, y (iii) se valida el uso de la metodología propuesta debido a que entrega excelentes resultados en la tasa de reconocimiento, además de mejorar el desempeño de los demás métodos de reconocimiento como SIFT y SURF, aumentando la robustez de los métodos de reconocimiento

cuando se encuentran en presencia de ruido en la posición de los ojos (mal alineamiento) y oclusión.

Debido al bajo desempeño obtenido por los métodos de reconocimiento en el estudio comparativo cuando se obtenían imágenes térmicas en diferentes sesiones de captura, en interior y exterior, es que se analizó este fenómeno a través de una nueva base de datos llamada UCHThermalFaceIndoorOutdoor. Se estudió y analizó como influye la configuración de la cámara térmica para realizar la captura de los datos y además se seleccionó la mejor configuración de parámetros para poder realizar un reconocimiento robusto de rostros. El estudio fue desarrollado por medio de los métodos de reconocimiento seleccionados en el estudio comparativo y se utilizaron 3 representaciones para las imágenes térmicas: el rostro térmico, la red vascular *VascularThermalFace* y la metodología propuesta *VascularThermalFace*. Como resultados de este estudio se observó que: (i) se obtienen resultados de sobre el ~80% de reconocimiento, mejorando los resultados obtenidos en el estudio comparativo, (ii) los métodos basados en calces locales como WLD, LBP y GJD-BC operan de manera robusta para el reconocimiento en el tiempo, los métodos SIFT y SURF obtienen pobre desempeño, y (iii) se valida el uso de la metodología *VascularThermalFace*, debido a los buenos resultados obtenidos por los métodos de reconocimiento de rostros en el tiempo cuando se usa esta metodología, y además la metodología *VascularThermalFace* mejora los resultados de los métodos SIFT y SURF.

En base a lo realizado en la tesis y las conclusiones obtenidas en cada capítulo, de los 3 estudios de reconocimiento de rostros que fueron realizados en esta tesis, se obtienen como conclusiones finales:

- (i) El método a seleccionar para realizar un reconocimiento de rostros que sea robusto y que sea invariante a expresiones faciales, oclusión, iluminación, a cambios de aspecto, que opere en tiempo real, con tiempos pequeños de procesamiento y que no sea afectado por el lapso temporal en la adquisición de las imágenes de rostro térmico es el método histogramas de WLD, debido a que posee el mejor “*trade off*” entre tiempo de procesamiento y resultados de reconocimiento. Los métodos de histogramas de LBP y GJD-BC, operan con buenos resultados de reconocimiento y son robustos ante variaciones temporales pero se ven más afectados ante factores adversos como oclusión y mal alineamiento. Los métodos SIFT y SURF poseen mejores aspectos de robustez que el método WLD para el caso de ruido en la posición de los ojos, pero son más lentos y no obtienen un desempeño alto en el reconocimiento de rostros cuando existe una variabilidad temporal considerable. Además, dependiendo de la configuración de se desee optar entre velocidad y desempeño, se pueden seleccionar combinaciones de los métodos de reconocimiento y la metodología propuesta *VascularThermalFace*, debido a que ésta aumenta la robustez de los métodos de reconocimiento en algunos casos de oclusión y mala alineación.
- (ii) Se valida el uso de la metodología propuesta *VascularThermalFace*. Esta nueva representación del rostro permite obtener tasas altas de reconocimiento, aumentando la robustez de los métodos de reconocimiento de rostros. Se demuestra el hecho que *VascularThermalFace* permite obtener mejoras considerables en el desempeño de los métodos seleccionados en base a los resultados obtenidos de los experimentos de

reconocimiento de rostros que usan las bases de datos UCHThermalFace, UCHThermalTemporalFace y UCHThermalFaceIndoorOutdoor. Se observó que la metodología propuesta eleva los resultados obtenidos para la base UCHThermalFace y genera tasas altas de reconocimiento para las demás bases de datos cuando se combina con un método de reconocimiento de rostros, como por ejemplo: GJD-BC + *VascularThermalFace* obtiene un 96.6% en el desempeño de reconocimiento para la base de datos UCHThermalTemporalFace.

- (iii) Es posible realizar el reconocimiento de rostros en el tiempo, con una galería de imágenes adquirida en diferente periodo que las imágenes de test, ya que depende solamente de la elección del método de reconocimiento y de una buena configuración de la cámara térmica. Se observó un buen desempeño de los métodos de reconocimiento ante las variaciones temporales obtenidas de 69 días de captura de las imágenes térmicas. Se observó también que la metodología propuesta *VascularThermalFace* mejora la estabilidad de los métodos de reconocimiento, cuando existe un mal alineamiento y oclusión, aumentando considerablemente el desempeño de los métodos de reconocimiento, en especial para las variantes de SIFT y SURF. Cuando es aplicado un mal alineamiento del rostro en conjunto con una oclusión parcial de alguna zona de la cara, la mayoría de los métodos tienden a disminuir su desempeño de reconocimiento, sin embargo el método WLD es el más robusto ante estas 2 perturbaciones aplicadas conjuntamente.
- (iv) Se muestra que la metodología *VascularNetwork*, puede ser usada como una característica del rostro, debido a que se mantiene en el tiempo. Sin embargo, los resultados de reconocimiento obtenidos por esta metodología, no son adecuados para los métodos seleccionados, debido a que se obtienen tasas de reconocimiento que no superan el desempeño de los métodos de reconocimiento cuando se usa el rostro térmico o la metodología propuesta *VascularThermalFace*. Como se observó, algunos de los métodos no eran aptos para realizar el reconocimiento de rostros con este tipo de características y otros sí.
- (v) Al realizar el reconocimiento de rostros con imágenes obtenidas en múltiples sesiones de interior y exterior, el resultado de reconocimiento es bueno para algunos métodos de reconocimiento (los métodos basados en calces locales) y para otros métodos es bastante bajo (métodos basados en calces globales). Se observó que cuando existen diferentes sesiones de interior y exterior, el desempeño de los métodos de reconocimiento mejora considerablemente al utilizar el enfoque de *VascularThermalFace*, sobre todo para el caso del método WLD, que obtiene la mejor tasa de reconocimiento promedio. Este resultado, mejora considerablemente el desempeño obtenido en el primer experimento de galerías de interior y exterior del estudio comparativo (Capítulo 3), lo que demuestra lo analizado respecto a que la configuración de la cámara térmica influye mucho en la adquisición de las imágenes térmicas. Este es un tema importante que debe ser considerado en los análisis de reconocimiento de rostros debido a que los factores propios y ambientales de la cámara térmica limitan el desempeño de los métodos si no se encuentra configurada adecuadamente.

Trabajo futuro

Se cree que aún falta mejorar en muchos aspectos el reconocimiento de rostros térmicos en ambientes no controlados y el reconocimiento de rostros con un lapso de tiempo prolongado en la adquisición de las imágenes térmicas (variación temporal). Las principales preguntas en el caso térmico son: (i) ¿cómo alcanzar invariancia a condiciones ambientales tales como temperatura y saturación de las imágenes debido al calentamiento de la cámara producido por largos periodos de operación o bajo diferentes condiciones ambientales? (ejemplo operación en exterior bajo directa exposición al sol), (ii) ¿cómo combinar el uso de los diferentes métodos para alcanzar, al mismo tiempo, altas tasas de reconocimiento y velocidad de procesamiento?, (iii) ¿cuál es la influencia de la resolución de la cara en los procesos de reconocimiento de rostros térmico?, (iv) ¿qué puede ser aprendido a través de un análisis profundo del efecto de la expresión facial en el reconocimiento de rostros?, (v) ¿cómo se puede desarrollar un método que sea robusto a rotaciones? (vi) ¿es posible seleccionar sólo características importantes que se mantengan en el tiempo para la representación de la red vascular?, y (vii) ¿cómo mezclar información vascular en una fusión de datos térmicos y visibles? Además se quiere extender este estudio comparativo aplicando nuevos métodos de clasificadores de textura ([75][76]) para reconocimiento de rostros usando imágenes visibles y térmicas, como también explorar el uso de enfoques basados en reconocimiento multimodal, los que utilizan la fusión de la información visible y térmica para desarrollar el reconocimiento ([77]).

7 Referencias

- [1] D. Socolinsky and A. Selinger, "A comparative analysis of face recognition performance with visible and thermal infrared imagery", in *Proc. of the Int. Conf. on Pattern Recognition (ICPR)*, Quebec, Canada, 2002.
- [2] A. Selinger and D. Socolinsky, "Appearance-Based Facial Recognition Using Visible and Thermal Imagery: A Comparative Study", *Tech. Rep., Equinox Corporation*, 2001.
- [3] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition", *IEEE Trans. on Patt. Analysis and Machine Intell.*, vol. 28, no.12, pp. 2037-2041, 2006.
- [4] J. Chen et al., "WLD: A Robust Local Image Descriptor", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no.9, pp. 1705-1720, 2012.
- [5] J. Zou, Q. Ji, and G. Nagy, "A Comparative Study of Local Matching Approach for Face Recognition", *IEEE Trans. on Image Processing*, vol. 16, no. 10, pp. 2617-2628, 2007.
- [6] D. Lowe, "2004. Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [7] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding (CVIU)*, vol. 110, no. 3, pp. 346-359, 2008.
- [8] P. Buddharaju, I. Pavlidis, and I. Kakadiaris, "Physiology-Based Face Recognition", *In Proc. of the IEEE Conf. on Advanced Video and Signal Based Surveillance, Lake Como, Italy*, pp. 354-359, 2005.
- [9] P. Buddharaju, I. Pavlidis, and I. Kakadiaris, "Pose-Invariant Physiological Face Recognition in the Thermal Infrared Spectrum", *In Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, New York, USA*, pp. 53-60, 2006.
- [10] D. Socolinsky, L. Wolff, J. Neuheisel, and C. Eveland, "Illumination invariant face recognition using thermal infrared imagery", in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2001.
- [11] O. Arandjelovic, R. Hammoud, and R. Cipolla, "Towards Person Authentication by Fusing Visual and Thermal Face Biometrics", *Face Biometrics for Personal Identification, Springer*, pp. 75-90, 2007.

- [12] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi, “Recent advances in visual and infrared face recognition - A review”, *Computer Vision and Image Understanding* 97(1), pp. 103-135, 2005.
- [13] D. Socolinsky and A. Selinger, “Thermal Face Recognition Over Time”, *Proc. of the, 17th Int. Conf. on Pattern Recognition (ICPR '04)*, vol. 4, pp. 187-190, 2004.
- [14] X. Chen, P. Flynn, and K. Bowyer, “IR and visible light face recognition”, *Computer Vision and Image Understanding*, vol. 99 n.3, pp. 332-358, 2005.
- [15] G. Hermosilla, J. Ruiz-del-Solar, R. Verschae, and M. Correa, “A Comparative Study of Thermal Face Recognition Methods in Unconstrained Environments”, *Pattern Recognition*, vol. 45, no. 7, pp. 2445-2459, 2012.
- [16] G. Hermosilla, J. Ruiz-del-Solar, R. Verschae, and M. Correa, “Face Recognition using Thermal Infrared Images for Human-Robot Interaction Applications: A Comparative Study”, *In the 6th IEEE Latin American Robotics Symposium – LARS 2009, Oct. 29 - 30, Valparaíso, Chile (CD Proceedings)*., 2009.
- [17] G. Hermosilla, P. Loncomilla, and J. Ruiz-del-Solar, “Thermal Face Recognition using Local Interest Points and Descriptors for HRI Applications”, *Lecture Notes in Computer Science 6556 (RoboCup Symposium 2010)*, pp. 25-35, 2010.
- [18] J. Ruiz-del-Solar, P. Loncomilla, and C. Devia, “A New Approach for Fingerprint Verification based on Wide Baseline Matching using Local Interest Points and Descriptors”, *Lecture Notes in Computer Science 4872 (PSIVT 2007)*, pp. 586-599, 2007.
- [19] M. Correa, G. Hermosilla, R. Verschae, and J. Ruiz-del-Solar, “Human Detection and Identification by Robot using Thermal and Visual Information in Domestic Environments”, *Journal of Intelligent and Robotic Systems*, 66, 1-2 (April 2012), 223-243. <http://dx.doi.org/10.1007/s10846-011-9612-2>., 2012.
- [20] N. Diakedes and J. Bronzino, “Medical Infrared Imaging”, *Chapter 2, Francis E. Ringand, Bryan F. Jones, The Historical Development of Thermometry and Thermal Imaging in Medicine*, 2006.
- [21] H. Qi and N. Diakedes, “Infrared Imaging in Medicine”, *Article, Book*, vol. 3, pp. 108-113, 1998.
- [22] W. Wolfe, “Introduction to Infrared System Design”, *SPIE Optical Engineering Press, Bellingham*, 1996.
- [23] C. Conaire, N. O’Connor, and A. Smeaton, “Thermo-visual feature fusion for object tracking using multiple spatiogram trackers”, *Machine Vision and Applications, Springer-Verlag*, pp. 1-12, 2007.
- [24] B. Jones and P. Plassmann, “Digital Infrared Thermal Imaging of Human Skin”, *Engineering in Medicine and Biology Magazine, IEEE*, pp. 41-48, 2002.
- [25] P. Kruse and D. Skatrud, “Uncooled infrared imaging arrays and systems in Semiconductors and Semimetals”, *R.K. Willardson, E.R. Weber, Academic Press, New York*, 1997.

- [26] D. Scribner, M. Kruer, and J. Killiany, “Infrared focal plane array technology”, *Proceeding of the IEEE*, vol. 79, no.1, pp. 66-85, 1991.
- [27] G. Holst, “CCD Arrays, Cameras and Displays”, *SPIE Optical Engineering Press, Bellingham*, 1996.
- [28] “Focal Plane Array”, [Online]. Available: http://www.dipity.com/LL_timeline/MIT-Lincoln-Laboratory-60th-Anniversary-Timeline/. [Accessed: 01-Mar-2012].
- [29] “Sistema de captura de un IRFPA”, [Online]. Available: <http://nuc.die.udec.cl/publications/slideshows/2003-PosterSeminarioDIE.pdf>. [Accessed: 01-Mar-2012].
- [30] A. Milton, F. Barone, and M. Kruer, “Influence of nonuniformity on infrared focal plane array performance”, *Optical Engineering*, vol. 24, pp. 855-862, 1985.
- [31] J. Mooney, F. Shepherd, W. Ewing, J. Murqui, and J. Silverman, “Responsivity nonuniformity limited performance of infrared starting cameras”, *Optical Engineering*, vol. 28, pp. 1151-1161, 1989.
- [32] A. El Gamal, “Introduction to Image Sensors and Digital Cameras (Lecture Notes)”, *Stanford University, CA*, 2001.
- [33] R. Reeves, “Compensación de Ruido Espacial en Sistemas de Imágenes con Sensores infrarrojos”, *Memoria de Título, Departamento de Ingeniería Eléctrica, Universidad de Concepción, Chile*, 2002.
- [34] G. Hermosilla, “Estimación indirecta de ganancia para la corrección de la uniformidad de las imágenes infrarrojas”, *Memoria de Título, Departamento de Ingeniería Eléctrica, Universidad de La Frontera, Chile*, 2007.
- [35] A. Friedenber and I. Goldblatt, “Nonuniformity two-point linear correction errors in infrared focal plane arrays”, *Optical Engineering*, vol. 3, pp. 1251-1253, 1998.
- [36] S. Tzimopoulou and A. Lettington, “Scene based techniques for nonuniformity correction of infrared focal plane array”, *Proceedings of the SPIE*, vol. 3436, pp. 172-183, 1998.
- [37] C. Jara, “Evaluación y Comparación de Desempeño de Métodos de Corrección de No-Uniformidad de Imágenes Infrarrojas”, *Memoria de Título, Departamento de Ingeniería Eléctrica, Universidad de Concepción, Chile*, 2003.
- [38] P. Narendra, “Reference-free Nonuniformity Compensation for IR Imaging Arrays”, *Proc. of the SPIE*, vol. 252, pp. 10-17, 1980.
- [39] P. Narendra and N. Foss, “Shutterless Fixed Pattern Noise Correction for Infrared Imaging Arrays”, *Proc. of the SPIE*, vol. 282, pp. 44-51, 1981.

- [40] D. Scribner, K. Sarkady, J. Caulfield, M. Kruer, G. Katz, and C. Gridley, “Nonuniformity correction for starting focal plane arrays using scene techniques”, *Proceedings of the SPIE*, vol. 1308, pp. 224-233, 1990.
- [41] F. Torres, S. Torres, and C. San Martín, “A Recursive Least Square Adaptive Filter for Nonuniformity Correction of Infrared Image Sequences”, *Lecture Notes on Computer Science*, vol. 3773, pp. 540-546, 2005.
- [42] X. Chen, P. Flynn, and K. W. Bowyer, “PCA-Based face recognition in infrared imagery: Baseline and comparative studies”, *IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures, Nice, France*, pp. 127-134, 2003.
- [43] “Cámara Térmica Jade UC”, [Online]. Available: http://www.edevis.de/products/products_ircameras_jade_uc_en.php. [Accessed: 01-Mar-2012].
- [44] “Cámara Térmica FLIR TAU 320”, [Online]. Available: <http://www.flir.com/cvs/cores/uncooled/products/tau>. [Accessed: 01-Mar-2012].
- [45] A. Gyaourova, G. Bebis, and I. Pavlidis, “Fusion of Infrared and Visible Images for Face Recognition”, *8th European Conference on Computer Vision, Prague, Czech Republic*, pp. 456-468, 2004.
- [46] J. Ruiz-del-Solar, R. Verschae, and M. Correa, “Recognition of Faces in Unconstrained Environments: A Comparative Study”, *EURASIP Journal on Advances in Signal Processing, special issue, Recent Advances in Biometric Systems: A Signal Processing Perspective*, vol. 2009, no. Article ID 184617, 19 pages, doi:10.1155/2009/184617, 2009.
- [47] D. Socolinsky and A. Selinger, “Thermal Face Recognition in an Operational Scenario”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA*, pp. 1012-1019, 2004.
- [48] H. Mendez, C. San Martín, J. Kittler, Y. Plasencia, and E. García, “Face Recognition with LWIR Imagery using Local Binary Patterns”, *LNCS*, vol. 5558, pp. 327 – 336, 2009.
- [49] “Sitio web corporación Equinox”, [Online]. Available: http://www.equinoxsensors.com/products/multispec_faces.html. [Accessed: 01-Mar-2012].
- [50] “Base de datos Equinox”, [Online]. Available: <http://www.equinoxsensors.com/products/HID.html>. [Accessed: 01-Mar-2012].
- [51] J. Ahmad, U. Ali, and R. J. Qureshi, “Fusion of Thermal and Visual Images for efficient Face Recognition using Gabor Filter”, *In the 4th ACS/IEEE International Conference on Computer Systems and Applications, March 8-11, Dubai/Sharjah, UAE*, pp. 135 – 139, 2006.
- [52] P. Buddharaju and I. Pavlidis, “Multi-Spectral Face Recognition - Fusion of Visual Imagery with Physiological Information”, *Face Biometrics for Personal Identification: Multi-Sensory Multi-Modal Systems*, R.I. Hammoud, B.R. Abidi, and M.A. Abidi, editors, Springer, January, pp. 91-108, 2007.

- [53] P. Buddharaju, I. Pavlidis, and C. Manohar, "Face Recognition Beyond the Visible Spectrum", *Advances in Biometrics: Sensors, Algorithms and Systems*, Springer, pp. 157-180, 2007.
- [54] M. Akhloufi and A. Bendada, "Thermal Faceprint: A New Thermal Face Signature Extraction for Infrared Face Recognition", In *CRV 2008: Fifth Canadian Conference on Computer and Robot Vision*, (Windsor, Ontario), 28-30 May, pp. 269-272, 2008.
- [55] S. Y. Cho, L. Wang, and W. L. Ong, "Thermal imprint feature analysis for face recognition", *Industrial Electronics 2009 ISIE 2009 IEEE International Symposium*, vol. 2009, pp. 1875-1880.
- [56] G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis, "Face recognition by fusing thermal infrared and visible imagery", *Image Vision Computing*, vol. 24, no. 7, pp. 727-742, 2006.
- [57] S. Desa and S. Hati, "IR and Visible Face Recognition using Fusion of Kernel Based Features", In *the 19th Int. Conf. on Pattern Recognition (ICPR 2008)*, December 8-11. Tampa, Florida, USA, pp. 1-4.
- [58] F. M. Pop, M. Gordan, C. Florea, and A. Vlaicu, "Fusion based approach for thermal and visible face recognition under pose and expresivity variation", *Roedunet Int. Conf. (RoEduNet)*, pp. 61-66, 2010.
- [59] B. Abidi, S. Huq, and M. Abidi, "Fusion of visual, thermal, and range as a solution to illumination and pose restrictions in face recognition", *Proc. of IEEE Carnahan Conf. on Security Technology, Albuquerque, NM, Oct*, pp. 325-330, 2004.
- [60] O. K. Kwon and S. G. Kong, "Multiscale fusion of visual and thermal images for robust face recognition", In *Proc. IEEE Int. Conf. Comp. Intell. for Homeland Security and Personal Safety, FL, Mar.*, vol. IV, pp. 112-116, 2005.
- [61] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu, and M. Kundu, "Optimum fusion of visual and thermal face images for recognition", *Sixth Sixth International Conference on Information Assurance and Security (IAS 2010)*, Atlanta, USA, *IEEE Intelligent Transportation Systems Society*, Aug 23-25, pp. 311-316, 2012.
- [62] M. M. Khan, R. D. Ward, and M. Ingleby, "Infrared Thermal Sensing of Positive and Negative Affective States", In *the Proceedings of the 2006 IEEE Conference on Robotics Automation and Mechatronics*, pp. 1-6, 2006.
- [63] L. Trujillo, G. Olague, R. Hammoud, and B. Hernandez, "Automatic Feature Localization in Thermal Images for Facial Expression Recognition", *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, June 20-26, p. 14, 2005.
- [64] S. Wang, S. Lv, and X. Wang, "Infrared Facial Expression Recognition Using Wavelet Transform", *Int. Symposium on Computer Science and Computational Technology*, vol. 2, pp. 327-330, 2008.

- [65] Y. Yoshitomi, N. Miyawaki, S. Tomita, and S. Kimura, "Facial expression recognition using thermal image processing and neural network", *Proc. IEEE Int. Workshop Robot Hum. Commun*, pp. 380-385, 1997.
- [66] Y. Yoshitomi, K. Sung-Il, T. Kawano, and T. Kilazoe, "Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face", *Proc. of 9th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 178-183, 2000.
- [67] J. Ruiz-del-Solar and J. Quinteros, "Illumination Compensation and Normalization in Eigenspace-based Face Recognition: A comparative study of different pre-processing approaches", *Pattern Recognition Letters*, vol. 29, no. 14, pp. 1966-1979, 2008.
- [68] X. Tan, S. Chen, Z. H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey", *Pattern Recognition*, vol. 39, pp. 1725-1745, 2006.
- [69] "OpenSURF", [Online]. Available: <http://www.chrisevansdev.com/computer-vision-opensurf.html>. [Accessed: 01-Mar-2012].
- [70] R. Gross, "Face Databases", in: *Li, S., Jain, A., (Eds). Chapter 13, Handbook of Face Recognition, Springer-Verlag, 2005.*
- [71] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 511-518, 2001.
- [72] M. Jones and J. Rehg, "Statistical color models with application to skin detection", *International Journal of Computer Vision (IJCV)*, vol. 46, no. 1, pp. 81-96, 2002.
- [73] S. Wu, W. Song, and L. Jiang, "Infrared face recognition by using blood perfusion data", *Infrared face recognition by using blood perfusion data, Rye Brook, NY, USA*, pp. 320-328, 2005.
- [74] A. Bhowmik et al., "Thermal Infrared Face Recognition – A Biometric Identification Technique for Robust Security System", *Reviews, Refinements and New Ideas in Face Recognition*, 2011.
- [75] W. Lan, "Face Recognition System Based on Spatial Constellation Model and Support Vector Machine", *Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Master's Thesis*, 2010.
- [76] H. Lategahn, S. Gross, T. Stehle, and T. Aach, "Texture classification by modeling joint distributions of local patterns with Gaussian mixtures", *IEEE Trans. on Image Processing*, vol. 19, no. 6, pp. 1548-1557, 2010.
- [77] M. He et al., "Performance evaluation of score level fusion in multimodal biometric systems", *Pattern Recognition*, vol. 43, no. 5, pp. 1789-1800, 2010.