



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO
CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN



Registro de imágenes digitales multimodales

T E S I S

QUE PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN

PRESENTA

Ing. Edgar López Jasso

DIRECTOR DE TESIS

Dr. Edgardo Manuel Felipe Riverón

Ciudad de México, enero de 2016



INSTITUTO POLITÉCNICO NACIONAL SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 12:00 horas del día 03 del mes de diciembre de 2015 se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

Centro de Investigación en Computación

para examinar la tesis titulada:

“Registro de imágenes digitales multimodales”

Presentada por el alumno:

LÓPEZ

Apellido paterno

JASSO

Apellido materno

EDGAR

Nombre(s)

Con registro:

B	1	3	0	0	9	5
----------	----------	----------	----------	----------	----------	----------

aspirante de: **MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN**

Después de intercambiar opiniones los miembros de la Comisión manifestaron **APROBAR LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA

Director de Tesis

Dr. Edgardo Manuel Felipe Riverón

Dr. Sergio Suárez Guerra

Dr. Oleksiy Pogrebnyak

Dr. Víctor Hugo Ponce Ponce

M. en C. Germán Téllez Castillo

Dr. Francisco Hiram Calvo Castro

PRESIDENTE DEL COLEGIO DE PROFESORES

Dr. Alfonso Villa Vargas



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA DE INVESTIGACIÓN
EN COMPUTACIÓN
DIRECCIÓN



INSTITUTO POLITÉCNICO NACIONAL
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

CARTA CESIÓN DE DERECHOS

En la Ciudad de **México** el día **9** del mes de **diciembre** del año **2015**, el que suscribe **C. Edgar López Jasso** alumno del Programa de **Maestría en Ciencias de la Computación** con número de registro **B130095**, adscrito al **Centro de Investigación en Computación**, manifiesta que es autor intelectual del presente trabajo de Tesis bajo la dirección del **Dr. Edgardo Manuel Felipe Riverón** y cede los derechos del trabajo intitulado **Registro de imágenes digitales multimodales**, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección **edgar.lopez.isc@gmail.com**. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

A handwritten signature in black ink, appearing to read 'E. López Jasso', is written over a horizontal line.

Edgar López Jasso

Resumen

El presente trabajo propone una metodología de registro de imágenes multimodales que sigue el enfoque de la selección de características invariantes a escala, conocida como SIFT por sus siglas en inglés. La directriz fundamental de este trabajo es lograr el registro de imágenes multimodales, y de manera particular conseguir el registro de imágenes de retina ópticas y fluoresceínicas requeridas para realizar estudios posteriores que, aunque fundamentales, su explicación queda fuera del alcance de este trabajo.

Las imágenes que se usarán para probar el desempeño de la implementación de la metodología propuesta constituye un conjunto de parejas de imágenes de retinas, de las cuales algunas presentan microaneurismas y otras no; cada par se compone de una imagen óptica y una fluoresceínica de una misma retina; estas imágenes se toman en condiciones no controladas, por lo que pueden presentar diferencias apreciables si se comparan a simple vista; de aquí surge la tarea de tratar de tenerlas en un estado similar para intentar registrarlas con el fin de usarlas en trabajos posteriores.

La metodología propuesta se conforma por cinco etapas, con el preprocesamiento como la etapa inicial y principal, ya que permite realizar las siguientes en el caso particular del registro de imágenes de retina. La segunda es la selección de características en la cual se implementa el método SIFT; como tercera etapa tenemos la correspondencia de características, la cual tiene como finalidad obtener la intersección de los conjuntos de características que estén presentes en ambas imágenes; la cuarta etapa busca encontrar la matriz de transformación que permita transformar las imágenes a registrar con base en las características de la imagen de referencia; como quinta y última etapa tenemos la transformación de la imagen a registrar.

Los resultados que se obtuvieron al probar la implementación de la metodología propuesta son los siguientes: de un total de 35 pares de imágenes de retina el registro fue viable para 26 de ellos, lo cual corresponde a un 74.28%, el porcentaje restante no fue viable porque no se pudieron obtener características significativas de ellos; de esos 26 pares registrados, 14 pares que corresponden a un 53.85% eran de imágenes de retinas sin microaneurismas y 46.15 % eran imágenes de retina con microaneurismas.

Abstract

The present work propose a registration methodology of multimodal images, which follow an approach of scale invariant features selection known as SIFT. The fundamental aim of this work is to achieve the multimodal image registration, and particularly achieve the registration of optical retinal images required for further studies, his explanation is beyond the scope of this paper.

The set of images that will be used to test the performance of the implementation of the proposed methodology is a set of pairs of retinal images, some of which have microaneurysms and the rest have not; each pair consists of an optical image and a fluorescein of the same; These images are taken in uncontrolled conditions so they can present significant differences; hence the attempt is to have these images in a similar state to try register them in order to use them in later works.

The proposed methodology is made up of five stages, the initial and main stage is the preprocessing, it allows the next stages; the second is the features selection which is done using SIFT method; the third stage is feature correspondence which aims to obtain the intersection of the features sets that are present in both images; the fourth step is try to find the transformation matrix capable of convert the sensed image; at the fifth and final stage we transform the sensed image.

The results obtained with the implementation of this methodology are: it was possible to register 26 pairs of a total 35 pairs of images of the retina, which corresponds to 74.28% of total. The remaining percentage was feasible because it was not possible to obtained image features. From the registered 26 pairs, 14 pairs, which correspond to 53.85%, were retinal images without microaneurysms, and 46.15% were retinal images with microaneurysms.

Agradecimientos

Agradezco a mi familia por su apoyo incondicional a lo largo de este proceso. En particular te agradezco Magdalena, la comprensión, paciencia y ayuda que me has brindado y que sin duda fueron fundamentales para lograr este objetivo; a ti Gabriel, te agradezco especialmente por ser la fuerza que me impulsa a seguir.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) y al Instituto Politécnico Nacional (IPN) por brindarme la oportunidad, el apoyo y el patrocinio para realizar mis estudios de maestría.

Agradezco a los profesores del Centro de Investigación en Computación (CIC) por sus observaciones, correcciones y por compartir sus conocimientos para la realización de este trabajo de tesis, en especial al Dr. Edgardo Manuel Felipe Riverón por la guía en este trabajo y por los consejos brindados.

Índice

SIP.....	I
Sesión de derechos.....	II
Resumen.....	III
Abstract.....	IV
Agradecimientos.....	V
Capítulo I.....	1
1.1 Planteamiento del problema.....	2
1.2 Justificación.....	2
1.3 Hipótesis.....	3
1.4 Objetivos.....	3
1.4.1 Objetivo general.....	3
1.4.2 Objetivos particulares.....	4
1.5 Consideraciones generales.....	4
1.5.1 Acerca de las imágenes.....	4
1.5.2 Acerca del software.....	4
1.6 Organización del documento.....	5
Capítulo II.....	6
2.1 Punto de partida.....	7
2.2 La metodología base.....	9
2.2.1 Opciones para la detección de características.....	9
2.2.2 Opciones para la correspondencia de las características.....	10
2.2.3 Estimación del modelo de transformación y transformación de la imagen.....	11
2.3 Tendencias actuales.....	12
Capítulo III.....	15

Registro de imágenes digitales multimodales.

3.1	Procesamiento digital de imágenes.....	16
3.1.1	Imagen digital.....	16
3.1.2	Fundamentos de operaciones sobre imágenes.....	17
3.1.3	Negativo de una imagen.....	17
3.1.4	Ecualización del histograma.....	18
3.1.5	Fundamentos del filtrado espacial.....	20
3.2	Transformación de características invariantes a la escala.....	21
3.2.1	Detección del espacio de escala.....	22
3.2.2	Localización de puntos de interés.....	23
3.2.3	Asignación de la orientación.....	24
3.2.4	Descriptor de los puntos de interés.....	25
3.3	Medidas de correspondencia.....	26
3.3.1	Distancia Euclidiana.....	26
3.3.2	Transformada de Hough.....	27
3.4	Mínimos cuadrados.....	29
3.5	Funciones de transformación (geométricas).....	30
3.5.1	Transformaciones afines.....	31
3.5.2	Interpolación.....	33
3.6	Cálculo del error.....	34
Capítulo IV	36
4.1	Metodología del registro.....	38
4.1.1	Preprocesamiento.....	39
4.1.2	Selección de características.....	42
4.1.3	Correspondencia de características.....	45
4.1.4	Ajuste de características.....	52

Registro de imágenes digitales multimodales.

4.1.5 Transformación	54
Capítulo V	56
5.1 Par de imágenes de retina 2_c1	58
5.1.1 Experimento 1a	59
5.1.2 Experimento 1b	65
5.2 Experimento 2 (Bidireccional).....	69
5.2.1 Experimento 2a	70
5.2.2 Experimento 2b	72
5.3 Experimento 3 (Óptica – Fluoresceínica)	74
5.4 Experimento 4 (Fluoresceínica – Óptica)	77
5.5 Tabla de resultados	79
Capítulo VI	88
6.1 Conclusiones	89
6.2 Aportaciones	90
6.3 Trabajo futuro	91
Referencias	92

Lista de figuras

Figura 1 Ejemplo de selección de características (esquinas).....	8
Figura 2 Ejemplo de correspondencia de características (esquinas).	8
Figura 3 Ejemplo del modelo de transformación y transformación de la imagen.....	9
Figura 4 Ejemplos de detección de puntos de interés.....	10
Figura 5 Técnicas basadas en áreas y basadas en puntos usadas simultáneamente.	13
Figura 6 Representación de una imagen digital.	16
Figura 7 Ejemplo de transformación de una imagen mediante la transformada de Fourier.	17
Figura 8 Ejemplo del negativo de una imagen.	18
Figura 9 Ejemplo de ecualización del histograma.....	20
Figura 10 Ejemplo de filtro Gaussiano.....	21
Figura 11 Ejemplo de filtro mediana.....	21
Figura 12 Espacio de escala de una imagen.	23
Figura 13 Diferencia de Gaussianos.....	23
Figura 14 Asignación de posición, escala y orientación clave.	25
Figura 15 Histograma de los gradientes de una imagen.....	25
Figura 16 Distancia entre dos puntos.	26
Figura 17 Par de puntos sobre un segmento de línea.	27
Figura 18 Líneas que pasan por un punto dado.....	28
Figura 19 Representación de los espacios de imagen y de parámetros en la transformada de Hough.	29
Figura 20 Transformaciones geométricas.....	31
Figura 21 Ejemplo de interpolación lineal.	34
Figura 22 Error de predicción.....	34
Figura 23 Error cuadrático medio.....	35

Registro de imágenes digitales multimodales.

Figura 24 a) Imagen óptica, b) Imagen en escala de grises, c) Imagen fluoresceínica.	38
Figura 25 Proceso de registro de imágenes.	40
Figura 26 Diagrama de bloques del preprocesamiento de las imágenes ópticas.	41
Figura 27 Diagrama de bloques del preprocesamiento de las imágenes fluoresceínicas.	41
Figura 28 Obtención de la función diferencia de Gaussianos.	43
Figura 29 Generación del espacio de escala para una imagen óptica de retina.	43
Figura 30 Detección de los puntos clave.	44
Figura 31 Generación de descriptores.	45
Figura 32 Algoritmo del vecino más cercano mediante un árbol <i>kd</i>	47
Figura 33 Construcción de un árbol <i>k-2</i>	48
Figura 34 Árbol <i>k-2</i>	48
Figura 35 Algoritmo de la transformada de Hough generalizada.	49
Figura 36 Detección de líneas rectas en una imagen digital.	51
Figura 37 Cálculo del ángulo y la longitud.	51
Figura 38 Espacio de Hough.	52
Figura 39 Resultados de la etapa de preprocesamiento.	60
Figura 40 Resultados de la selección de características.	60
Figura 41 Correspondencia Euclidiana entre las imágenes a registrar y de referencia.	61
Figura 42 Correspondencia de Hough entre la imagen a registrar y la de referencia.	62
Figura 43 Ajuste y transformación de la imagen a registrar sobre la imagen de referencia.	63
Figura 44 Imagen de referencia e imagen resultante de la transformación.	63
Figura 45 Unión de la imagen a registrar y la imagen de referencia.	64
Figura 46 Graficación del error.	64
Figura 47 Imagen fluoresceínica a registrar (negativo) y la imagen óptica de referencia (canal verde).	65

Registro de imágenes digitales multimodales.

Figura 48 Selección de características de la imagen a registrar y la imagen de referencia..	66
Figura 49 Correspondencia Euclidiana.	66
Figura 50 Correspondencia de Hough.	67
Figura 51 Resultado de aplicar la matriz de transformación hallada sobre la imagen a registrar.	68
Figura 52 Imagen de referencia e imagen resultante de la transformación.	68
Figura 53 Unión lógica de la imagen a registrar sobre la imagen de referencia.	68
Figura 54 Graficación del error.	69
Figura 55 a), b) pares de imagenes de retina con las que no fue posible lograr el registro..	86

Lista de tablas

Tabla 1 Artículos que fundamentan la investigación.	13
Tabla 2 Relación espacio de la imagen y espacio de parámetros.	28
Tabla 3 Tabla de resultados.	79
Tabla 4 a), b), c) Resumen de los resultados de las pruebas.	87

Capítulo I

Introducción

El presente trabajo se centra en el registro de imágenes digitales multimodales, en particular en el registro de dos imágenes de la misma retina, una óptica y la otra fluoresceínica. El proceso de registro se llevará a cabo mediante varias etapas que de manera general tiene la finalidad de seleccionar características comunes, corresponderlas y encontrar los parámetros que permitan transformar la imagen a registrar para que se corresponda adecuadamente con la imagen de referencia, con el interés de que en una etapa posterior de la investigación, ya aparte de esta tesis, se puedan realizar otros estudios una vez que ambas imágenes sean coincidentes.

1.1 Planteamiento del problema

Al realizar estudios que por su naturaleza impliquen la manipulación de imágenes, uno de los problemas con los cuales nos podemos enfrentar, es que en la mayoría de los casos las imágenes no fueron tomadas bajo condiciones controladas, por lo que si es necesario compararlas de alguna forma, se hace sumamente difícil. Las diferencias principales entre las imágenes oscilarán en que una está desplazada, rotada o escalada respecto a la otra y en un caso extremo podría ser una combinación de lo anterior o que alguna de ellas tenga algún tipo de deformación. A partir de esto se impone la necesidad de alinear o hacer corresponder cada vez dos imágenes; esta es la operación que en particular se conoce como registro de imágenes, el cual por definición tiene precisamente la finalidad de hacer corresponder la imagen que se desea registrar a partir de sus características, con las características de la otra imagen que tomará el rol de imagen de referencia.

Pero al intentar conseguir esto, se presentan diversas cuestiones que hay que resolver, por ejemplo, ¿qué tipo de características pueden servir?, ¿cuántas de ellas son necesarias?, ¿cómo se corresponderán? En la literatura abierta se pueden encontrar diversas variantes para lograr posibles soluciones, de las que se deberá hacer una selección, probar su funcionamiento y evaluar la calidad del registro hecho.

1.2 Justificación

Los estudios en los que la información entrante son conjuntos de imágenes diversas que van desde un simple filtrado para eliminar ruido y así observar adecuadamente alguna escena, hasta reconocer objetos dentro de una imagen mejorada con la finalidad de realizar algún estudio definitivo o como etapa previa para realizar otros estudios.

Tal es el caso del registro de imágenes de retina, el cual debe ser realizado como paso previo a otros estudios. Un ejemplo de ello está relacionado con el interés de detectar automáticamente microaneurismas en imágenes ópticas de retina, los que desafortunadamente por su tamaño es prácticamente imposible detectarlos en ellas.

La aparición de los microaneurismas en la retina es considerada como el primer síntoma de la retinopatía diabética, segunda enfermedad retiniana que a nivel mundial, de forma silenciosa, en 10-12 años deja ciegos a los pacientes que padecen de diabetes mellitus tipo 2, cuando no se la atienden adecuadamente.

Por lo tanto, para detectar los microaneurismas se exige disponer de la información que normalmente se observa en las imágenes fluoresceínicas de retina. La obtención de las imágenes fluoresceínicas se lleva a cabo mediante un procedimiento cruento, donde el paciente debe ser inyectado por el oftalmólogo con una sustancia fluoresceínica que pone opaca la sangre. Debido a eso, en estas imágenes los microaneurismas se detectan fácilmente por ser pequeñísimas gotas de sangre producto de arteriolas y vénulas debilitadas por la enfermedad. Sin embargo, las imágenes ópticas y fluoresceínicas de las retinas de pacientes enfermos de diabetes mellitus tipo 2, son tomadas en diferentes momentos y bajo diferentes condiciones.

Hasta el momento la identificación de los microaneurismas en ambas imágenes se ha hecho manualmente. Esto implica que si el conjunto de pares de imágenes a estudiar es demasiado grande, el tiempo necesario se incrementa considerablemente, y los resultados siempre estarán sujetos a cierto grado de subjetividad por parte del investigador que realiza el estudio.

De lo anterior se impone la necesidad de contar con un mecanismo de registro automático que permita en particular registrar las imágenes de retina ópticas y fluoresceínicas tomadas de la misma retina, para después comparar la presencia de los microaneurismas fácilmente visibles en las imágenes fluoresceínicas con los detectados con dificultad en las imágenes ópticas. Ya con esto se podrá evaluar fácilmente la calidad del método de detección de los microaneurismas

1.3 Hipótesis

Mediante técnicas de procesamiento digital de imágenes y visión por computadora, es posible registrar imágenes multimodales con base en la selección de características similares entre ellas, la correspondencia de las mismas y la determinación de los parámetros necesarios para transformar la imagen a registrar.

1.4 Objetivos

El objetivo general y los particulares se exponen a continuación.

1.4.1 Objetivo general

Diseñar e implementar una metodología que permita el registro de imágenes digitales multimodales.

1.4.2 Objetivos particulares

- Lograr el registro de las imágenes de retina fluoresceínicas y ópticas.
- Implementar los algoritmos de registro necesarios para resolver la tarea.
- Diseñar y desarrollar una interfaz gráfica que permita manipular fácil y confiablemente la metodología propuesta.
- Diseñar pruebas para validar el desempeño de la implementación propuesta.

1.5 Consideraciones generales

A continuación se exponen las consideraciones generales bajo las cuales está basada la metodología de registro de imágenes digitales multimodales.

1.5.1 Acerca de las imágenes

El conjunto de pares de imágenes reales de retina ópticas y fluoresceínicas fue conseguido a través de un oftalmólogo amigo. Está compuesto por 35 pares de imágenes de diferentes retinas en la que cada par se compone de una imagen óptica y una imagen fluoresceínica de la misma retina. Las imágenes tienen una resolución espacial de 720 x 576 píxeles; las imágenes ópticas se muestran en el modelo de color RGB (*Red-Green-Blue*) y las imágenes fluoresceínicas en escala de grises.

Para poder iniciar el proceso de registro, se deben satisfacer las siguientes condiciones:

- Las imágenes deben tener el mismo tamaño.
- Las imágenes deben estar en escala de grises.
- Las imágenes deben tener una buena calidad.
- Las imágenes deben tener una resolución espacial eventualmente alta de al menos 800 x 600 píxeles por pulgada.

1.5.2 Acerca del software

El desarrollo de la aplicación de registro de imágenes multimodales fue construido sobre el lenguaje de programación M; para la interfaz gráfica se usaron tanto las bibliotecas que ofrece MATLAB en su versión 2013 y algunos elementos gráficos de JAVA, además del conjunto de herramientas *vl-feat* que ofrece opciones de selección de características. Para el módulo de procesamiento se usó C++ Builder, V. 6.0.

1.6 Organización del documento

Este trabajo consta de seis capítulos. A continuación se da un breve resumen de su contenido.

Capítulo I Introducción. En este capítulo se encuentra el planteamiento del problema que aborda este trabajo, así como su justificación y la hipótesis de la cual se parte. También se encuentran los objetivos general y específicos del trabajo y las consideraciones generales tomadas en cuenta para su desarrollo.

Capítulo II Estado del arte. Dentro de este capítulo se describen algunos trabajos previos ya publicados sobre el registro de imágenes y variantes del mismo relacionados con el trabajo desarrollado en esta tesis.

Capítulo III Marco teórico. En este capítulo se explican los conceptos y técnicas utilizados por la metodología desarrollada en este trabajo. Entre otros, nos referimos al mejoramiento de contraste, filtrado, selección de características, correspondencia, ajuste por mínimos cuadrados, etc.

Capítulo IV Desarrollo. En este capítulo se describe en detalle la metodología propuesta para conseguir el registro de imágenes multimodales. El capítulo sigue la secuencia de etapas requeridas para conseguir el registro, que son: preprocesamiento, selección de características, correspondencia de características, ajuste de características y transformación.

Capítulo V Experimentos y resultados. En este capítulo se presentan los experimentos y los resultados de las pruebas diseñadas para validar el funcionamiento de la metodología propuesta para realizar el registro.

Capítulo VI Conclusiones y trabajo futuro. En este capítulo se resumen las conclusiones del trabajo realizado y se exponen algunas mejoras y nuevos experimentos a tomar en cuenta en futuros desarrollos.

Capítulo II

Estado del arte

En este capítulo se presentan algunos de los trabajos más relevantes que fundamentan la presente investigación y que permiten direccionar el trabajo respecto a las tendencias observadas en trabajos similares. Es importante tener presente que el estudio del registro de imágenes es bastante amplio, por lo que aquí sólo se busca discutir el desarrollo que ha tenido, así como la situación actual que presenta la temática.

2.1 Punto de partida

Para establecer una referencia que permita dar una idea de lo que ha implicado el estudio del registro de imágenes, cabe señalar que se llevan más de 20 años de estudio; muestra de esto es el trabajo publicado por Brown, 1992, en el que se presenta el estudio de las técnicas de registro con las que se contaba hasta ese momento o el trabajo publicado por Zitová y Flusser, 2003, el cual es un estudio amplio sobre los métodos de registro; en particular, este último trabajo, a consideración personal, representa el material más completo sobre el tema; en la actualidad Saxena y Kumar, 2014, presentan un estudio más en el que proporciona un condensado de los métodos recientes de registro de imágenes.

De lo anterior es importante destacar que en los primeros trabajos se establecen adecuadamente clasificaciones del registro de imágenes que permiten diferenciar puntualmente los desarrollos que se han logrado. A continuación se presenta un listado con el nombre y la descripción de las clasificaciones más relevantes propuestas.

- *Clasificación basada en la aplicación.* Esta clasificación fue propuesta por Zitová y Flusser, 2003, la cual se divide a su vez en relación con el punto de vista de las imágenes, el momento en el que fueron tomadas, la forma en que fueron capturadas y el modelo de la escena.
- *Clasificación en un esquema de nueve dimensiones.* Esta clasificación es usada para imágenes médicas; las dimensiones más relevantes a destacar son la naturaleza del registro, la naturaleza de la transformación y el dominio de la transformación.
- *Clasificación basada en lo esencial.* Propuesta también por Zitová y Flusser, 2003, en la que se clasifican las técnicas de registro como basadas en áreas y basadas en características.

A partir de esos trabajos también se establece de forma genérica que el proceso de registro de imágenes se realiza en cuatro etapas, pero dejan en claro también que no se puede establecer un proceso universal para conseguir el registro, ya que se presentan diversos factores que influyen considerablemente en el registro en dependencia de lo que se desea realizar. Las cuatro etapas son:

- *Detección de características.* Extracción de estructuras sobresalientes que en general pueden ser regiones, líneas o puntos, y se definen descriptores para las

características identificadas. En la figura 1 se ejemplifica esta etapa (Zitová y Flusser, 2003).



Figura 1 Ejemplo de selección de características (esquinas).

- *Correspondencia de características.* Establecer una correspondencia entre las características presentes en ambas imágenes mediante la similitud de los descriptores de las características. En la figura 2 se ejemplifica este proceso (Zitová y Flusser, 2003).



Figura 2 Ejemplo de correspondencia de características (esquinas).

- *Estimación del modelo de transformación.* Se establecerán los parámetros de la función de tal forma que permitan alinear la imagen a registrar con la imagen de referencia.
- *Transformación de la imagen.* Se aplicará la función definida en el paso anterior para conseguir el registro. Esto se puede observar en la figura 3 (Zitová y Flusser, 2003).



Figura 3 Ejemplo del modelo de transformación y transformación de la imagen.

Es importante destacar que en cada etapa se presentan diferentes problemas que repercuten en los resultados. Estas complicaciones comienzan desde el momento en el que se decide qué tipo de característica es apropiada para una tarea dada, ya que deben tener suficientes elementos en común, los cuales se espera poder identificarlos en ambas imágenes considerando que pudiera existir algún tipo de degradación; en la siguiente etapa el tener una cantidad pequeña de características limita los resultados, o si es muy grande la cantidad de características el discriminar entre ellas puede resultar una tarea ardua; además, el tipo de descriptor que se seleccione deberá proporcionar información que permita identificar y establecer la función que transforme la imagen a registrar.

2.2 La metodología base

Las metodologías del registro de imágenes pueden variar, pero de manera genérica el proceso completo se divide en cuatro etapas, cada una de las cuales puede presentar diferentes problemáticas por lo que existen diversas formas de realizar cada una de éstas. En el siguiente apartado se abordan las más relevantes.

2.2.1 Opciones para la detección de características

Se debe tener en cuenta que se busca contar con características que estén distribuidas en toda la imagen y que efectivamente se puedan detectar en ambas imágenes. El número de elementos comunes en el conjunto de características detectadas debe de ser alto. A continuación se enlista tres variantes de detección. Ejemplo de esto se muestra en la figura 4.

- *Detección de regiones características.* Las regiones características pueden ser bosques, áreas urbanas, reservas, entre otras, las cuales se representan por medio de su centroide y son detectadas mediante métodos de segmentación que a su vez dependen de la tarea a realizar. Con el detector de esquinas de Harris y Stephens, 1988, se ha conseguido detectar potenciales regiones relevantes.

- *Detección de líneas características.* Este tipo de característica se puede representar por segmentos de líneas, contornos, caminos o estructuras anatómicas en imágenes médicas. El detector de contornos de Canny, 1986, es considerado como un algoritmo estándar de detección de contornos, aunque otro muy popular es el de Marr y Hildreth, 1980; un estudio comparativo entre algoritmos de este tipo se puede encontrar en el trabajo de Vishwakarma y Katiyar, 2011.
- *Detección de puntos de interés.* Se consideran puntos de interés aquellos puntos que unen curvas (o rectas) con gran curvatura; aquí se pueden considerar tanto esquinas como objetos irregulares. Sobre esto se explicará más ampliamente en los capítulos siguientes. También se puede observar una comparación de los métodos en el trabajo de Liu y otros, 2009.

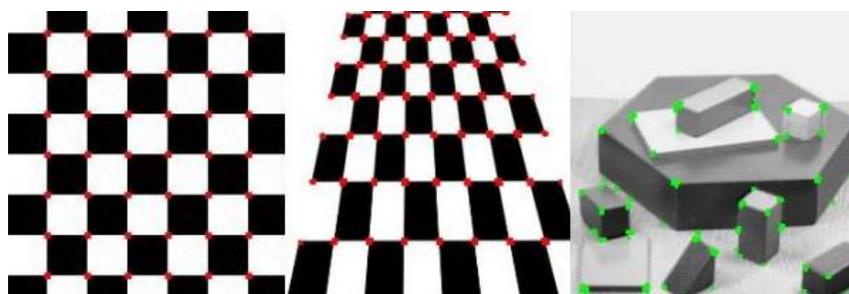


Figura 4 Ejemplos de detección de puntos de interés.

2.2.2 Opciones para la correspondencia de las características

Basado en la clasificación brindada por Zitová y Flusser, 2003, podemos realizar esta etapa mediante los métodos basados en área y los basados en características. A continuación se presenta lo más relevante en cada uno de ellos.

1. Métodos basados en área

- *Basados en los píxeles o en su correlación.* La correlación cruzada es el principio básico en este tipo de enfoque, la cual se entiende también como una correspondencia de plantillas; la correlación cruzada es un tipo de medida de similitud; en el caso particular de imágenes la correlación cruzada se define de la siguiente forma:

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

Donde $w(x, y)$ corresponde a la transformación a realizar e $f(x, y)$ es la imagen sobre la que se aplica la transformación (Gonzalez y Woods 2008).

- *Basados en la transformada de Fourier.* La transformada de Fourier de una imagen $f(x, y)$ es una función compleja en la que se obtiene una parte real $R(\omega x, \omega y)$ y una parte imaginaria $I(\omega x, \omega y)$ para cada frecuencia $(\omega x, \omega y)$ del espectro de frecuencia. La transformada de Fourier se define de la siguiente forma:

$$F(\omega x, \omega y) = |F(\omega x, \omega y)| e^{-j\varphi(\omega x, \omega y)}$$

Donde $|F(\omega x, \omega y)|$ es la magnitud y $\varphi(\omega x, \omega y)$ es el ángulo de fase. El registro con este tipo de métodos es dedicado fundamentalmente a imágenes trasladadas; el desarrollo de la transformada de Fourier está ampliamente documentado en el trabajo de Ahmed, 1995.

2. Métodos basados en características

Estos métodos usan características detectadas previamente las cuales pueden ser de diversos tipos; se presentan a continuación los más sobresalientes:

- *Métodos usando relación espacial.* Estos métodos se aplican si la característica detectada no es clara o sus vecinos están distorsionados; se usa información espacial y de similitud para conseguir el registro; detalles de esto se pueden encontrar en el trabajo de Sharma y Goyal, 2013.
- *Métodos usando descriptores invariantes.* Estos métodos también exploran información espacial, pero en este caso se consigue estableciendo correspondencias entre los descriptores de las características; la alternativa predominante ha sido el uso de SIFT (*Scale Invariant Feature Transform*), que en español se traduce como Transformación de características invariantes a la escala. Este método fue publicado por Lowe, 2004, pero hoy en día existe una gran cantidad de variantes del mismo que proporcionan diversos resultados. Este tipo de método se explicará en detalle en el próximo capítulo, ya que es parte fundamental del presente trabajo de investigación.

2.2.3 Estimación del modelo de transformación y transformación de la imagen

Después de que se han detectado las características y correspondido adecuadamente se prosigue a seleccionar el tipo de función de mapeo y sus parámetros (Saxena y Kumar, 2014). En este caso existen dos tipos diferentes de modelos que son:

- *Modelo Global.* Este modelo usa todos los puntos de correspondencia para calcular los parámetros de la función que será aplicada sobre toda la imagen; también es conocido como transformación rígida y consiste en transformaciones de rotación, traslación y escala.

- **Modelo Local.** Este modelo consiste en dividir la imagen en diferentes regiones, en la que los parámetros de la función de mapeo son calculados para cada región; esto es conocido como transformación elástica.

El remuestreo es llevado a cabo por medio de la convolución de la imagen con funciones de interpolación tales como la bilineal o bicúbica, entre otras.

2.3 Tendencias actuales

Como se habrá podido observar existen diversas formas con las que se ha resuelto el problema del registro de imágenes desde hace varios años; actualmente existen variantes que de forma novedosa tratan de realizarlo combinando la detección de dos tipos de características, las basadas en áreas y las basadas en puntos. Ejemplo de ello es el trabajo de Jin y Feng, 2014, donde se utilizó una variante de SIFT para realizar la selección de características basadas en puntos.

En estos momentos es necesario establecer de manera pertinente y clara la aportación del mecanismo SIFT, dado a conocer por Lowe, 2004, en el que propone una detección de características invariantes a la escala, pero además invariantes a la rotación en un rango de distorsión, a un cierto nivel de ruido y a cambios de iluminación; este tipo de propiedades representa ventajas considerables sobre otro tipo de métodos que se tenía hasta el momento; el proceso que se lleva a cabo para conseguir esto es complejo, por lo que se han realizado diferentes modificaciones del SIFT que en su mayoría han consistido en optimizaciones. De forma general se enlistan a continuación los pasos que se realizan; los detalles de su implementación se comentarán en el siguiente capítulo.

- *Detección del espacio de escala.* En este primer paso se realiza una búsqueda sobre todas las escalas de una imagen para identificar potenciales puntos de interés; esto se consigue por medio del uso de la función *Difference-of-Gaussians* (Diferencia de Gaussianos).
- *Localización de puntos de interés.* Para cada punto candidato se establece una localización y una escala.
- *Asignación de una orientación.* Se asignan una o más orientaciones para cada punto de interés con base en las direcciones del gradiente local de la imagen.
- *Descriptor del punto de interés.* Se crea un histograma del gradiente local para cada punto de interés.

En la figura 5 se muestra la detección de características basadas en áreas y en puntos.

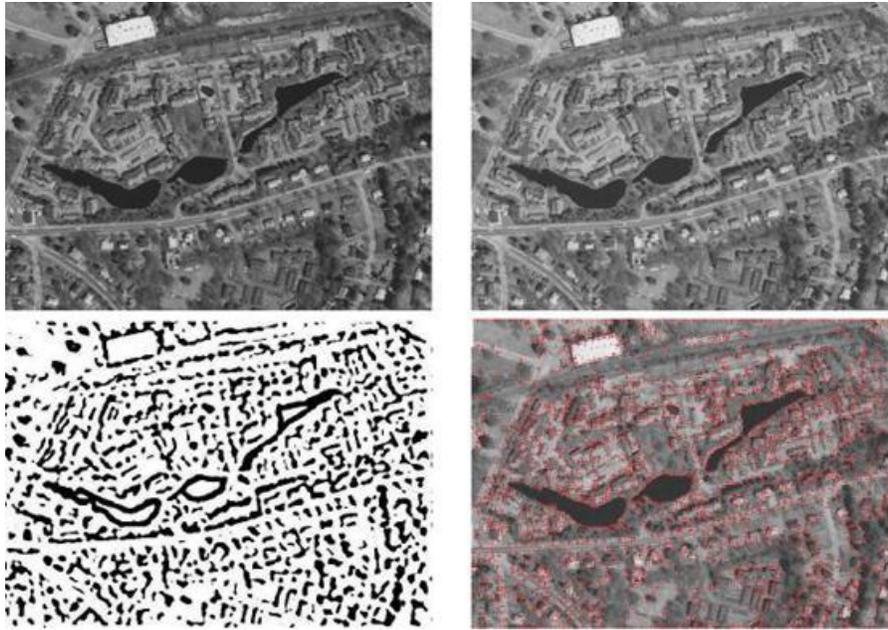


Figura 5 Técnicas basadas en áreas y basadas en puntos usadas simultáneamente.

En la tabla 1 se enlistan los artículos más relevantes que le han dado fundamento al desarrollo de este trabajo y que han permitido conocer el estado del arte de la problemática del registro de imágenes digitales.

Tabla 1 Artículos que fundamentan la investigación.

Nombre	Autor(es)	Año	Publicado en	Resumen
Image registration methods: a survey	Barbara Zitová and Jan Flusser	2003	Image and Vision Computing, Vol. 21, pp. 977-1000	Una revisión de los métodos clásicos y recientes.
Distinctive Image Features from Scale-Invariant Keypoints	David G. Lowe	2004	International Journal of Computer Vision 60(2), pp. 91–110	Descripción detallada sobre el desarrollo de SIFT.
Low-Cost Orthographic Imagery	Peter Pesti, Jeremy Elson, Jon Howell, Drew Steedly, Matt Uyttendaele	2008	ACM SIGSPATIAL International Conference, Microsoft Research	Presentan un sistema de registro de imágenes.
A robust multisource image automatic registration system based on the SIFT descriptor	Li Wang, ZhengNiu, Chaoyang Wu, RenweiXie and Huabing Huang	2012	International Journal of Remote Sensing, Vol. 33, pp. 3850-3869	Un sistema basado en SIFT y propuesta de un registro de imágenes multifuente.
Auto-Align – Multi-Modality Fluorescence	William T. E. Pitkeathly Natalie	2012	Traffic, Vol. 13, pp. 204 – 217	Presenta una herramienta para el

Microscopy Image Co-registration	S. Poulter, ElaClaridge and Joshua Z. Rappoport			registro automático de imágenes de microscopio.
A Survey of Recent and Classical Image Registration Methods	Siddharth Saxena and Rajeev Kumar Singh	2014	International Journal of Remote Sensing, Vol. 33, pp. 3850-3869	Una revisión de los métodos recientes de registro de imágenes.
Image Registration Algorithm Using Mexican Hat Function-Based Operator and Grouped Feature Matching Strategy	Feng Jin and Dazheng Feng	2014	Open Access PLOS ONE, Vol. 9	Propuesta de un método que combina la detección de características basadas en puntos y áreas.
A Quantitative Evaluation Method of Ground Control Points for Remote Sensing Image Registration	Wenting Ma, Jian Yang, Xia Ning, and Wei Gao	2014	Progress In Electromagnetics Research, Vol. 34, pp. 55-62	Describe un método para la evaluación cuantitativa de grupos de puntos de control.

Capítulo III

Marco teórico

Este capítulo integra, entre otras cuestiones, la teoría que fundamenta el proceso de registro de imágenes y en particular el proceso que se lleva a cabo desde el preprocesamiento hasta la obtención completa del registro. La organización de este capítulo seguirá un orden con base en las etapas de la metodología establecida que a grandes rasgos se enmarca en el preprocesamiento de las imágenes, la selección de las características, la correspondencia de tales características y la función de transformación (Gonzalez y Woods, 2008).

3.1 Procesamiento digital de imágenes

La visión es el sentido más avanzado del humano, por lo que juega un rol importante en el proceso de percepción de nuestro entorno; no es de sorprender que en la actualidad se realicen diversos estudios para entender cabalmente su funcionamiento.

El campo de estudio del procesamiento digital de imágenes se refiere al procesamiento de una imagen digital por medio de una computadora digital, considerando que las imágenes se componen de un número finito de elementos cada uno de los cuales tiene una localización y un valor (Gonzalez y Woods, 2008).

3.1.1 Imagen digital

Una imagen digital se puede definir como una función bidimensional, $f(x, y)$, donde x, y son coordenadas espaciales finitas y discretas y la amplitud de f en cualquier par de coordenadas (x, y) es un valor positivo, finito y discreto de la intensidad o nivel de gris de la imagen en ese punto. Cada uno de estos elementos es denominado pixel o pel (del inglés *picture element*). Por lo tanto, cada pixel posee coordenadas espaciales (x, y) en el plano de la imagen y un valor (Gonzalez y Woods, 2008). En la figura 6 se observa la representación de una imagen digital.

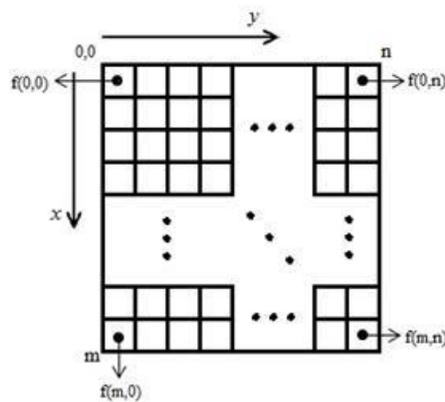


Figura 6 Representación de una imagen digital.

A continuación, mostramos algunos ejemplos de procesamiento sencillos de imágenes digitales en niveles de gris.

3.1.2 Fundamentos de operaciones sobre imágenes

Se debe tener presente que se pueden realizar operaciones sobre imágenes con diferentes enfoques; ejemplo de esto son las que se realizan en el dominio espacial o las que se realizan en el dominio de la frecuencia. Generalmente se opta por realizar operaciones en el dominio espacial, las cuales se pueden representar con la siguiente expresión.

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

Donde $f(x, y)$ es la imagen de entrada, $g(x, y)$ es la imagen de salida y T es la operación sobre f definida sobre un punto (x, y) ; la operación puede ser aplicada a una sola imagen o a n imágenes (Gonzalez y Woods, 2008). En la figura 7 se presenta a la derecha la transformada de Fourier de la imagen de la izquierda.



Figura 7 Ejemplo de transformación de una imagen mediante la transformada de Fourier.

3.1.3 Negativo de una imagen

El negativo de una imagen con niveles de intensidad $[0, L - 1]$ es obtenido mediante el uso de una transformación negativa, la cual se puede representar por la siguiente expresión (Gonzalez y Woods, 2008).

$$S = L - 1 - r$$

Invirtiendo los niveles de intensidad de una imagen de esta manera, se produce el equivalente a un negativo fotográfico. En la figura 8 se presenta el resultado de aplicar una transformación negativa.



Figura 8 Ejemplo del negativo de una imagen.

3.1.4 Ecuación del histograma

El histograma de una imagen digital con niveles de intensidad en un rango $[0, L - 1]$ es una función discreta $h(r_k) = n_k$, donde r_k es el k -ésimo valor de intensidad y n_k es el número de píxeles en la imagen con intensidad r_k ; una práctica común para normalizar el histograma es dividir cada uno de sus componentes por el número total de píxeles en la imagen, denotado por el producto de MN , donde M y N son la cantidad de filas y columnas de la imagen, respectivamente. La normalización del histograma está dada por $p(r_k) = \frac{n_k}{MN}$, para $k = 1, 2, \dots, L - 1$; en otras palabras, $p(r_k)$ es una estimación de la probabilidad de ocurrencia del nivel de intensidad r_k en una imagen. Una de las aplicaciones más importantes de la ecualización del histograma es el mejoramiento del contraste (Gonzalez y Woods, 2008).

La ecualización del histograma parte de lo siguiente, consideremos valores de intensidad continuos donde la variable r denota las intensidades de una imagen; asumimos que r está en el rango $[0, L - 1]$, donde $r = 0$ representa el color negro y $r = L - 1$ representa el blanco. Para que r satisfaga esa condición, consideramos la transformación (mapeo de intensidad) denotada por $T(r)$ $0 \leq r \leq L - 1$, que produce una salida de nivel de intensidad s para cada píxel de la imagen de entrada que tiene intensidad r ; en esto asumimos que:

- a) $T(r)$ es una función monótonicamente creciente en el intervalo $0 \leq r \leq L - 1$
- b) $0 \leq T(r) \leq L - 1$ para $0 \leq r \leq L - 1$.

Si consideramos la inversa de s se establece que:

$$r = T^{-1}(s) \text{ para } 0 \leq s \leq L - 1$$

El requerimiento en la condición a) garantiza que el valor de intensidad de salida nunca será menor que el correspondiente valor de entrada; la condición b) garantiza que el rango de intensidades de la salida es el mismo que el de la entrada.

Los niveles de gris de una imagen pueden ser vistos como una variable aleatoria en el intervalo $[0,1]$. Uno de los descriptores fundamentales de las variables aleatorias es la función de densidad de probabilidades (FDP).

Una función de transformación de variables aleatorias de particular importancia en el procesamiento de imágenes es la función de distribución acumulativa (FDA). Como las FDP son siempre positivas y recordando que la integral de una función está dada por el área bajo la curva, la función de transformación es de valor único y monótonicamente creciente, además de que se mantienen en el intervalo $[0,1]$ (satisfacen las condiciones a y b).

Para valores discretos trabajamos con probabilidades y sumatorias en lugar de funciones de densidad de probabilidad e integrales. La probabilidad de ocurrencia de un nivel de gris r_k en una imagen se aproxima como $p_r(r_k) = \frac{n_k}{n}$ para $k = 0, 1, 2, \dots, L - 1$, donde n es el número total de píxeles de la imagen, n_k es el número de píxeles que tienen el nivel de gris r_k , y L es el número total de posibles niveles de gris en la imagen.

La versión discreta de la función de transformación, es decir, de la función de distribución acumulativa es:

$$s_k = T(r_k) = \sum_{j=0}^k p_r(r_j)$$

$$s_k = \sum_{j=0}^k \frac{n_j}{n} \text{ para } k = 0, 1, 2, \dots, L - 1$$

La imagen de salida se obtiene mapeando cada pixel con valor r_k de la imagen de entrada a su correspondiente nivel s_k de la imagen de salida. A esta transformación se le llama ecualización del histograma o linealización del histograma, la cual satisface también las condiciones a y b.

El resultado de la ecualización del histograma es similar a la expansión del contraste, pero ofrece la ventaja de ser totalmente automático sin necesidad de dar ningún parámetro adicional, pues la ecualización del histograma determina automáticamente una función de transformación para producir una nueva imagen $p_s(s)$ (imagen de salida), con un histograma uniforme (en la práctica cercano a ser uniforme), independientemente de la forma de $p_r(r)$ (imagen de entrada).

La ecualización del histograma determina automáticamente la función de transformación y busca producir una imagen de salida que tenga un histograma aproximadamente uniforme, como se muestra en la figura 9.

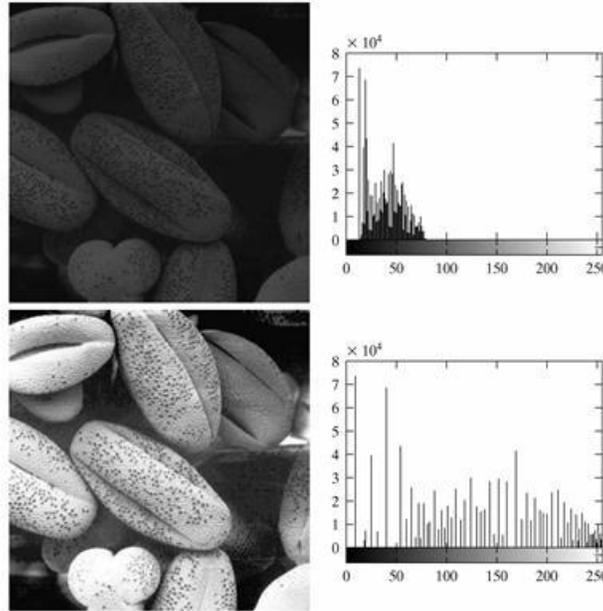


Figura 9 Ejemplo de ecualización del histograma.

3.1.5 Fundamentos del filtrado espacial

El filtrado espacial es una de las principales herramientas usadas en el campo del procesamiento digital de imágenes. El filtrado espacial se compone de dos elementos, una vecindad, típicamente representada por un rectángulo pequeño, y una operación predefinida que se realizará sobre los píxeles de la imagen que componen su vecindad. El filtro crea un nuevo píxel con coordenadas iguales a las coordenadas de la vecindad y su valor es el resultado de la operación de filtrado. Una imagen filtrada es la obtenida tras la aplicación del filtro en cada píxel que compone la imagen de entrada.

De forma general, un filtro espacial lineal para una imagen de $M \times N$ píxeles está dado por la expresión:

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

Donde x y y varían en la vecindad w de cada píxel en f . En la figura 10 observamos un ejemplo de un filtrado Gaussiano (filtro pasa-bajas): la primera es la imagen de entrada, la segunda es la imagen una vez que se le ha aplicado el filtro Gaussiano con una cierta

desviación estándar, y la tercera imagen corresponde a la aplicación del filtro a la misma imagen de entrada pero con un valor de desviación estándar mayor.



Figura 10 Ejemplo de filtro Gaussiano.

Otro tipo de filtro es el filtro mediana. En particular un filtro mediana consiste en sustituir el valor de la mediana de los pixeles vecinos de cada pixel de la imagen situados bajo la ventana del filtro. Este filtro es idóneo para reducir los efectos del ruido multiplicativo de sal y pimienta, tal como se puede observar en la figura 11, donde se tiene una imagen con un alto grado de contaminación; las dos imágenes subsecuentes fueron suavizadas aplicando una y dos veces el filtro mediana, respectivamente.



Figura 11 Ejemplo de filtro mediana.

3.2 Transformación de características invariantes a la escala

Para detectar puntos clave en una imagen digital se sigue un enfoque de filtrado en cascada que usa algoritmos eficientes para identificar candidatos que serán examinados más adelante en el proceso de registro. En nuestro caso esto se llevó a cabo con la ayuda del software de transformación de características invariante a la escala, SIFT (*Scale Invariant Feature Transform* por sus siglas en inglés). El primer paso para realizar la detección de puntos clave es identificar la posición y la escala que pueden ser repetidamente asignadas bajo diferentes vistas del mismo objeto. La detección de puntos clave invariantes a cambios de escala de una imagen puede realizarse con la búsqueda de características estables a

través de todas las posibles escalas, usando una función continua de escala que se conoce como espacio de escala (Lowe, 2004).

3.2.1 Detección del espacio de escala

El espacio de escala es una teoría formal para el manejo de estructuras de la imagen a diferentes escalas, en la que se representa a una imagen como una familia de un solo parámetro de imágenes suavizadas; la representación del espacio de escala está parametrizada por el tamaño del kernel de suavizado utilizado para suprimir estructuras a una escala más fina.

La noción del espacio de escala se aplica a las señales de un número arbitrario de variables. El caso más común en la literatura se aplica a las imágenes bidimensionales. Para una imagen dada $f(x, y)$, la representación lineal (Gaussiana) del espacio de escala es una familia de señales derivadas $L(x, y; \sigma)$ definida por la convolución de la función $f(x, y)$ con el kernel Gaussiano:

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma}$$

de tal manera que:

$$L(x, y; \sigma) = G(x, y; \sigma) * f(x, y)$$

donde el punto y coma en el argumento de L implica que la convolución se realiza sólo sobre las variables x, y , mientras que el parámetro de escala σ después del punto y coma sólo indica qué nivel de escala se está definiendo. Esta definición de L funciona para un conjunto de escalas $t \geq 0$, aunque típicamente sólo un conjunto discreto finito de niveles sería considerado realmente en la representación del espacio de escala.

El parámetro de escala $t = \sigma^2$ es la varianza del filtro Gaussiano y como límite para $t = 0$ el filtro G se convierte en una función de impulso de tal manera que $L(x, y; 0) = f(x, y)$, es decir, la representación del espacio de escala al nivel de escala $t = 0$ es la misma imagen f . Cuando t aumenta, L es el resultado del suavizado de f con un filtro cada vez mayor, eliminando así más y más los detalles que contiene la imagen (Wiley, 2009). En la figura 12 se muestra el espacio de escala de una imagen digital.

Una de las formas de implementar el espacio de escala es mediante la función Diferencia de Gaussianos, la cual puede ser calculada a partir de la diferencia de dos escalas consecutivas separadas por una constante k , esto es:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$

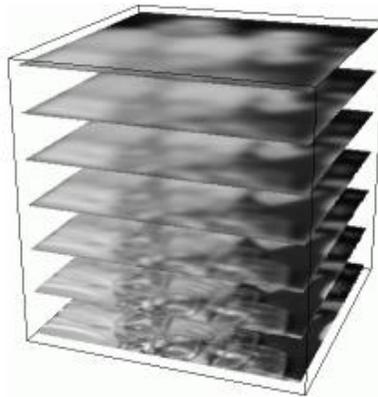


Figura 12 Espacio de escala de una imagen.

La figura 13 ejemplifica la $D(x, y, \sigma)$ que estableció Lowe, 2004, en su trabajo.

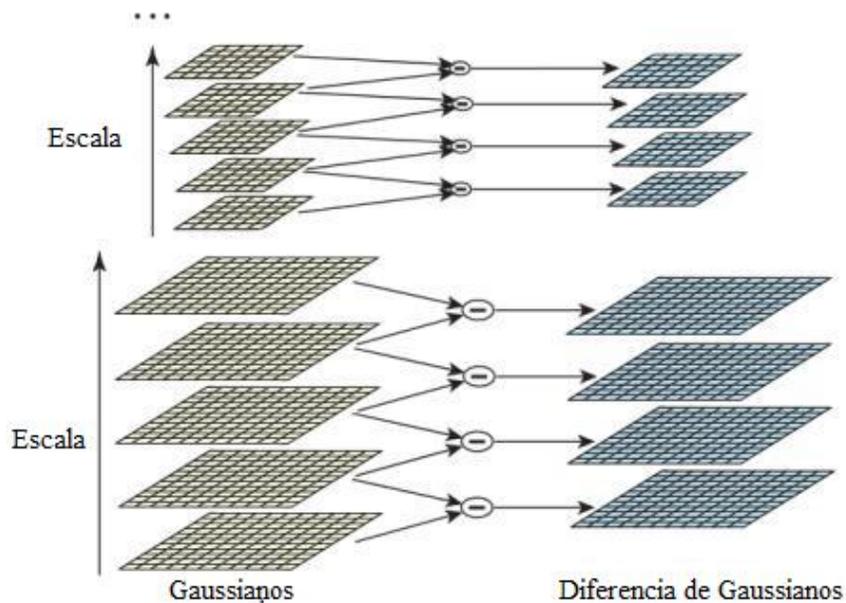


Figura 13 Diferencia de Gaussianos.

3.2.2 Localización de puntos de interés

Detector. Los marcos de SIFT (puntos clave) se extraen basándose en los extremos locales (picos) del espacio de escala de $D(x, y, \sigma)$. Numéricamente, los extremos locales son todos los elementos cuyos $3 \times 3 \times 3$ vecinos en el espacio de escala tienen el valor más pequeño o más grande. Una vez extraídos, los extremos locales se interpolan de forma cuadrática. Esto es muy importante especialmente en las escalas de menor resolución, con el fin de tener la

localización exacta del punto clave en la resolución completa. Por último, se filtran los puntos clave para eliminar las respuestas de bajo contraste o respuestas cercanas a los bordes; después de esto último se asignan las orientaciones.

Eliminación de las respuestas de bajo contraste. Los picos demasiado cortos pueden haber sido generados por el ruido y se deben descartar. Esto se hace mediante la comparación del valor absoluto del espacio de escala $D(x, y, \sigma)$ con respecto a cada pico y con un umbral preseleccionado.

Eliminación de las respuestas de borde. Los picos demasiado planos suelen generarse por los bordes y no se consideran como características estables. Estos picos se detectan y eliminan como se indica a continuación. Dado un pico (x, y, σ) , el algoritmo evalúa las (x, y) mediante una matriz Hessiana (matriz que contiene las segundas derivadas parciales que sirven para verificar si el punto crítico del que se está hablando es máximo, mínimo, punto silla o no puede determinarse) del espacio de escala $D(x, y, \sigma)$ en una cierta escala t ; enseguida se realiza una votación calculada con la siguiente expresión (similar a la detección de bordes que realiza Harris y Stephens, 1988):

$$\frac{(tr D(x, y, \sigma))^2}{\det D(x, y, \sigma)}, D = \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial x \partial y} \\ \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial y^2} \end{bmatrix}$$

Donde $tr(D(x, y, \sigma))$ representa la suma de los elementos de la diagonal principal de una matriz, y $\det D(x, y, \sigma)$ representa el cálculo del determinante de una matriz.

Esta votación tiene un mínimo (igual a 4) cuando ambos valores propios de la Jacobiana son iguales (pico curvado) y aumenta a medida que uno de los valores propios crece y el otro se mantiene pequeño, esto es, para aproximar linealmente a la función en un punto, un grupo y una variedad algebraica asociada a la curva, la cual puede ser embebida. Los picos se mantienen si la puntuación es inferior a $(t_e + 1)(t_e + 1)/t_e$, donde t_e es el umbral de borde. Debe tenerse en cuenta que esta cantidad tiene un mínimo igual a 4 cuando $t_e = 1$, el cual crece a partir de entonces. Por lo tanto el rango del umbral de borde es $[1, \infty)$.

3.2.3 Asignación de la orientación

Un pico en el espacio de escala $D(x, y, \sigma)$ fija 2 parámetros del punto clave: la posición y la escala. Queda por elegir una orientación. Con este fin SIFT calcula un histograma de las orientaciones del gradiente en una ventana Gaussiana con una desviación estándar 1,5 veces más grande que la escala σ del punto clave. Para darnos una idea de esto observemos la figura 14 donde percibimos que para cada punto clave se le asigna una orientación (ángulo) y una escala.

Este histograma es entonces normalizado y se selecciona el máximo (moda). Además de la moda mayor, se seleccionan otras tres modas cuya amplitud esté dentro del 80% de la moda mayor, las cuales son devueltas como orientaciones adicionales (Vedaldi y Fulkerson, 2008).

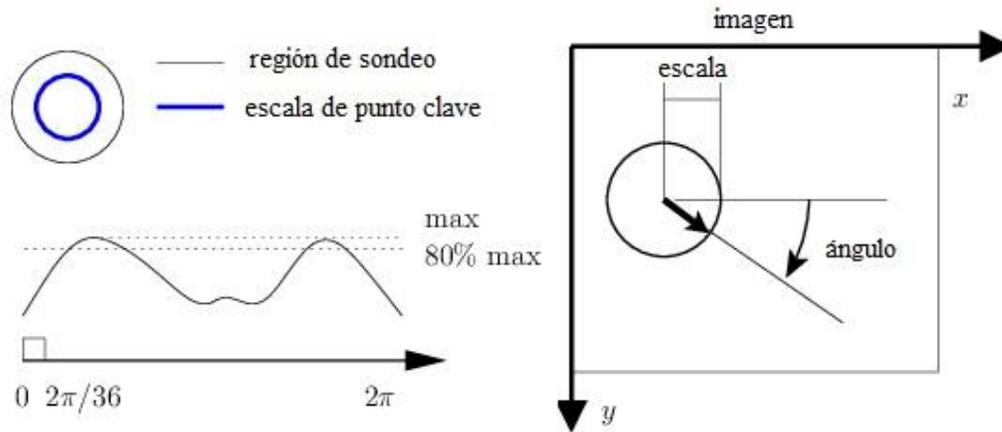


Figura 14 Asignación de posición, escala y orientación clave.

3.2.4 Descriptor de los puntos de interés

Un descriptor SIFT de una región local (punto clave) es un histograma espacial 3-D de los gradientes de la imagen. El gradiente en cada pixel se considera como parte de un vector de características primarias tridimensional, compuesto por la ubicación del pixel y la orientación del gradiente. Las muestras son ponderadas con la norma del gradiente y se acumulan en un histograma 3-D; además se normaliza nuevamente con base en una función Gaussiana con la finalidad de otorgar mayor importancia a los elementos centrales del descriptor, (Vedaldi y Fulkerson, 2008), como se muestra en la figura 15.

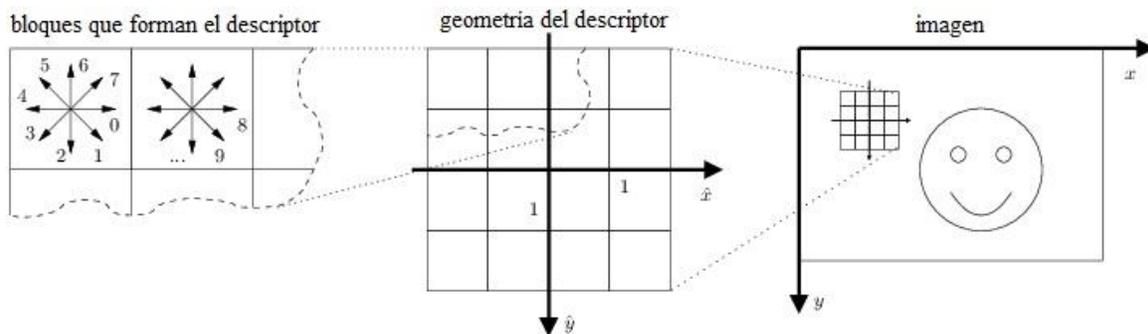


Figura 15 Histograma de los gradientes de una imagen.

3.3 Medidas de correspondencia

Cuando en el proceso de registro de imágenes se han detectado las características que se decidieron usar, se debe establecer un mecanismo de correspondencia que pueda discriminar aquellas características que no coincidan.

El algoritmo base para realizar esto fue propuesto por Lowe, 2004, el cual se basa en obtener los descriptores de la imagen de referencia y los de la imagen que se desea registrar; una vez que se tienen los vectores que representan los descriptores, por cada uno de estos en la imagen de referencia se calcula la distancia Euclidiana hacia cada uno de los descriptores de la imagen a registrar, y se define un umbral con el que se discriminarán aquellas características que resultaron más diferentes. A mayor distancia entre los descriptores, más diferentes resultarán las características.

3.3.1 Distancia Euclidiana

En matemáticas, la distancia Euclidiana o Euclídea, es la distancia entre dos puntos de un espacio Euclídeo, la cual se deduce a partir del teorema de Pitágoras.

Por ejemplo, en un espacio bidimensional, la distancia Euclidiana entre dos puntos p_1 y p_2 , de coordenadas cartesianas (x_1, y_1) y (x_2, y_2) respectivamente, es:

$$d_E(p_1, p_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

En general, la distancia Euclidiana entre los puntos $P = (p_1, p_2, \dots, p_n)$ y $Q = (q_1, q_2, \dots, q_n)$ del espacio Euclídeo n -dimensional, se define como:

$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

La figura 16 muestra gráficamente la distancia en un sistema de coordenadas cartesianas.

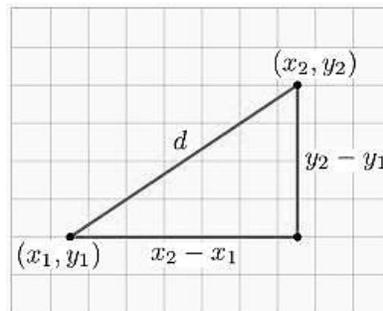


Figura 16 Distancia entre dos puntos.

3.3.2 Transformada de Hough

La transformada de Hough es un enfoque general para la localización de cualquier forma que pueda ser descrita paramétricamente dentro de una distribución de puntos, por ejemplo formas geométricas como líneas, círculos y elipses que pueden ser descritas usando ecuaciones con pocos parámetros. La Transformada de Hough es utilizada mayormente para detectar segmentos de línea en segmentos de bordes.

Para clarificar esto consideremos lo siguiente: una línea puede ser descrita con dos parámetros mediante la ecuación de pendiente-intercepto $y = kx + d$ donde k es la pendiente y d el intercepto en el eje y , como lo podemos observar en la figura 17.

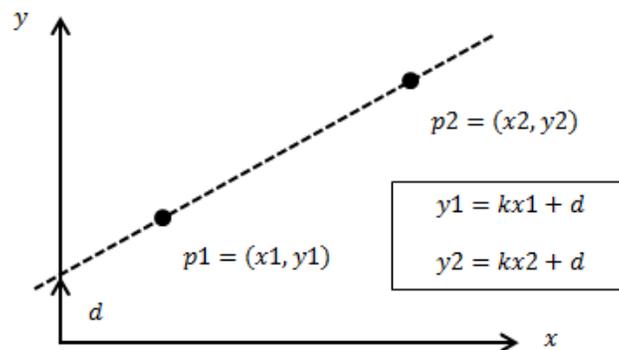


Figura 17 Par de puntos sobre un segmento de línea.

Un segmento de línea que pasa a través de dos puntos dados $p_1 = (x_1, y_1)$ y $p_2 = (x_2, y_2)$ debe satisfacer las siguientes condiciones $y_1 = kx_1 + d$ y $y_2 = kx_2 + d$ para $k, d \in \mathbb{R}$, donde el objetivo es encontrar valores de k y d tales que tantos más puntos de un borde se superpongan a la línea que estos describen. Pero para determinar la cantidad de puntos de un borde que se superponen a una línea, se debe considerar la generación de un espacio de parámetros como se describe a continuación.

Espacio de parámetros.

El enfoque de la transformada de Hough examina todos los posibles segmentos de línea que pasan a través de un punto dado en la imagen, lo cual se puede observar en la figura 18:

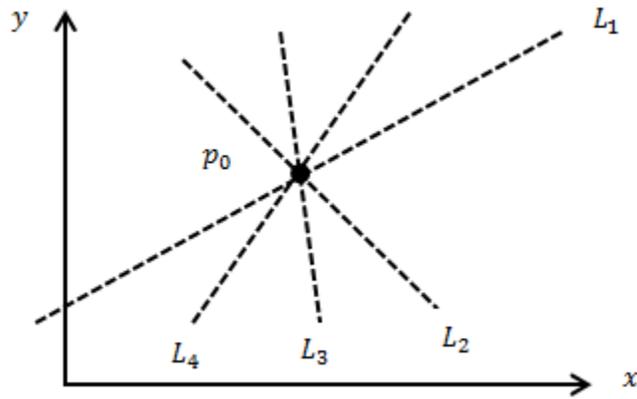


Figura 18 Líneas que pasan por un punto dado.

Cada línea L_j que pasa a través de un punto $p_o = (x_o, y_o)$ debe satisfacer la condición $L_j : y_o = k_j x_o + d_j$ para todo valor apropiado k_j, d_j . Las posibles soluciones para k_j, d_j corresponden al conjunto infinito de líneas que pasa a través de un punto p_o . Para una k_j dada la solución d_j es $d_j = -x_o k_j + y_o$, la cual es otra ecuación para una línea donde k_j, d_j ahora son las variables y x_o, y_o las constantes de la ecuación. El conjunto $\{(k_j, d_j)\}$ describe los parámetros de todas las posibles líneas L_j que pasan a través del punto en la imagen $p_o = (x_o, y_o)$. Para un punto arbitrario, $p_i = (x_i, y_i)$ describe la línea $M_i: d = -x_i k + y_i$ con los parámetros $-x_i, y_i$ en el llamado espacio de parámetros.

La relación entre el espacio de la imagen y el espacio de parámetros se puede resumir en la siguiente tabla.

Tabla 2 Relación espacio de la imagen y espacio de parámetros.

	Espacio de la imagen (x, y)	Espacio de parámetros (k, d)	
Punto	$p_i = (x_i, y_i)$	$M_i: d = -x_i k + y_i$	Línea
Línea	$L_j : y = k_j x + d_j$	$q_j = (k_j, d_j)$	Punto

Cada punto de la imagen P_i y su línea límite asociada corresponde exactamente a una línea M_i en el espacio de parámetros. La siguiente figura ilustra cómo la línea M_1 y M_2 intersectan en la posición $q_{12} = (k_{12}, d_{12})$ en el espacio de parámetros, lo cual significa que (k_{12}, d_{12}) son los parámetros de la línea en el espacio de la imagen que pasa por ambos puntos p_1 y p_2 . Las líneas M_i que intersectan en un solo punto en el espacio de parámetros son la cantidad de puntos en el espacio de la imagen que se sobreponen sobre un segmento

de línea. Para poder identificar la intercepción de los puntos en el espacio de parámetros se debe hacer uso de un acumulador, el que se detalla a continuación. En la figura 19 podemos observar del lado izquierdo la información del espacio de la imagen y del lado derecho la información correspondiente al espacio de parámetros.

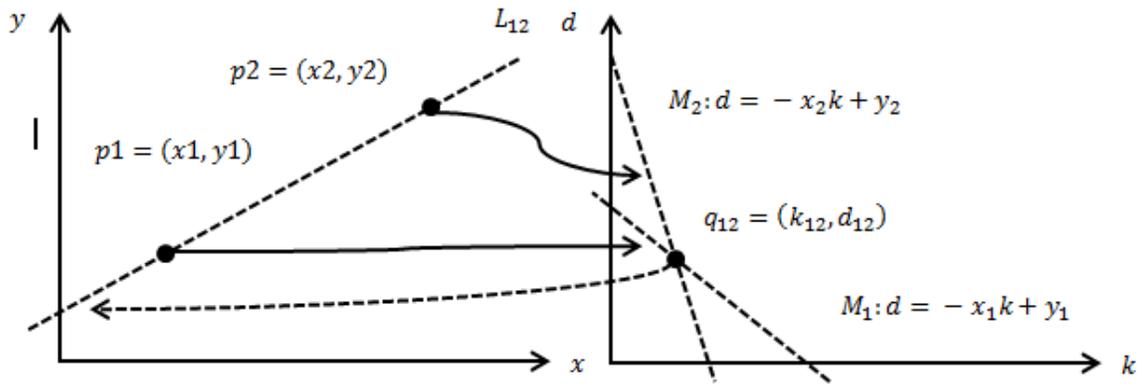


Figura 19 Representación de los espacios de imagen y de parámetros en la transformada de Hough.

Acumulador

Encontrar las líneas dominantes en la imagen puede ahora reformularse como encontrar todas los puntos en el espacio de parámetros donde un número considerable de líneas se intersectan. Este es el objetivo de la transformada de Hough. Para poder llevar a cabo esto, se debe definir una representación discreta del espacio de parámetros continuos seleccionando un apropiado tamaño para los ejes (k, d) . Una forma de realizar esto es mediante un arreglo. Este arreglo será usado para mantener el control del número de veces que se intersectan líneas en el espacio de parámetros. Cada línea en el espacio de parámetros es analizada y colocada en el acumulador, al final de lo cual las casillas del acumulador tendrán contabilizado el número de líneas del espacio de parámetros que pasaron por un mismo punto.

3.4 Mínimos cuadrados

El término mínimos cuadrados describe un enfoque usado para resolver sistemas de ecuaciones especificados inexactamente en algún sentido apropiado. En lugar de resolver las ecuaciones exactamente, se busca solamente minimizar la suma de los cuadrados de los residuos.

Si se hacen suposiciones probabilísticas apropiadas acerca de la distribución del error, el enfoque de mínimos cuadrados produce lo que se conoce como la estimación de máxima verosimilitud de los parámetros. Existen diversos problemas que pueden ser abordados

usando el enfoque de mínimos cuadrados; un ejemplo común de esto es el problema del ajuste de curvas a un conjunto de datos dados, el cual se detalla a continuación.

Suponiendo que tenemos m observaciones $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, donde $y_i = y(x_i), i = 1, 2, \dots, m$, la idea es modelar $y(x)$ por medio de una combinación de n funciones base $f_1(x), f_2(x), \dots, f_n(x)$. Es decir, debemos suponer que las funciones se ajustan a los datos en una combinación lineal de la forma $y(x_i) \approx c_1 f_1(x_i) + c_2 f_2(x_i) + \dots + c_n f_n(x_i), i = 1, 2, \dots, m$.

La última expresión constituye un sistema de m ecuaciones y n incógnitas que en forma matricial puede expresarse de la siguiente forma:

$$\begin{bmatrix} f_1(x_1) & f_2(x_1) & \dots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_m) & f_2(x_m) & \dots & f_n(x_m) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, m \geq n.$$

A la matriz de este sistema $A = (a_{ij})$ con $a_{ij} = f_j(x_i)$ se le denomina matriz de diseño y dado que $m \geq n$, entonces el sistema tiene más renglones (ecuaciones) que columnas (incógnitas). A este tipo de sistemas se les denomina sobredeterministas y generalmente no tienen solución. Las funciones base $f_i(x)$ pueden escogerse de muchas maneras; algunas elecciones comunes pueden ser polinomios, funciones racionales, exponenciales, Gaussianas, etc.; una vez establecida la función a ajustar se pueden determinar sus parámetros.

- En aproximación de funciones, aparte de la norma de mínimos cuadrados (L_2), se utilizan usualmente la norma L_1 y la llamada norma L_∞ ; la aproximación de funciones que minimiza la norma L_∞ se conoce como aproximación minimax. Cuando deseamos una aproximación a una función en un intervalo por otra más sencilla, la aproximación minimax es quizás la más razonable, ya que limita el error máximo cometido en un punto arbitrario del intervalo. Sin embargo, cuando tenemos puntos experimentales afectados de un error estadístico, entonces la aproximación de mínimos cuadrados, en la versión de mínimo x^2 es la única justificada desde el punto de vista estadístico.

3.5 Funciones de transformación (geométricas)

Una vez que se conocen las características que más se corresponden, se prosigue a establecer una función que sobreponga la imagen a registrar sobre la imagen de referencia; esto se realiza mediante funciones de transformación, cuyo objetivo es deformar una imagen alterando su geometría; ejemplos de estas son la translación, la rotación y el escalamiento, (Burger y Burge, 2008); parte de esto se puede observar en la figura 20 para la translación y la escala.



Figura 20 Transformaciones geométricas.

3.5.1 Transformaciones afines

El objetivo de las operaciones geométricas es deformar una imagen alterando su geometría; ejemplos de éstas son el desplazamiento (o traslación), la rotación y el cambio de escala. En un principio, una transformación geométrica a una imagen dada I da como resultado una nueva imagen I' por medio de la modificación de las coordenadas de los píxeles de la imagen, esto es:

$$I(x, y) \rightarrow I'(x', y');$$

El valor de la función I que originalmente estaba localizado en la coordenada (x, y) es movido a la posición (x', y') en la función I' .

Para modelar este proceso se necesita una función de transformación:

$$T : R^2 \rightarrow R^2$$

que especifique para cada coordenada original del punto $px = (x, y)$, el correspondiente punto $px' = (x', y')$ en la nueva imagen I' , esto es:

$$x' = T(x)$$

Es importante notar que las coordenadas (x, y) y (x', y') especifican el valor real de los puntos en el plano de la imagen continua $R \times R$; sin embargo, el problema en la transformación digital de imágenes es que los píxeles $I(u, v)$ no están definidos sobre un plano continuo sino en una trama discreta $Z \times Z$, por lo que una coordenada transformada $(u', v') = T(u, v)$ producida por la función de mapeo $T()$ no caerá sobre la trama discreta; la solución para este problema es calcular valores intermedios de píxeles para la imagen transformada por medio de un proceso llamado interpolación.

Función de mapeo.

La función de mapeo $T()$ es una función continua arbitraria que por razones de simplicidad es a menudo especificada en dos funciones separadas;

$$\begin{aligned} x' &= T_x(x, y) \\ y & \\ y' &= T_y(x, y). \end{aligned}$$

Las funciones de mapeo simples son la traslación, el escalamiento, y la rotación, definidas de la siguiente forma:

- La traslación respecto a un vector (d_x, d_y) se puede representar de las siguientes formas:

$$\begin{aligned} T_x: x' &= x + d_x; T_y: y' = y + d_y \\ \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \end{pmatrix}. \end{aligned}$$

- El escalamiento a lo largo del eje x o y por un factor s_x o s_y respectivamente, se puede representar de las siguientes formas:

$$\begin{aligned} T_x: x' &= x * s_x; T_y: y' = y + s_y \\ \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix} * \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix}. \end{aligned}$$

- La rotación por un ángulo α (considerando al origen como el centro de rotación) se puede representar de las siguientes formas:

$$\begin{aligned} T_x: x' &= x * \cos \alpha - y \sin \alpha; T_y: y' = x * \sin \alpha + y \cos \alpha \\ \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} * \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

Coordenadas homogéneas.

Las ecuaciones enlistadas en la sección previa, constituyen una importante clase de funciones de mapeo conocidas también como transformaciones afines. Para poder concatenar estas operaciones es conveniente especificar las operaciones en forma de multiplicación de matrices; sin embargo, algunas como la traslación no pueden ser descritas de esta forma. Para ello se usan las coordenadas homogéneas. Para convertir las coordenadas dentro de una representación de coordenadas homogéneas, cada vector es extendido agregando un elemento h .

$$x = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{Se convierte en} \quad x' = \begin{pmatrix} x' \\ y' \\ h \end{pmatrix} = \begin{pmatrix} hx \\ hy \\ h \end{pmatrix}.$$

3.5.2 Interpolación

En el subcampo matemático del análisis numérico, se denomina interpolación a la obtención de nuevos puntos a partir del conocimiento previo de un conjunto discreto de puntos contiguos.

En ingeniería y algunas ciencias es frecuente disponer de un cierto número de puntos obtenidos por muestreo o a partir de un experimento, para después pretender construir una función que los ajuste.

Otro problema estrechamente ligado con el de la interpolación es la aproximación de una función complicada por una más simple. Si tenemos una función cuyo cálculo resulta costoso, podemos partir de un cierto número de sus valores e interpolar dichos datos construyendo una función más simple. En general, por supuesto, no obtendremos los mismos valores evaluando la función obtenida que si evaluamos la función original, aunque dependiendo de las características del problema y del método de interpolación usado, la ganancia en eficiencia puede compensar el error cometido.

En todo caso se trata de que a partir de n parejas de puntos (x_k, y_k) , obtener una función f que verifique:

$$f(x_k) = y_k, k = 1, 2, \dots, n$$

a la que se denomina función interpolante de dichos puntos. A los puntos x_k se les llama nodos. Algunas formas de interpolación que se utilizan con frecuencia son la interpolación lineal, la interpolación polinómica (de la cual la anterior es un caso particular), la interpolación por medio de *spline* o la interpolación polinómica de Hermite.

Uno de los métodos de interpolación más sencillo es el lineal. En general, en la interpolación lineal se utilizan dos puntos, (x_a, y_a) y (x_b, y_b) para obtener un tercer punto interpolado (x, y) a partir de la siguiente fórmula:

$$y = y_a + (x - x_a) \frac{(y_b - y_a)}{(x_b - x_a)}$$

La interpolación lineal es rápida y sencilla, pero en ciertos casos no es muy precisa, como se observa en la figura 21.

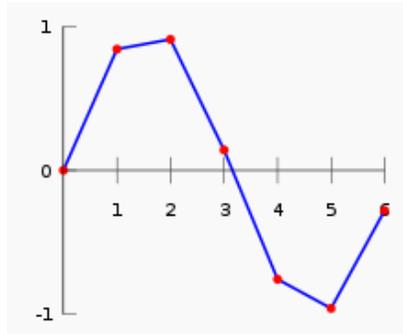


Figura 21 Ejemplo de interpolación lineal.

3.6 Cálculo del error

Considerando el método de regresión que se usa para predecir y a partir de valores de x , se conoce que los valores difieren, ¿pero en que magnitud difieren?; en este punto es donde el error cuadrático medio permite medir el tamaño global de las diferencias que se dan entre los valores de la variable y las predicciones; a estas diferencias se les denomina residuos, que si hablamos gráficamente será la distancia a la que está una observación real por encima o por debajo de la recta de regresión; este residuo es el equivalente gráfico del error de predicción cometido por el método de regresión; la siguiente gráfica muestra la representación del error.

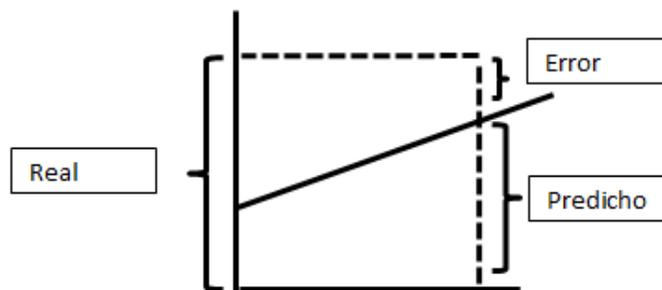


Figura 22 Error de predicción.

El tamaño global de los errores se mide calculando la raíz cuadrada de la media de los cuadrados de las diferencias de los valores; el resultado recibe el nombre de error cuadrático medio considerado un estimador T para estimar $g(\theta)$ y se define como:

$$E_{\theta} [(T - g(\theta))^2] = \begin{cases} \sum_{i=1}^N (T(x_1, \dots, x_N) - g(\theta))^2 P_{\theta}(x_1) \dots P_{\theta}(x_N) & \text{(caso directo)} \\ \int_{\mathbb{R}} (T(x_1, \dots, x_N) - g(\theta))^2 f_{\theta}(x_1) \dots f_{\theta}(x_N) dx_1 \dots dx_N & \text{(caso continuo)} \end{cases}$$

Es claro que un estimador será más efectivo cuanto más pequeño sea su error cuadrático medio.

El siguiente desarrollo nos aporta una fórmula sencilla para el cálculo de este error:

$$\begin{aligned} E_{\theta} \left[(T - g(\theta))^2 \right] &= E_{\theta} \left[(T - E_{\theta}[T] + E_{\theta}[T] - g(\theta))^2 \right] \\ &= E_{\theta} \left[(T - E_{\theta}[T])^2 \right] + (E_{\theta}[T] - g(\theta))^2 \\ &= V_{\theta}(T) + (\text{Sesgo}(T))^2 \end{aligned}$$

Donde $\text{Sesgo}(T) = E_{\theta}[T] - g(\theta)$.

De esta última igualdad vemos que podemos reducir el error cuadrático medio, bien reduciendo la varianza del estimador o reduciendo su sesgo para estimar $g(\theta)$. Podemos eliminar completamente el sesgo al trabajar con estimadores insesgados.

El error cuadrático medio nos indica lo alejado que se halla, por encima o por debajo de una recta de regresión, un punto típico de la nube de puntos, tal como se observa en la 23.

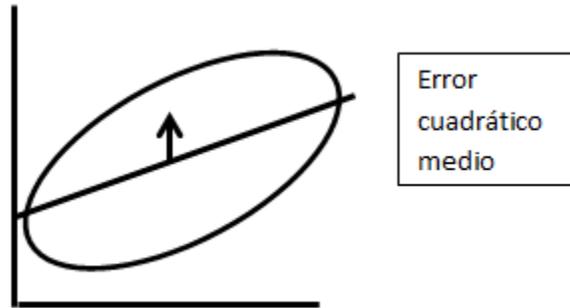


Figura 23 Error cuadrático medio.

Capítulo IV

Desarrollo

Este capítulo abarca el desarrollo de la metodología propuesta en cada una de sus etapas, con el fin de describir adecuadamente cada una de ellas. Antes de abordar el desarrollo es importante dar a conocer el caso de estudio que justificó este trabajo, el cual requirió el registro de 35 parejas de imágenes de retina, en la que cada par está compuesto por una imagen óptica de retina y una imagen fluoresceínica de la misma retina (figura 24); para comprender este caso de estudio se deberá considerar lo siguiente:

Los microaneurismas son pequeñas gotas de sangre al final de las vénulas y arteriolas de la retina, generadas por la debilidad del endotelio. Por su tamaño, los microaneurismas son prácticamente imposibles de detectar en la imagen óptica en colores del ojo de una persona al inicio de la retinopatía diabética, enfermedad producida por malos cuidados de la diabetes mellitus tipo dos. Sin embargo, la presencia de microaneurismas es el primer síntoma de la retinopatía diabética en la retina humana; cuando han tomado mayor tamaño, estos se pueden observar en una imagen óptica en colores como pequeños puntos oscuros en los extremos de las vénulas y arteriolas en la retina de una persona enferma de diabetes mellitus tipo 2.

La presencia de microaneurismas en la zona alrededor de la mácula lútea es una señal de alerta para los médicos en el sentido de tomar oportunamente las medidas necesarias, ya que normalmente después de la presencia de microaneurismas y derrames sobre la retina comienzan a aparecer exudados duros, los cuales son depósitos de lípidos que debe evitarse su presencia cerca de la mácula, ya que a medida que proliferan pueden causar ceguera a los pacientes diabéticos que no han cuidado adecuadamente su padecimiento.

Con la finalidad de poder hacer visibles los microaneurismas durante exámenes a personas propensas a padecer retinopatía diabética o a personas enfermas de diabetes mellitus tipo 2, se les inyecta intravenosamente con 5 ml de fluoresceína sódica al 10-20% a través de un catéter. Después de 10 - 20 segundos el colorante aparece en el ojo fluyendo lentamente a través de las arterias coroideas; de esta forma es posible capturar unas secuencias de imágenes para ser estudiadas posteriormente por el oftalmólogo. Esta sustancia hace a la sangre opaca y fluorescente, por lo que la sangre podrá ser visible si se analiza detenidamente la retina; en este caso las venas y arterias de la red vascular se observarán de color blanco y los microaneurismas se podrán observar como pequeños puntos blancos localizados al final de las vénulas y arteriolas (con el endotelio debilitado) ubicadas cerca del centro de la mácula lútea. Al ser un método invasivo, sólo es recomendado a pacientes con diabetes mellitus tipo 2 que sean propensos a padecer la retinopatía diabética.

En la figura 24 se muestran tres imágenes del mismo ojo: (a) muestra una imagen óptica en colores; (b) muestra la imagen de la misma retina pero en escala de grises y finalmente en (c) se muestra la misma retina de una persona enferma de retinopatía diabética después de ser inyectada con la sustancia fluoresceínica. Como se puede observar, en (a) y (b) los microaneurismas no se pueden apreciar, la red vascular y la mácula lútea aparecen

oscuras, mientras que la papila óptica aparece como un disco claro. Los objetos claros en las imágenes (a), (b) son exudados duros, el segundo síntoma de la retinopatía diabética. Por otro lado, en la figura (c) se muestra la imagen fluoresceínica del mismo ojo, donde los microaneurismas aparecen como pequeños puntos blancos, la red vascular y la papila óptica aparecen claras, mientras que la mácula lútea continúa viéndose oscura. En resumen, las imágenes (b) y (c) presentan diferentes características radiométricas.

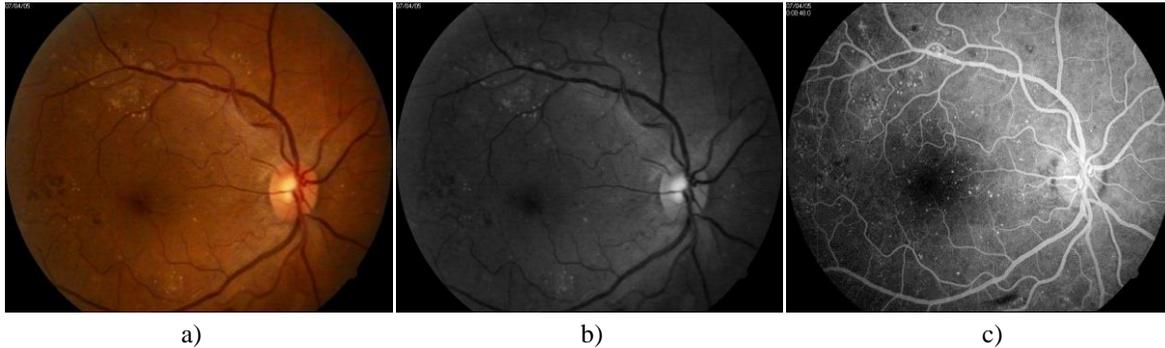


Figura 24 a) Imagen óptica, b) Imagen en escala de grises, c) Imagen fluoresceínica.

El registro de imágenes de retinas capturadas a partir de la misma fuente, pero con diferentes características radiométricas, es el objetivo de este trabajo, lo cual es necesario para poder realizar un conteo de los microaneurismas que se intentan detectar sobre imágenes ópticas de retina de pacientes con retinopatía diabética; este es un trabajo que se encuentra en proceso. Hasta ahora, el procedimiento de conteo se ha realizado a mano, pero si se consigue el registro de imágenes de retina ópticas y fluoresceínicas correspondientes, el proceso podría plantearse en un futuro de forma automática.

El reto se encuentra en poder registrar imágenes que a simple vista son diferentes partiendo del supuesto de que ambas imágenes poseen ciertas características en común. Para poder conseguirlo se propone la siguiente metodología que sigue el enfoque del registro de imágenes basadas en SIFT, (sección 3.2).

4.1 Metodología del registro

En la literatura observamos variadas metodologías del registro de imágenes dependiendo del problema que se esté abordando; en nuestro caso nos basaremos en el descriptor SIFT como se había mencionado anteriormente, el que de manera genérica consta de las cuatro etapas siguientes:

- Selección de características.
- Correspondencia de características.
- Ajuste de características.
- Transformación.

Durante el desarrollo del trabajo se implementará cada etapa en su orden después de una etapa previa de preprocesamiento, la cual se agrega para acercar las imágenes óptica y fluoresceínica a un nivel en que sean más similares, es decir, que se parezcan lo suficiente para facilitar la realización del registro y así obtener los resultados esperados.

Es importante destacar que el registro de imágenes en general se puede llevar a cabo sin considerar el módulo de preprocesamiento, tomando tan solo en cuenta ciertas consideraciones que se presentarán en el capítulo 5.

La inclusión de la primera etapa de preprocesamiento constituye un aporte de nuestro trabajo, ya que en nuestro caso sin esta etapa previa no era posible realizar el registro de las imágenes ópticas y fluoresceínicas.

El siguiente diagrama de bloques de la figura 25 describe el proceso de registro para conseguir el objetivo deseado.

4.1.1 Preprocesamiento

El módulo de preprocesamiento, de forma general, tiene la finalidad de mejorar y modificar independientemente las imágenes a registrar, con el fin de incrementar la semejanza de ambas y con ello la posibilidad de lograr un registro exitoso.

En el caso de las imágenes ópticas de retina, el preprocesamiento consiste en los siguientes pasos:

- Aplicar el filtro mediana para reducir *outliers* (pixeles con niveles de gris extremos muy diferentes a los de los pixeles que los rodean).
- Aplicar la ecualización de su histograma.
- Realizar la división de canales
- Seleccionar el canal verde.

La figura 26 muestra el diagrama de bloques que representa este proceso paso a paso.

Para el caso de las imágenes fluoresceínicas de retina, el preprocesamiento consiste en los siguientes pasos:

- Aplicar el filtro mediana.
- Obtener el negativo de la imagen.

La figura 27 muestra el diagrama de bloques que describe este proceso paso a paso.

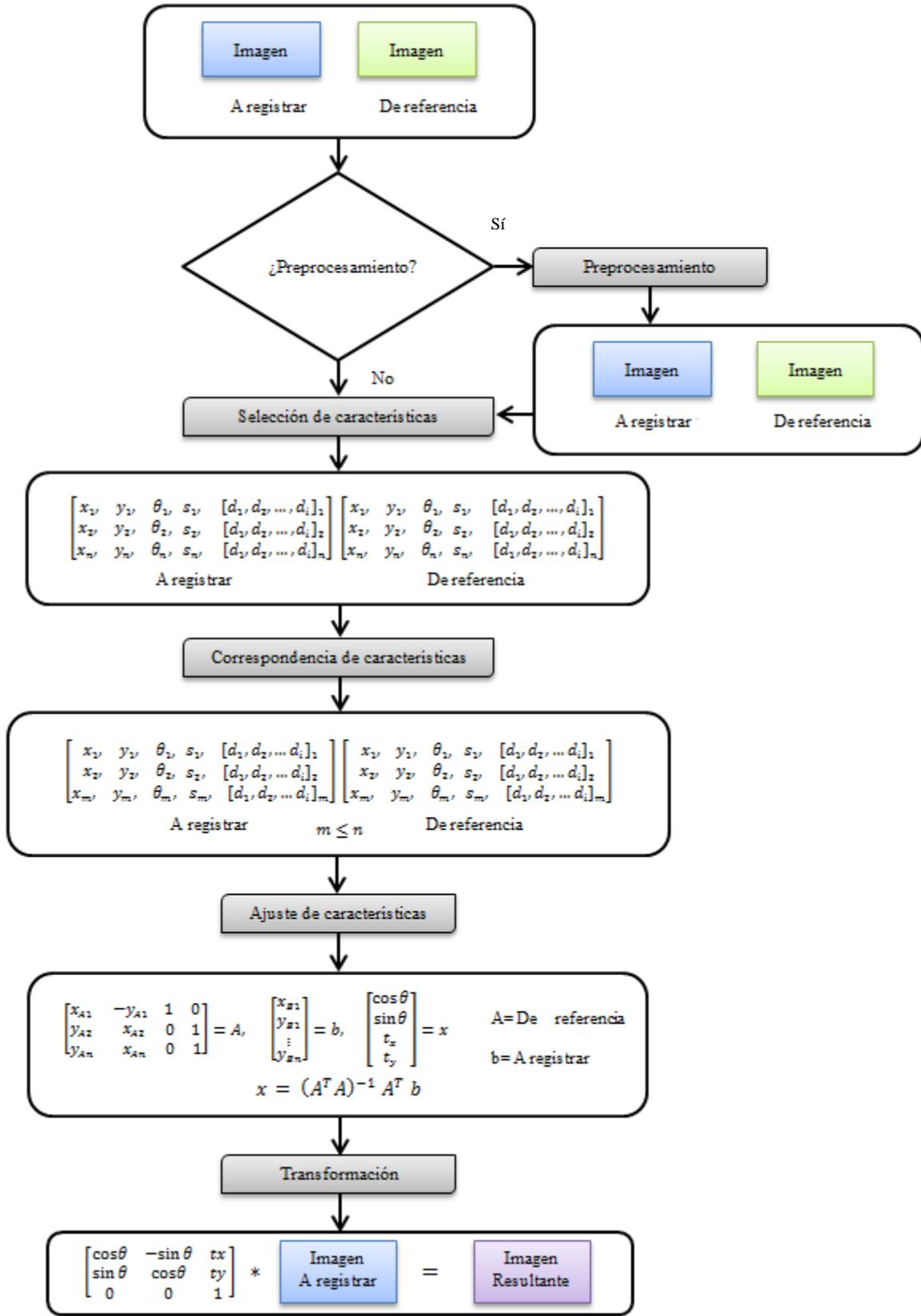


Figura 25 Proceso de registro de imágenes.

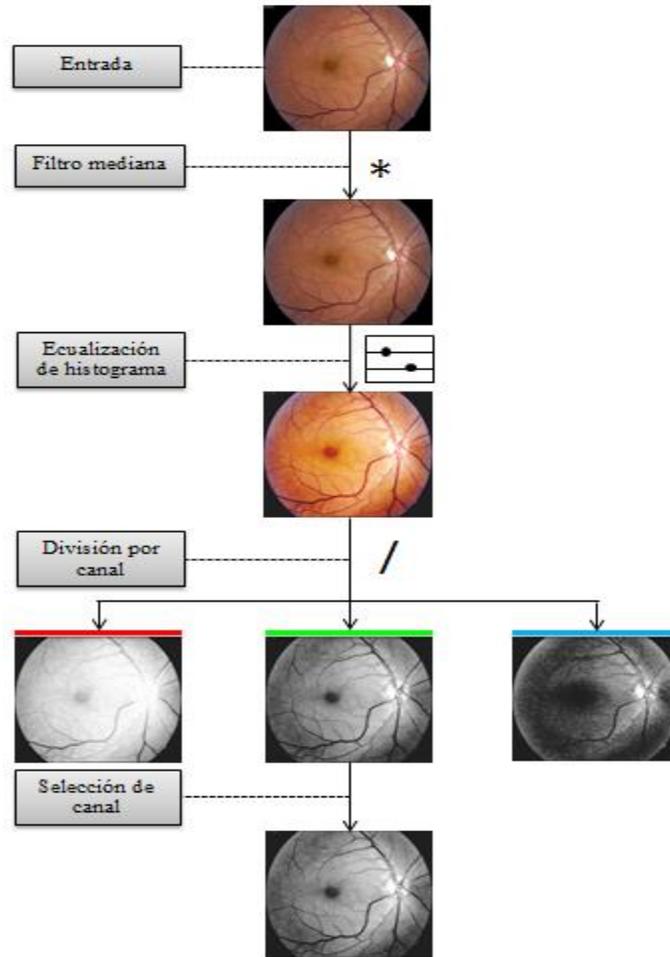


Figura 26 Diagrama de bloques del preprocesamiento de las imágenes ópticas.

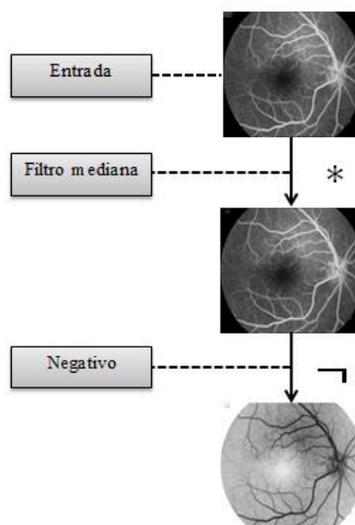


Figura 27 Diagrama de bloques del preprocesamiento de las imágenes fluorescéticas.

A continuación se detalla cada una de las etapas restantes de este desarrollo.

4.1.2 Selección de características

El proceso de selección de características tiene la finalidad de localizar en la imagen aquellas características o puntos clave con los que se intentará hacer corresponder ambas imágenes. Este proceso se puede realizar de diversas formas, pero para este trabajo se optó por usar una transformación de características invariante a la escala (*Scale Invariant Feature Transform*) conocido como SIFT por su sigla en inglés; el proceso que realiza SIFT consiste en buscar puntos clave sobre un espacio de escala generado por la función Diferencia de Gaussianos. Los pasos para obtener los puntos clave y sus descriptores son los siguientes:

- Creación del espacio de escala y detección de los puntos clave.
- Asignación de coordenadas, ángulo y escala.
- Definición de descriptores.

A continuación se describirá e ilustrará cada uno de estos pasos. Si se desea conocer con mayor detalle alguno de estos pasos, se sugiere revisar el apartado del capítulo tres donde se expone información detallada sobre la transformación SIFT (sección 3.2).

4.1.2.1 Creación del espacio de escala y detección de los puntos clave

El espacio de escala se obtiene bajo un enfoque de filtrado en cascada mediante la función Diferencia de Gaussianos. En este paso la imagen de entrada será convolucionada n veces con un filtro Gaussiano para distintos valores de la desviación estándar; como resultado de esto se obtendrán n imágenes del Gaussiano, de las cuales se obtendrán las diferencias de pares consecutivos. Este proceso se muestra esquemáticamente en la figura 28. De lo anterior, podremos concluir que lo que se desea es obtener un conjunto de imágenes borrosas que representen la variación de la escala de una misma imagen (figura 29). Este procedimiento, como los que siguen, se le realiza tanto a la imagen óptica como a la fluorescénica preprocesadas.

En la figura 30 se ejemplifica de manera gráfica este proceso sobre las correspondientes imágenes ópticas de retina.

La detección de los puntos clave sobre el espacio de escala se realiza identificando máximos y mínimos mediante la comparación de cada pixel con sus 26 vecinos (sus 9 vecinos de la escala superior, los 8 vecinos de su escala y los 9 vecinos de la escala inferior) pertenecientes a la escala superior, a la escala actual y a la escala inferior, respectivamente, lo cual se puede observar en la figura 30. El pixel bajo análisis será

considerado como punto clave si y solo si es mayor o menor que sus correspondientes ubicados en las escalas vecinas.

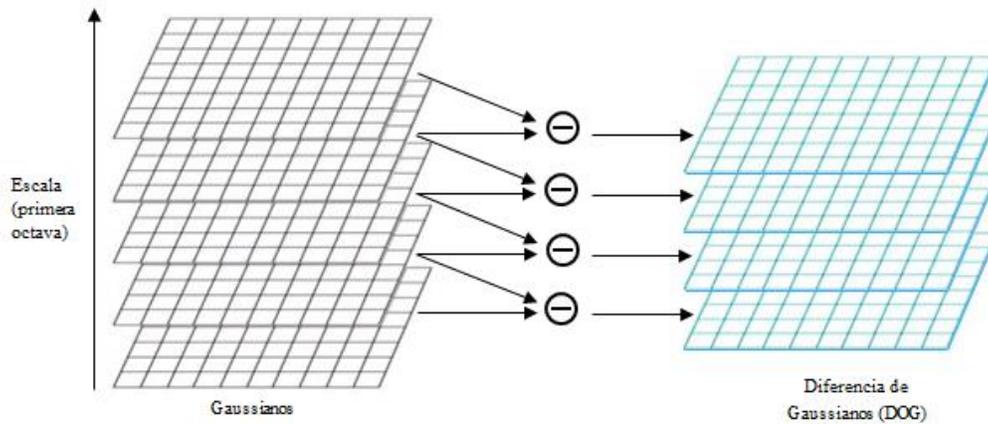


Figura 28 Obtención de la función diferencia de Gaussianos.

En la figura 30 se muestran los puntos clave detectados; la imagen de retina contiene círculos de color verde de diferentes tamaños, lo cual indica la escala en la que se dejó de encontrar, esto es, los círculos pequeños representan características que se dejaron de encontrar en escalas iniciales del espacio de escalas, mientras que los círculos más grandes representan características que se dejaron de encontrar en escalas superiores del espacio de escala.

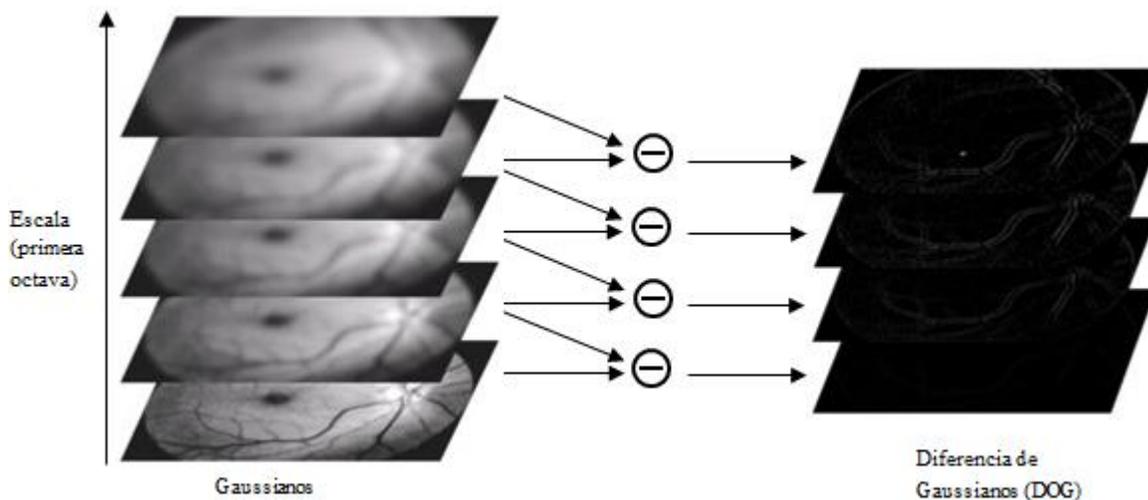


Figura 29 Generación del espacio de escala para una imagen óptica de retina.

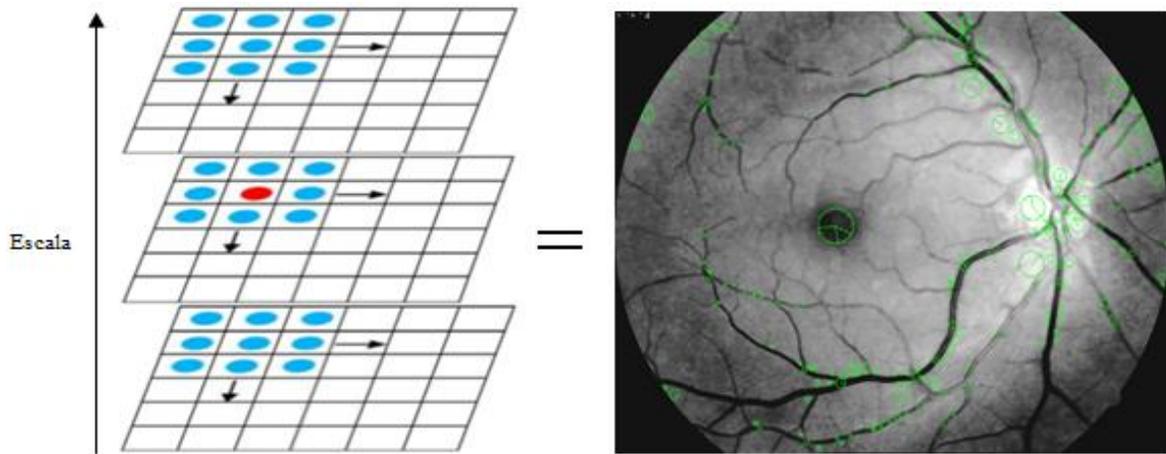


Figura 30 Detección de los puntos clave.

4.1.2.2 Asignación de coordenadas, ángulo y escala

Durante el proceso de detección de puntos clave se almacenan las coordenadas x, y donde fueron localizados; a su vez se identifica y se asigna como el parámetro escala, la última escala en la que el punto clave se mantuvo; el siguiente paso ahora es definir la magnitud y la dirección del gradiente de cada punto clave. Las ecuaciones relacionadas con este proceso se detallaron en el capítulo 3 en el apartado de selección de características (sección 3.2.1). Como resultado de este procedimiento se genera la matriz asociada, que por cada punto clave contendrá sus correspondientes x, y, θ, s (coordenadas, ángulo y escala):

$$\begin{bmatrix} x_1, & y_1, & \theta_1, & s_1 \\ x_2, & y_2, & \theta_2, & s_2 \\ \vdots & \vdots & \vdots & \vdots \\ x_n, & y_n, & \theta_n, & s_n \end{bmatrix}$$

4.1.2.3 Definición de descriptores

En este paso, por cada punto clave, se definirá un descriptor el cual representa el histograma de gradientes de un punto clave en una determinada escala. En la figura 30 se pueden observar las características seleccionadas marcadas en círculos verdes. En la figura 31 se muestran los descriptores como ventanas azules, cuyo tamaño dependerá de la escala de cada punto clave.

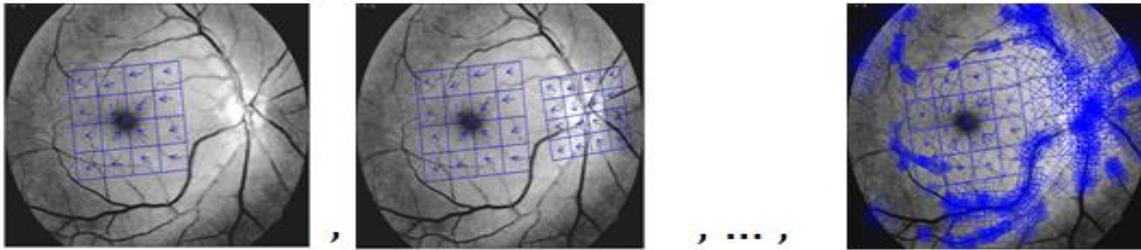


Figura 31 Generación de descriptores.

El resultado final de esta etapa corresponde al conjunto de características compuestas por un x, y, θ, s y un descriptor asociado a los puntos clave como se aprecia en la siguiente matriz.

$$\begin{bmatrix} x_1, & y_1, & \theta_1, & s_1, & [d_1, d_2, \dots, d_i]_1 \\ x_2, & y_2, & \theta_2, & s_2, & [d_1, d_2, \dots, d_i]_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n, & y_n, & \theta_n, & s_n, & [d_1, d_2, \dots, d_i]_n \end{bmatrix}$$

4.1.3 Correspondencia de características

Una vez que se cuenta con las características de la imagen a registrar y de la imagen de referencia, el siguiente paso es hacerlas corresponder, para lo que se realizan dos procedimientos de discriminación: el primero utiliza la distancia Euclidiana como medida de similitud y el algoritmo árbol kd ($k-d$ tree); el segundo mediante un procedimiento de votación usando el algoritmo de la transformada de Hough.

El primer procedimiento consiste en medir los descriptores asociados con los puntos clave, bajo el supuesto de que los descriptores poseen información relevante que está presente alrededor de un punto clave. El procedimiento para poder discriminar características y hacer corresponder otras, se basa en el siguiente algoritmo descrito en pseudocódigo (figura 32):

Algoritmo: El vecino más cercano en un árbol kd

Entrada: **kd** un árbol kd
objetivo \in dominio
hr un hiperrectángulo
distanciam flotante

Salida: **mascercano** tipo **nodo** ejemplar
distancia flotante

Precondición: elArbolEsLegal(**kd**)

Poscondición: el **mascercano** será el ejemplar más cercano al **objetivo**, ambos se encontrarán dentro del hiperrectángulo **hr** y dentro de una **distanciam** del **objetivo** siendo la **distancia** del punto más cercano

- 1: *Si kd* está vacío
 Entonces
 distancia = ∞
 Si No
- 2: s = divide campo de *kd*
- 3: **pivote** = dominio del campo *kd*
- 4: cortar **hr** en **hr-izquierdo** y **hr-derecho**; el corte sobre el **pivote** es perpendicular a la división s
- 5: **objetivo en izquierda** = $objetivo_s \leq pivote_s$
- 6: *Si objetivo* en izquierda
 Entonces
 cercano-kd = campo izquierdo de *kd*
 cercano-hr = **hr-izquierdo**
 lejano-kd = campo derecho de *kd*
 lejano-hr = **hr-derecho**
- 7: *Si No*
 cercano-kd = campo derecho de *kd*
 cercano-hr = **hr-derecho**
 lejano-kd = campo izquierdo
 lejano-hr = **hr-izquierdo**
- 8: Recursivamente llamar a **vecinoMasCercano()** con los parámetros (**cercano-kd**, **objetivo**, **cercano-hr**, **distanciam**) y guardando el resultado en **mascercano** y **distancia**
- 9: **distanciam** = **min (distanciam, distancia)**
- 10: Un punto cercano solo podrá estar en **lejano-kd** si estaba en alguna parte de **lejano-hr** dentro de una distancia **distanciam**
 Si (*puntocercano* \in **lejano – kd** & **lejano – hr**) dentro de una **distanciam**
 Entonces
 Si ($(pivote - objetivo)^2 < distancia$)
 Entonces
 mascercano = **pivote**, **rangodelcampo kd**
 distancia = $(pivote - objetivo)^2$
 Recursivamente llamar a **vecinoMasCercano()** con

(*cercano* – *kd, objetivo, cercano* – *hr, distanciam*)
 almacenar resultado *mascercano* & *distancia*
 Si (*distanciatemp* < *distancia*)
 Entonces
 mascercano = *mascercano temp*
 distancia = *distancia temp*

Figura 32 Algoritmo del vecino más cercano mediante un árbol *kd*.

La finalidad del algoritmo anterior es identificar el descriptor más parecido, en un principio con base en sus coordenadas y en segundo lugar con base en los valores del descriptor.

Para ejemplificar esto consideremos lo siguiente: tenemos un conjunto de datos de entrada representados por pares (x, y) , de los que se conforma el correspondiente árbol *k-2* y se observa la división del plano respecto a la formación del árbol (figura 34); para esto último se debe tener en cuenta lo siguiente:

- El primer valor (x, y) de entrada será el pivote para la construcción del árbol; cada nivel del árbol se dividirá en planos x o y , comenzando con el pivote en x y el siguiente nivel en y , alternando cada nivel del árbol con estos planos.
- Los siguientes nodos de entrada (x, y) se ubicarán en el árbol comparando su correspondiente valor respecto a x o y , en dependencia del nivel (x o y) en que se encuentre para su ubicación en el árbol. En la figura 33 se muestra un ejemplo de la construcción de un árbol *k-2* en el que tenemos como pivote a $(3,7)$ perteneciente al plano x ; el siguiente valor $(8,1)$ comienza su comparación en el pivote respecto a x ; como 8 es mayor que 3, se coloca en la ramificación derecha; el siguiente nodo de entrada es $(6,6)$ y comienza su comparación en el pivote respecto a x ; como 6 es mayor que 3, pasa a la ramificación derecha y se encuentra al nodo $(6,6)$ con el que se tendrá que comparar respecto a y ; como 6 es mayor que 1 se ubica en la ramificación derecha; en la figura 33 se puede observar el árbol completo.
- Cada vez que se inserta un nodo en el árbol se aplicará el mismo criterio de comparación antes de ser ubicado en el árbol.

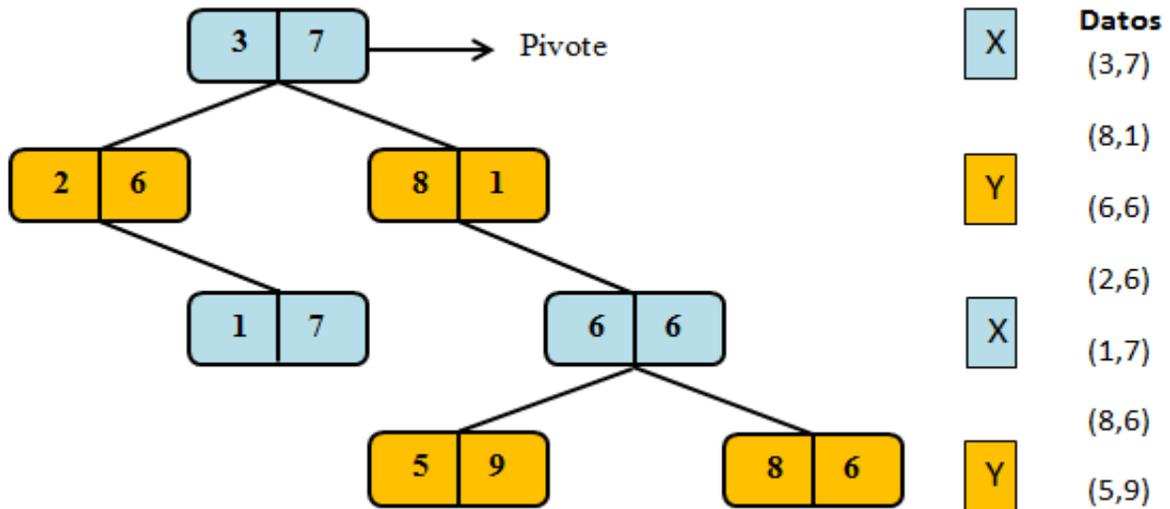


Figura 33 Construcción de un árbol $k-2$.

A continuación se muestra la gráfica del árbol $k-2$ (figura 34), en el que los nodos ubicados en el plano x del árbol se grafican respecto al eje x (líneas verticales azules) y los nodos ubicados en el plano y se grafican respecto a y (líneas horizontales naranjas).

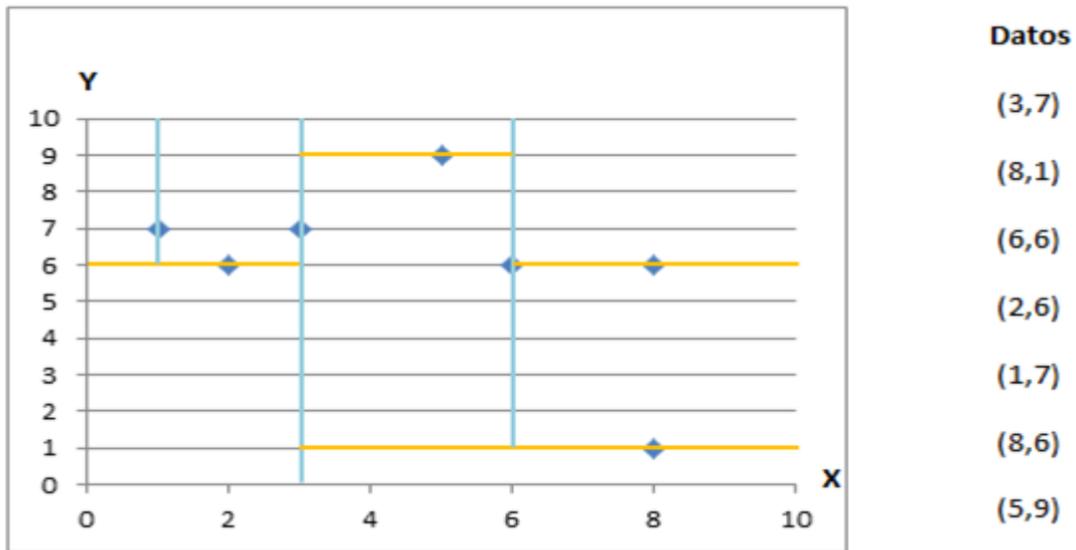


Figura 34 Árbol $k-2$.

El segundo procedimiento consiste en que una vez que se ha discriminado una cantidad considerable de características, es necesario contar con un conjunto consistente de ellas. Un conjunto consistente representa aquellas características que están presentes en ambas imágenes; para poder conseguir esto se hace uso de la transformada de Hough.

La transformada de Hough es usada para localizar cualquier forma que pueda ser definida paramétricamente con una distribución de puntos.

El problema de encontrar un objeto (descrito a través de un modelo) en una imagen puede ser resuelto encontrando su posición en la imagen; con la transformada de Hough generalizada el problema de encontrar la posición del modelo es transformado en un problema de encontrar los parámetros de transformación que mapean el modelo dentro de la imagen; al momento de conocer tales parámetros, la posición del modelo puede ser determinada. Para poder obtener la generalización se considera la siguiente representación $r = \{x, y, \theta, s\}$, donde x, y es la referencia al origen, θ es su orientación y $s = (sx, sy)$ describe un factor de escala; con el algoritmo se obtendrá el mejor conjunto de parámetros para una forma dada.

Se construirá una tabla r la cual se compondrá de las posibles x, y, θ, s de una imagen dada; cada punto en la imagen de prueba y sus propiedades se buscará en la tabla r y si su correspondiente valor es recuperado, la celda apropiada del acumulador a es incrementada; la celda con mayor cantidad de votos en la matriz acumulador a serán los puntos existentes que asocian el objeto en la imagen de referencia.

En la figura 35 se presenta en pseudocódigo el algoritmo de la transformada de Hough generalizada para corresponder un objeto arbitrario dentro de una imagen (figura 36).

Algoritmo: Transformada de Hough generalizada

Entrada: **cr** Características de imagen de referencia representadas por $\{x, y, a, s\}$,
cp Características de imagen de prueba representadas por $\{s, y, a, s\}$,
r tabla de posibles valores con respecto a **cr**,
a acumulador

Salida: **pco** puntos que conforman al objeto

- 1:** Construcción de la tabla **r** con base en **cr**
- 2:** Iniciar **a**
- 3:** Para cada elemento **cp**
- 4:** Buscar su correspondiente valor en **r**
- 5:** Si el valor se corresponde
- 6:** Incrementar **a**
- 7:** **pco** igual a valor máximo de **a**

Figura 35 Algoritmo de la transformada de Hough generalizada.

Para aclarar lo anterior, a continuación se presenta un ejemplo de detección de líneas en una imagen; se usa este ejemplo por cuestiones de legibilidad, ya que presentar un ejemplo completo con el algoritmo generalizado implicaría colocar una gran cantidad de figuras que puede confundir más que aclarar la idea general de este procedimiento.

Se debe de considerar lo siguiente:

- Espacio de la imagen: conjunto de coordenadas (x, y) de la imagen.
- Espacio de Hough: conjunto de parámetros (a, d) que representan una línea recta.

En el espacio de la imagen, la recta se puede representar con la ecuación $y = m * x + n$ y se puede graficar para cada par (x, y) de la imagen.

En la transformada de Hough, la idea principal es considerar las características de una recta en término de sus parámetros (m, n) , y no como puntos de la imagen $(x_1, y_1), \dots, (x_n, y_n)$.

Basándose en lo anterior, la recta $y = m * x + n$ se puede representar como un punto (m, n) en el espacio de parámetros (pendiente e intercepto).

Sin embargo, cuando se tienen rectas verticales, los parámetros de la recta (m, n) no se definen. Por esta razón es mejor usar los parámetros que describen una recta en coordenadas polares, denotadas (ρ, θ) .

El parámetro ρ representa la distancia entre el origen de coordenadas y el punto (x, y) , mientras que θ es el ángulo del vector director de la recta perpendicular a la recta original y que pasa por el origen de coordenadas.

Usando esta parametrización, la ecuación de una recta se puede escribir de la siguiente forma: $y = \left(-\frac{\cos \theta}{\sin \theta}\right) * x + \left(\frac{\rho}{\sin \theta}\right)$, la que se puede reescribir como $\rho = x * \cos \theta + y * \sin \theta$.

Entonces, es posible asociar a cada recta un par (ρ, θ) que es único si $\theta \in [0, \pi)$ y $\rho \in \mathbf{R}$ o $\theta \in [0, 2 * \pi)$ y $\rho \geq 0$. El espacio (ρ, θ) se denomina espacio de Hough para el conjunto de rectas en dos dimensiones.

Para un punto arbitrario en la imagen con coordenadas (x_0, y_0) , las rectas que pasan por ese punto son los pares (ρ, θ) con $r = x * \cos \theta + y * \sin \theta$ donde ρ (la distancia entre la línea y el origen) está determinado por θ . Esto corresponde a una curva sinusoidal en el espacio (ρ, θ) , que es única para ese punto. Si las curvas correspondientes a dos puntos se intersecan, el punto de intersección en el espacio de Hough corresponde a una línea en el espacio de la imagen que pasa por estos dos puntos. Generalizando, un conjunto de puntos que forman una recta, producirán sinusoides que se intersecan en los parámetros de esa línea. Por tanto, el problema de detectar puntos colineales se puede convertir en un problema de buscar curvas concurrentes.

En la figura 36 se muestra en pseudocódigo el algoritmo para detectar rectas en una imagen digital.

Algoritmo: Detección de líneas rectas en una imagen digital

Entrada: **ie** imagen de entrada al acumulador

Salida: **cr** conjunto de rectas cuyos valores son los máximos en el acumulador

- 1: Para cada (x, y) en **ie**
- 2: Si (x, y) pertenece al borde
- 3: Para cada θ posible
- 4: Calcular **p** para (x, y) con θ
- 5: Incrementar la posición (p, θ) en el acumulador
- 6: **cr** igual a posiciones máximas en **a**

Figura 36 Detección de líneas rectas en una imagen digital.

De lo anterior podemos concluir lo siguiente, lo cual se muestra en las tablas y gráficas de la figura 37.

Por cada punto se dibujan un número de líneas que pasan por los mismos lugares, con distintos ángulos. Son las líneas continuas.

Por cada línea se dibuja una recta perpendicular a ésta que pasa por el origen de coordenadas. Son las líneas discontinuas. Se calcula la longitud y el ángulo de cada línea discontinua. Esto se repite para cada punto.

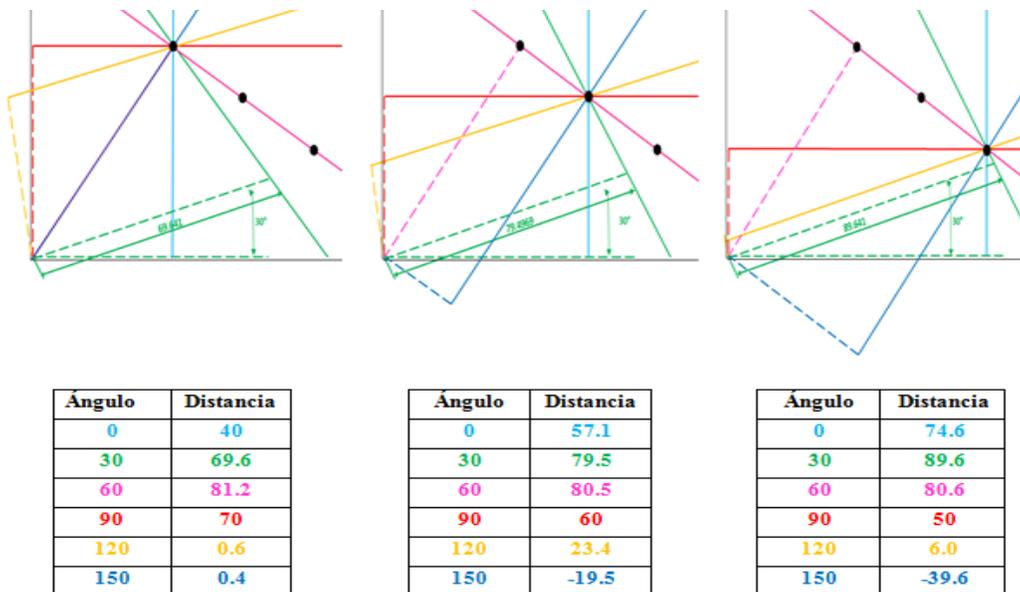
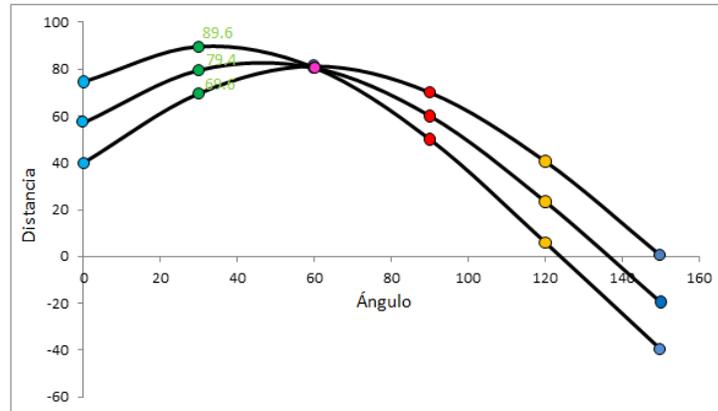


Figura 37 Cálculo del ángulo y la longitud.

A partir de aquí, se crea una gráfica con las longitudes de las líneas por cada ángulo, conocida como gráfica del espacio de Hough (figura 38).



Angulo	Distancia	Distancia	Distancia
0	40	57.1	74.6
30	69.6	79.05	89.6
60	81.2	80.5	80.6
90	70	60	50
120	40.6	23.4	6.0
150	0.4	-19.5	-39.6

Figura 38 Espacio de Hough.

El punto donde las curvas se intersecan en distancia y ángulo en el espacio de Hough, representa los puntos que conforman la línea recta en el espacio de la imagen

4.1.4 Ajuste de características

Una vez que se consiguió obtener un conjunto consistente de características, el siguiente paso es realizar un ajuste de características usando el método de mínimos cuadrados; con ello obtendremos los valores de la matriz de transformación que permite transformar la imagen a registrar.

En esta etapa del desarrollo se busca el ajuste de un modelo a un conjunto de datos. En el estado del arte se observan diversas formas de realizarlo, pero la más prometedora y documentada es con base en los mínimos cuadrados, que se basa en lo siguiente:

Los mínimos cuadrados consisten en un enfoque común para resolver sistemas de ecuaciones sobredeterministas o inexactamente especificados en algún sentido; con los mínimos cuadrados, en lugar de resolver las ecuaciones exactamente, se buscará minimizar la suma de los cuadrados de los residuos. Este enfoque produce lo que se conoce como la estimación de máxima verosimilitud (método habitual para ajustar un modelo y encontrar sus parámetros).

Al abordar este tipo de problemas mediante el uso de mínimos cuadrados se debe considerar lo siguiente:

Sea x una variable independiente y sea $y(x)$ una función desconocida de x la cual se quiere aproximar. Supongamos que tenemos m observaciones:

$$(x_1, y_1), \quad (x_2, y_2), \dots, (x_m, y_m),$$

Donde, $y_i = y(x_i)$, $i = 1, 2, \dots, m$; la idea es modelar $y(x)$ por medio de una combinación de n funciones base $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$, es decir, suponer que la función que se ajusta a los datos es una combinación lineal de la forma:

$$y(x) \approx c_1\phi_1(x) + c_2\phi_2(x) + \dots + c_n\phi_n(x)$$

Donde $m \geq n$. Entonces los datos deben satisfacer de manera aproximada $y(x_i) \approx c_1\phi_1(x_i) + c_2\phi_2(x_i) + \dots + c_n\phi_n(x_i)$, $i = 1, 2, \dots, m$.

La última expresión constituye un sistema de m ecuaciones con n incógnitas, c_1, c_2, \dots, c_n , que en forma matricial puede expresarse de la siguiente manera:

$$\begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \ddots & \phi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \dots & \phi_n(x_m) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, m \geq n.$$

$A \qquad \qquad \qquad x \qquad \qquad \qquad b$

A la matriz de este sistema $A = (a_{ij})$ con $a_{ij} = \phi_j(x_i)$ se le denomina matriz diseño. Dado que $m \geq n$, entonces el sistema tiene más filas (ecuaciones), que columnas (incógnitas). A este tipo de sistemas se les denomina sobredeterministas y generalmente no tiene solución. Las funciones base $\phi_i(x)$, $i = 1, \dots, n$, pueden ser funciones no lineales de x , pero los coeficientes y parámetros c_j aparecen en el modelo en forma lineal cuando se trata de un ajuste lineal, donde $\phi_i(x)$ pueden ser polinomios, funciones racionales, exponenciales, o Gaussianas.

Por lo tanto, para cualquier matriz A de tamaño $m \times n$ y cualquier vector b de tamaño m , existe un solución x de mínimos cuadrados para $Ax = b$. Pero suponiendo que A es un sistema de ecuaciones sobredeterminado (que no tiene solución) se deberá considerar que $b = b_p$, donde b_p es la proyección ortogonal de b sobre el espacio generado por las columnas de A , de tal forma que x será una solución por mínimos cuadrados si y solo si es una solución de ecuaciones normales $A^T A x = A^T b$; si A es invertible, la solución por mínimos cuadrados es única y puede calcularse mediante $x = (A^T A)^{-1} A^T b$.

En este caso:

A : Valores de características de la imagen de referencia

b : Valores de características de la imagen a registrar

x : Valores de la matriz de transformación

En este caso x representa alguna combinación de traslación, rotación y cambio de escala, conocidas como transformaciones afines, las que se pueden expresar en forma matricial de la siguiente forma:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} rs_{xx} & rs_{xy} & trs_x \\ rs_{yx} & rs_{yy} & trs_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Donde los términos rs_{jk} representan el ángulo y el factor de escala en la transformación y el elemento trs_x or y son las coordenadas de traslación.

4.1.5 Transformación

La última etapa que lleva por nombre Transformación consiste en operar la matriz de transformación que se obtuvo en la etapa anterior sobre la imagen a registrar; este paso se puede realizar de forma directa aunque el resultado final podría no ser el deseado.

Para obtener un mejor resultado después de operar la imagen con la matriz de transformación, se realiza un procedimiento de interpolación cuyo objetivo es rellenar la información faltante a partir de datos muestra; el algoritmo de interpolación bilineal es el siguiente:

Supóngase que se quiere encontrar el valor para la función f desconocida en el punto $P = (x, y)$. Conocemos el valor de f en los cuatro puntos $Q_{11} = (x_1, y_1)$, $Q_{12} = (x_1, y_2)$, $Q_{21} = (x_2, y_1)$ y $Q_{22} = (x_2, y_2)$.

Primero se hace una interpolación lineal en la dirección x . Esto genera:

$$f(R_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21})$$

donde:

$$R_1 = (x, y_1,)$$

$$f(R_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22})$$

donde:

$$R_2 = (x, y_2).$$

Después se hace una interpolación en la dirección y :

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2)$$

Esto proporciona una estimación de $f(x, y)$.

$$\begin{aligned} f(x, y) \approx & \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y_2 - y) + \\ & \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y_2 - y) + \\ & \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y - y_1) + \\ & \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y - y_1) + \\ & \frac{1}{(x_2 - x_1)(y_2 - y_1)} \left(\begin{aligned} & f(Q_{11})(x_2 - x)(y_2 - y) + \\ & f(Q_{21})(x - x_1)(y_2 - y) + \\ & f(Q_{12})(x_2 - x)(y - y_1) + \\ & f(Q_{22})(x - x_1)(y - y_1) \end{aligned} \right). \end{aligned}$$

Es importante hacer notar que se obtienen los mismos resultados si la interpolación se hace primero en la dirección y y después en la dirección x . Una vez que se obtiene la matriz de transformación y se opera la imagen a registrar, el resultado deberá ser una imagen que visualmente se parece a la imagen de referencia. Al momento en que cada etapa depende de la anterior, es necesario que para obtener buenos resultados se trate de que cada etapa se ejecute en los mejores términos.

Capítulo V

Experimentos
y resultados

Este capítulo tiene la finalidad de presentar las pruebas con las que se valida el funcionamiento de la metodología de registro propuesta, usando imágenes de retina.

Con el fin de facilitar su comprensión, se detallará el proceso de registro paso a paso, solo en el primer experimento que corresponde a un ejemplo bidireccional, ya que presentar la información de esa forma para cada prueba resultaría poco práctico y en gran medida redundante. Por lo tanto, en el resto de los experimentos se expondrá la información esencial y distintiva, resumiendo al final del capítulo todos los resultados en forma tabular.

Recordando que el proceso de registro de imágenes se compone de cinco etapas, a continuación se enlista lo que se mostrará en cada parte del proceso.

- Preprocesamiento. Se presentarán las imágenes resultantes en cada caso, ya que los pasos en esta etapa dependen del tipo de la imagen a preprocesar, es decir, si es óptica o fluoresceínica.
- Selección de características. Se presentarán imágenes en las que se observa el conjunto de características, las que estarán representadas mediante círculos de un color determinado, de acuerdo al tipo de imagen; cada característica contiene a su vez su coordenada correspondiente, la escala y el ángulo.
- Correspondencia de características. Se presentará una imagen compuesta por la imagen a registrar y la imagen de referencia, asociadas por las características presentes tanto en una como en la otra.
- Ajuste de características. En este punto se presentarán los valores de la matriz de transformación que se obtuvieron después del proceso de mínimos cuadrados y que permitirá transformar la imagen a registrar.
- Transformación. Se presentarán las imágenes resultantes; en particular la imagen en color amarillo representa la unión de la imagen de referencia y la imagen a registrar ya transformada.

La forma en la que se mide el resultado obtenido es mediante el cálculo del error medio cuadrático de los conjuntos de características consistentes; recordemos que estos conjuntos representan las características presentes tanto en la imagen a registrar como en la imagen de referencia. Para mostrar esto se presentará una gráfica compuesta por triángulos verdes que representan las características de la imagen a registrar distribuidas respecto a su posición en la imagen original; asteriscos azules también distribuidos respecto a su posición en la imagen original, y círculos rojos que representan a los triángulos verdes después de ser operados con los valores de la matriz de transformación; se deberá tener presente que mientras más cercanos (o ajustados) se encuentren los círculos rojos de los asteriscos azules, será indicativo de que se ha conseguido una mejor transformación.

La primera prueba que se presenta consiste en registrar un par de imágenes de retina con base en el caso de estudio ya comentado. El objetivo de esta prueba es mostrar cómo se

realiza el proceso de registro de imágenes de retina en ambos sentidos, esto es, tomando como imagen a registrar la imagen óptica y como imagen de referencia a la imagen fluoresceínica y viceversa. Cabe aclarar que este proceso bidireccional no se hará en general con todos los pares de imágenes de retina, sino sólo se presenta para tenerlo como referencia en discusiones posteriores.

Sobre el conjunto de imágenes utilizado en las pruebas

El conjunto de imágenes con las que se trabaja puede ser diferenciado entre imágenes que presentan microaneurismas e imágenes que no presentan microaneurismas; las imágenes fueron tomadas bajo ciertas condiciones particulares. El par de imágenes utilizado en cada prueba son de la misma retina y poseen características geométricas similares. El conjunto de imágenes utilizado para realizar los registros fue proporcionado por especialistas oftalmólogos que han trabajado en investigaciones anteriores realizadas junto con el Dr. Edgardo Manuel Felipe Riverón, tutor de esta tesis.

Sobre la información general para cada prueba

Al momento de realizar el proceso de registro se debe contar con una imagen de referencia y una imagen a registrar, la cual se intenta hacer parecer a la imagen de referencia. Para el caso particular de imágenes de retina ópticas y fluoresceínicas, la imagen de referencia puede ser tanto la imagen óptica como la imagen fluoresceínica; a partir de esto se observa que el registro será viable de forma bidireccional, de fluoresceínica a óptica o viceversa, lo cual se podrá observar en los experimentos detallados más adelante en este trabajo.

Cada prueba contendrá una pequeña descripción, un número de prueba, el tipo de prueba (si pertenece al conjunto de imágenes con microaneurismas o imágenes sin microaneurismas), una categoría (en la que se puede identificar si se consiguió realizar de forma bidireccional, de óptica a fluoresceínica o de fluoresceínica a óptica), además de indicar qué imagen fue la imagen de referencia y cuál la imagen a registrar. A continuación se presentan las pruebas realizadas con pares de imágenes de retina ópticas y fluoresceínicas.

La forma en que se evaluó el registro logrado fue mediante la medición del error cuadrático medio entre el conjunto consistente (intersección de los conjuntos de características) de las características de la imagen de referencia y el conjunto consistente de características de la imagen a registrar, este último utilizado después de ser operado con la matriz de transformación obtenida en el proceso de ajuste.

5.1 Par de imágenes de retina 2_c1

Descripción: Este par de imágenes de retina son del grupo de imágenes sin microaneurismas, lo cual se puede determinar visualmente al observar que las imágenes

fluoresceínicas no presentan puntos blancos en las terminaciones de las vénulas y arteriolas de la red vascular que caracterizan a los microaneurismas.

5.1.1 Experimento 1a

Tipo: Sin microaneurismas

Categoría: o-f (óptica-fluoresceínica)

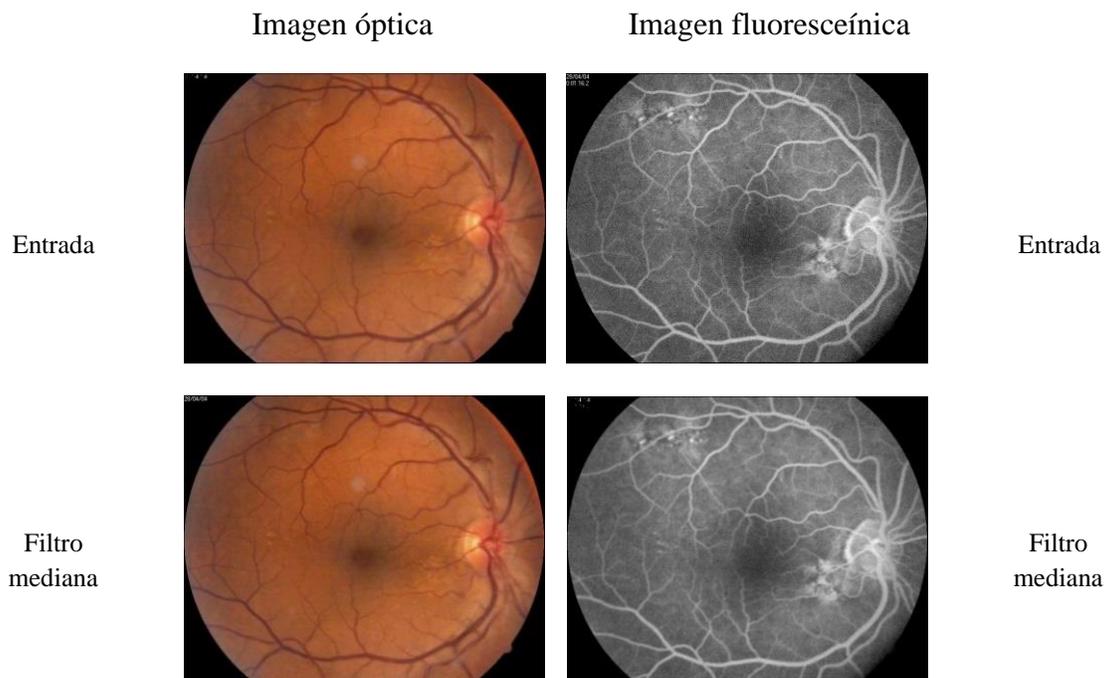
A registrar/nombre: Imagen óptica / 2a_c1

De referencia/nombre: Imagen fluoresceínica / 2b_c1

Etapa de preprocesamiento.

Bajo el supuesto de que las imágenes son de la misma retina, se realizó el siguiente preprocesamiento para intentar llevarlas a un punto en el que ambas sean más similares; con esto último nos referimos a que la retina en sí de ambas imágenes se observe en un tono de gris similar y la red vascular se observe de un color gris oscuro en las dos imágenes.

La figura 39 muestra los resultados que se obtienen al aplicarse la etapa de preprocesamiento sobre las imágenes óptica y fluoresceínica detallados en el capítulo IV (sección 4.1).



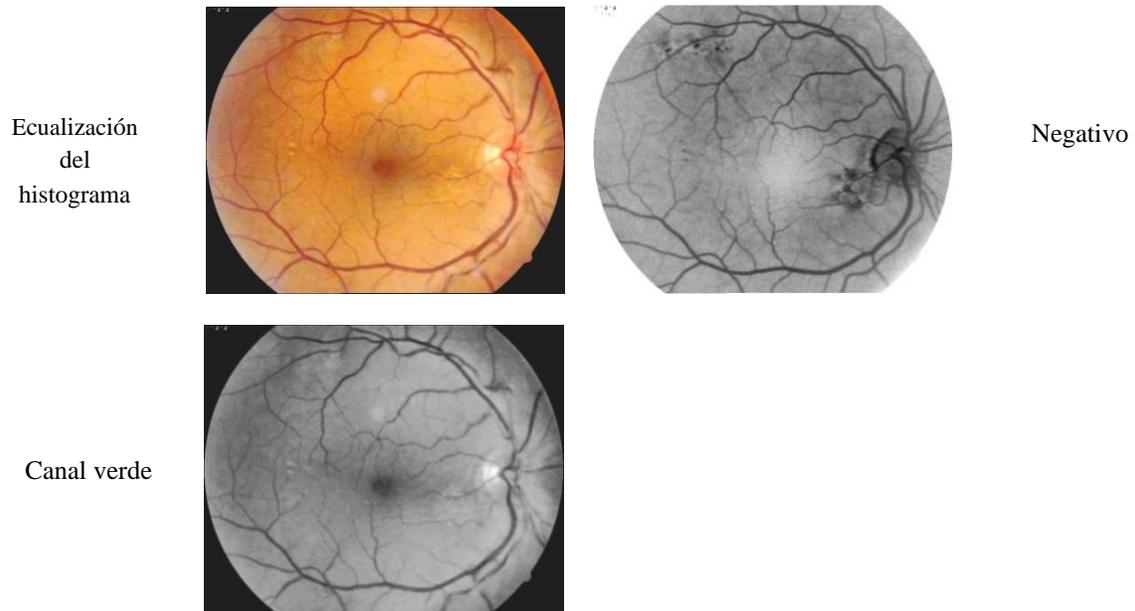


Figura 39 Resultados de la etapa de preprocesamiento.

Etapa de selección de características.

Esta etapa toma como entrada las imágenes resultantes de la etapa anterior. El proceso consiste en detectar los puntos clave sobre un espacio de escala generado por la función Diferencia de Gaussianos. El resultado se puede observar en la figura 40 en la que los puntos de color verde son las características detectadas en la imagen a registrar y los puntos azules representan las características detectadas en la imagen de referencia. Bajo estas imágenes se muestran las matrices que contienen la representación numérica de cada característica, en específico con su correspondiente coordenada (x, y) , ángulo (θ) , escala (s) y descriptor (d_1, d_2, \dots, d_i) .

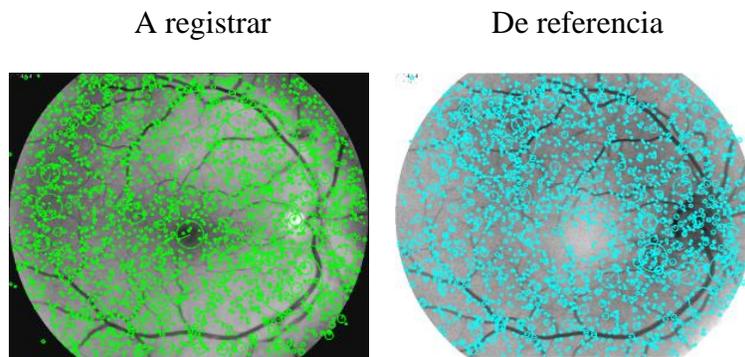


Figura 40 Resultados de la selección de características.

$$\begin{bmatrix} 7.0003, & 240.1772, & 1.9428, & 0.1123, & [39,1,\dots 3] \\ 10.1330, & 354.9647, & 2.1059, & -0.1383, & [48,0,\dots 8] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 494.9855, & 156.5693, & 86.5348, & -3.9231, & [0,0,\dots 8] \end{bmatrix} \begin{bmatrix} 8.0136, & 343.0283, & 2.0717, & -3.1922, & [1,2,\dots 4] \\ 9.1216, & 350.1780, & 1.9305, & -3.3261, & [6,1,\dots 0] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 203.6969, & 131.5683, & 80.3201, & -2.7671, & [10,4,\dots 0] \end{bmatrix}$$

Número de características: 2287

Número de características: 2264

El radio de los círculos depende de la escala en la que se dejaron de encontrar las características, lo cual significa que los círculos pequeños se dejaron de encontrar en las primeras escalas y los más grandes se siguieron encontrando en escalas superiores.

Etapa de correspondencia de características.

En esta etapa se hacen corresponder las características de la imagen a registrar y de la imagen de referencia, mediante dos procedimientos de discriminación: el primero mediante la distancia Euclidiana como medida de similitud (figura 41) y el segundo mediante la transformada de Hough (figura 42).

En esta etapa las características supuestamente correspondientes en cada imagen se unirán por una línea recta con un color seleccionado aleatoriamente, la cual se identificará con un número dado que se muestra siempre en color cian.

Correspondencia Euclidiana. Después de realizar la correspondencia se obtendrá un conjunto de características menor al original; se puede observar en la figura 41 que algunas características no se corresponden adecuadamente (no son relacionadas con rectas aproximadamente horizontales), por lo que para afinar el resultado se deberá aplicar la transformada de Hough. Las matrices que se muestran debajo de cada imagen en la figura 41 contienen las características que en este momento representan las características que se corresponden.

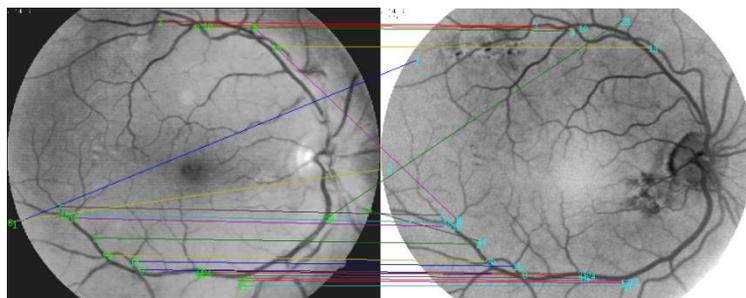


Figura 41 Correspondencia Euclidiana entre las imágenes a registrar y de referencia.

$$\begin{bmatrix} 13.6294, & 427.5871, & 2.0157, & -0.2432, & [0,0,\dots 39] \\ 169.2248, & 454.0475, & 1.9578, & -0.5195, & [41,0,\dots 2] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 377.4225, & 522.1591, & 4.6316, & -4.6303, & [0,0,\dots 21] \end{bmatrix} \begin{bmatrix} 68.1680, & 104.2701, & 2.4321, & -2.3774, & [14,9,\dots 102] \\ 186.8113, & 466.5283, & 2.0159, & -0.5941, & [46,0,\dots 8] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 398.5428, & 528.4573, & 4.9459, & -4.6613, & [43,12,\dots 39] \end{bmatrix}$$

Número de correspondencias: 24

Transformada de Hough. El objetivo de esta transformada es obtener el conjunto consistente de características que esté presente en las dos imágenes, lo cual se logra mediante un mecanismo de votación; esto se puede observar en la figura 42 en la que las características que se corresponden adecuadamente se identifican por líneas de colores; es evidente que el número de características es normalmente menor que el obtenido en el proceso anterior.

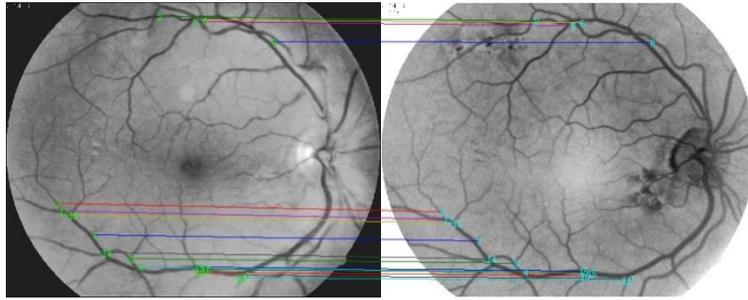


Figura 42 Correspondencia de Hough entre la imagen a registrar y la de referencia.

$$\begin{bmatrix} 169.2248, 454.0475, 1.9578, -0.5195, [41, 0, \dots 2] \\ 294.1991, 28.4813, 1.6904, 1.4156, [2, 1, \dots 110] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 377.4225, 522.1591, 4.6316, -4.6303, [0, 0, \dots 21] \end{bmatrix} \quad \begin{bmatrix} 186.8113, 466.5283, 2.0159, -0.5941, [46, 0, \dots 8] \\ 297.3088, 36.576138, 1.8832, 1.3676, [7, 3, \dots 121] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 398.5428, 528.4573, 4.9459, -4.6613, [43, 12, \dots 39] \end{bmatrix}$$

Número de características: 16

Etapa de ajuste – transformación.

El ajuste que se realiza en esta etapa se lleva a cabo mediante el método de mínimos cuadrados, con el que se obtienen los valores de la matriz de transformación que permite transformar la imagen a registrar para aproximarla a la imagen de referencia. Los valores resultantes de esta operación se muestran a continuación:

$$\text{Ecuación de mínimos cuadrados: } x = (A^T A)^{-1} A^T b$$

Dónde:

A es la imagen de referencia.

b es la imagen a registrar.

x es la matriz de transformación

Los datos numéricos son los siguientes:

$$A = \begin{bmatrix} 169.2248, & -454.0475 & 1 & 0 \\ 294.1992, & 28.4813 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 377.4226, & 522.1591, & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 186.8113 \\ 466.5284 \\ \vdots \\ 528.4573 \end{bmatrix}, \quad x = \begin{bmatrix} 0.9988 & 0.0322 & 2.9152 \\ -0.0332 & 1.0001 & 18.7701 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

En la etapa de transformación se opera la matriz de transformación x sobre la imagen a registrar ya preprocesada para acercarla a la imagen de referencia, el resultado se muestra en la figura 43.

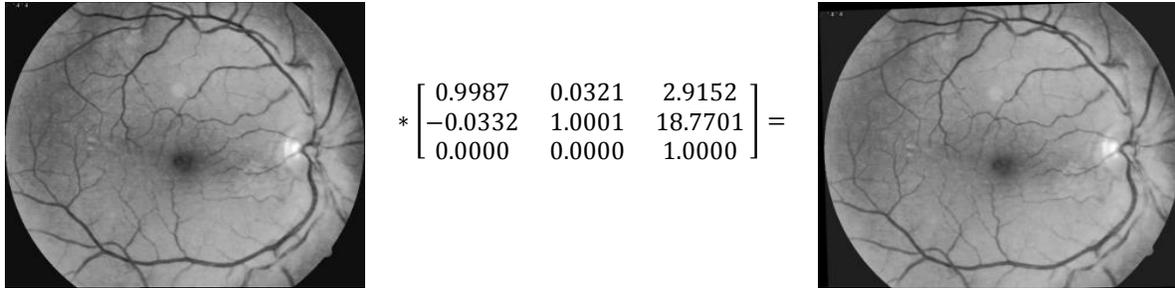


Figura 43 Ajuste y transformación de la imagen a registrar sobre la imagen de referencia.

Resultado

Para finalizar, en la figura 44 se presentan la imagen de referencia y la imagen resultante de la transformación sobrepuesta mediante la unión lógica de las dos (figura 45), en la que se puede observar que la red vascular de ambas aparece ajustada adecuadamente.

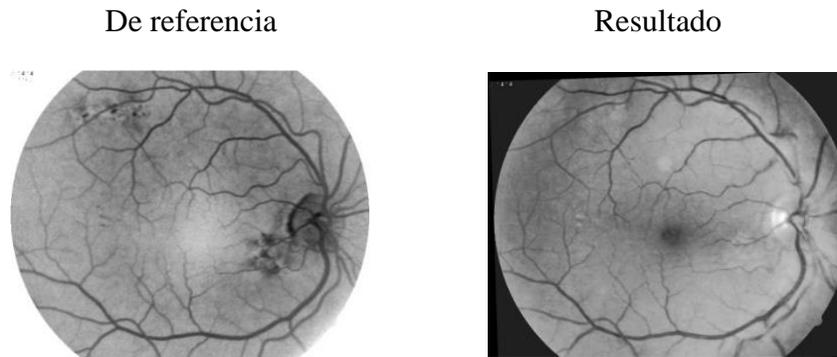


Figura 44 Imagen de referencia e imagen resultante de la transformación.

Unión

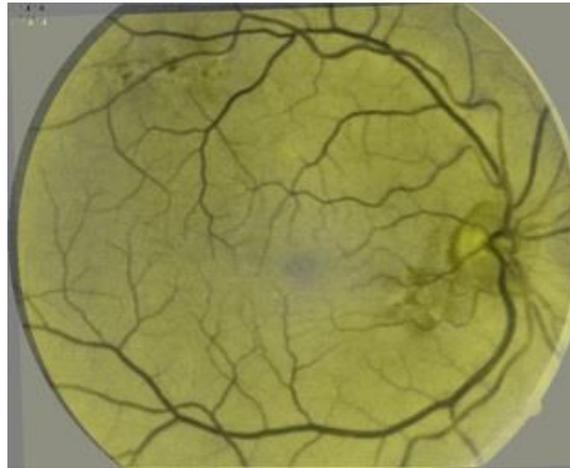


Figura 45 Unión de la imagen a registrar y la imagen de referencia.

Graficación del Error. Para evaluar qué tan bien se llevó a cabo el proceso de registro, se calcula el error medio cuadrático del conjunto consistente de características. En términos del error medio cuadrático y de interpretación, el error que se obtiene nos dice que cada pixel en la imagen registrada está a 1 pixel (0.96888 pixel) de diferencia del valor de referencia (figura 46). En esta figura se observa que las 16 características consistentes originalmente obtenidas (triángulos verdes), una vez transformadas (asteriscos azules) se trasladaron hacia dentro de los círculos rojos que representan las características correspondientes en la imagen de referencia.

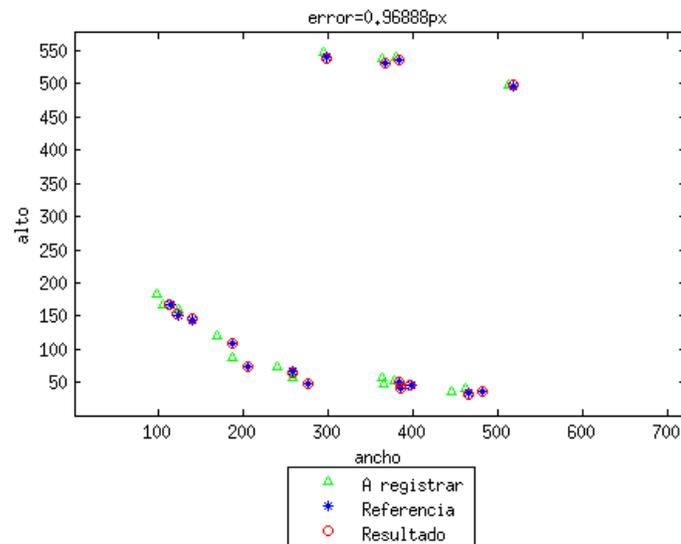


Figura 46 Graficación del error.

En este primer procedimiento se logró el registro de las imágenes considerando como imagen a registrar a la imagen óptica y como imagen de referencia a la imagen fluoresceínica.

A continuación se presenta la segunda parte de esta prueba de forma inversa, usando ahora como imagen a registrar a la imagen fluoresceínica y como imagen de referencia a la imagen óptica. Con esto se demostraría que para este par de imágenes de retina se puede realizar el registro de imágenes de forma bidireccional.

5.1.2 Experimento 1b

Tipo: Sin microaneurismas

Categoría: f-o (fluoresceínica-óptica)

A registrar/nombre: Imagen fluoresceínica / 2b_c1

De referencia/nombre: Imagen óptica / 2a_c1

Etapa de preprocesamiento.

En esta prueba no es necesario replicar el preprocesamiento inicial, ya que se utilizan las mismas imágenes, pero en sentido inverso; ahora la imagen a registrar es el negativo de la imagen fluoresceínica y la imagen de referencia es la imagen del canal verde de la imagen óptica (figura 47).

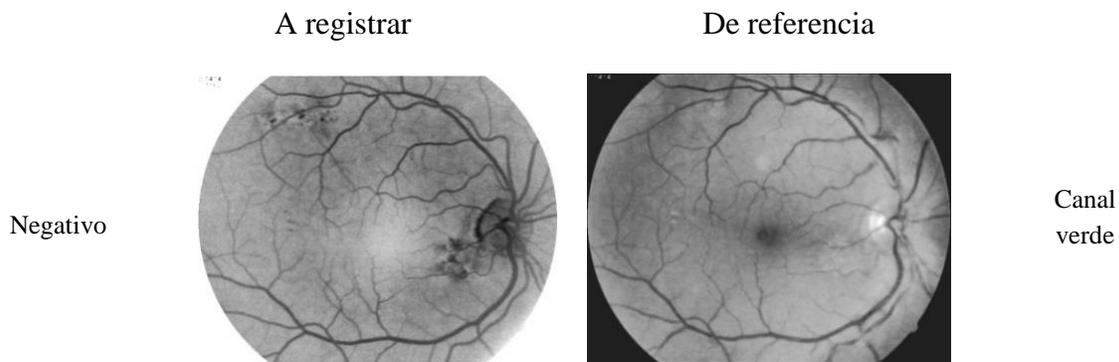


Figura 47 Imagen fluoresceínica a registrar (negativo) y la imagen óptica de referencia (canal verde).

Etapa de selección de características.

En la figura 48 se presentan ambas imágenes con sus correspondientes características identificadas y las correspondientes matrices asociadas a cada imagen que se utilizarán en las etapas posteriores.

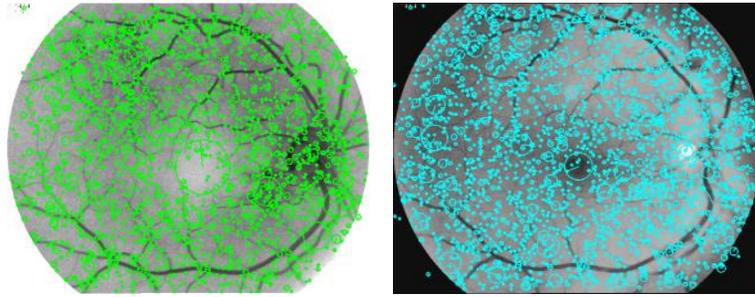


Figura 48 Selección de características de la imagen a registrar y la imagen de referencia.

$$\begin{bmatrix} 8.0136, & 343.0283, & 2.0717, & -3.1922, & [1, 2, \dots 4] \\ 9.1216, & 350.1780, & 1.9305, & -3.3261, & [6, 1, \dots 0] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 203.6969, & 131.5683, & 80.3201, & -2.7671, & [10, 4, \dots 0] \end{bmatrix} \quad \begin{bmatrix} 7.0003, & 240.1772, & 1.9428, & 0.1123, & [39, 1, \dots 3] \\ 10.1330, & 354.9647, & 2.1059, & -0.1383, & [48, 0, \dots 8] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 494.9855, & 156.5693, & 86.5348, & -3.9231, & [0, 0, \dots 8] \end{bmatrix}$$

Numero de características: 2264

Numero de características: 2287

Los puntos de color verde son las características detectadas en la imagen a registrar (la fluorescénica) y los puntos azules representan las características detectadas en la imagen de referencia (la óptica).

Etapa de correspondencia de características.

Correspondencia Euclidiana. En la figura 49 se muestran las correspondencias con líneas de colores y un número de identificación. A simple vista se observa que varias de ellas no se corresponden adecuadamente, por lo que será necesario aplicar la transformada de Hough para la obtención de un conjunto consistente.

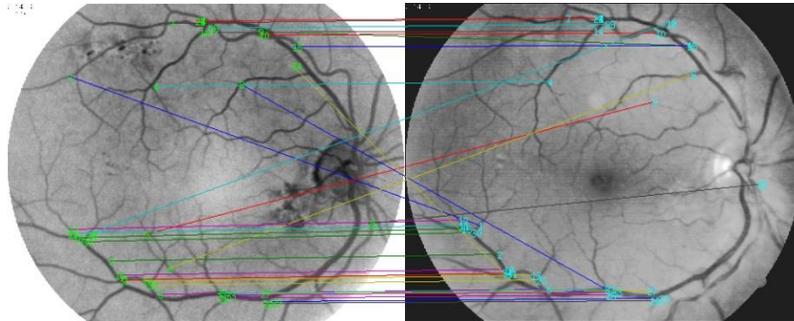


Figura 49 Correspondencia Euclidiana.

$$\begin{bmatrix} 113.9445, & 132.9085, & 1.7948, & -2.0011, & [0, 0, \dots 48] \\ 186.8113, & 466.5283, & 2.0159, & -0.5941, & [46, 0, \dots 8] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 655.5866, & 398.2567, & 27.1400, & 0.5500, & [5, 0, \dots 0] \end{bmatrix} \quad \begin{bmatrix} 134.4385, & 401.6919, & 2.2583, & 0.5840, & [10, 10, \dots 46] \\ 169.2248, & 454.0475, & 1.9578, & -0.5195, & [41, 0, \dots 2] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 638.5744, & 327.6811, & 2.3746, & -0.4059, & [50, 5, \dots 0] \end{bmatrix}$$

Número de correspondencias: 35

Transformada de Hough. Al realizar la transformada de Hough se discrimina una cantidad de características, teniendo ya como resultado un conjunto consistente (figura 50). En esta prueba se detectaron 21 características, cinco más que en la prueba de registro anterior.

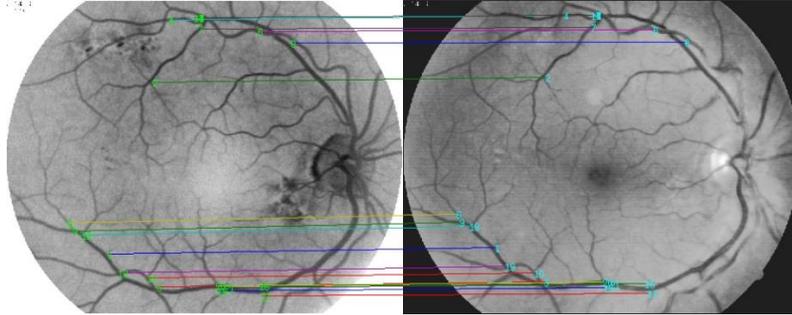


Figura 50 Correspondencia de Hough.

$$\begin{bmatrix} 186.8113, & 466.5283, & 2.0159, & -0.5941, & [46,0,\dots 48] \\ 268.4075, & 150.6906, & 2.1449, & 0.5148, & [2,2,\dots 44] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 398.5428, & 528.4573, & 4.9459, & -4.661380, & [43,12,\dots 39] \end{bmatrix} \begin{bmatrix} 169.2248, & 454.0475, & 1.9578, & -0.5195, & [10,10,\dots 46] \\ 260.6397, & 142.4797, & 2.0139, & 0.4779, & [41,0,\dots 2] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 377.4225, & 522.1591, & 4.6316, & -4.6303, & [50,5,\dots 0] \end{bmatrix}$$

Número de características: 21

Etapas de ajuste – transformación.

El ajuste que se realiza en esta etapa se lleva a cabo mediante el método de mínimos cuadrados, con el que se obtienen los valores de la matriz de transformación que permita transformar la imagen a registrar para aproximarla a la imagen de referencia. Los valores resultantes de esta operación se muestran a continuación:

$$x = (A^T A)^{-1} A^T b$$

Donde:

A es la imagen de referencia.

b es la imagen a registrar.

x es la matriz de transformación.

Los datos numéricos son los siguientes:

$$A = \begin{bmatrix} 186.8113, & 466.5284 & 0 & 1 \\ 268.4076, & 150.6907 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 398.5428, & 528.4573, & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 169.2248 \\ 454.0475 \\ \vdots \\ 522.1591 \end{bmatrix}, \quad x = \begin{bmatrix} 0.9992 & -0.0345 & -1.0544 \\ 0.0331 & 0.9962 & -17.4457 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

En la etapa de transformación se opera la matriz resultante x sobre la imagen a registrar ya preprocesada; el resultado se muestra en la figura 51.

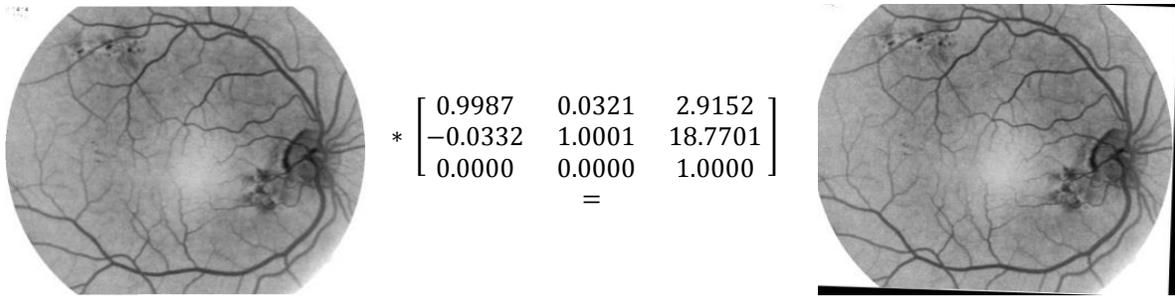


Figura 51 Resultado de aplicar la matriz de transformación hallada sobre la imagen a registrar.

Resultado

Para finalizar, en la figura 52 se presentan sobrepuestas la imagen de referencia y la imagen resultante de la transformación mediante la unión lógica de las dos (figura 53), en la que se puede observar que la red vascular de ambas aparece ajustada adecuadamente.

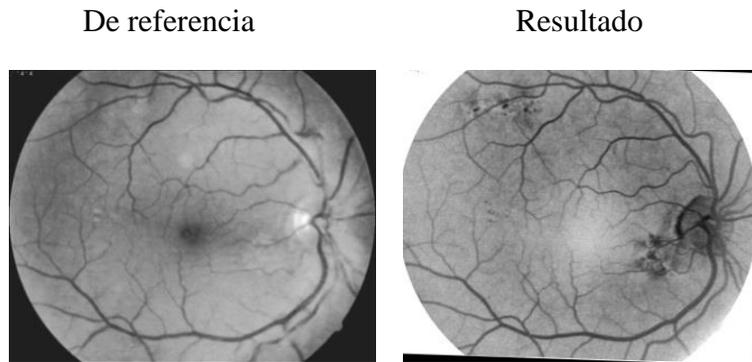


Figura 52 Imagen de referencia e imagen resultante de la transformación.

Unión

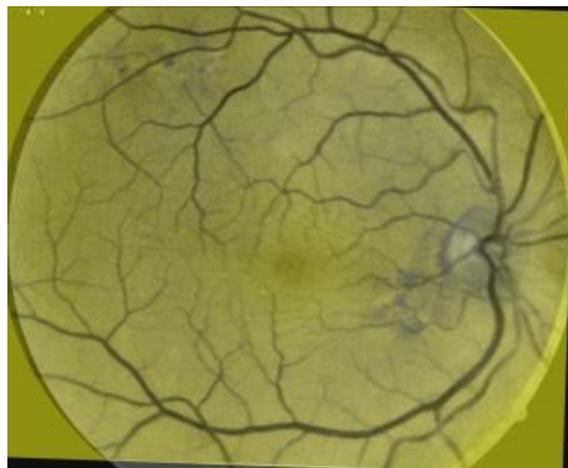


Figura 53 Unión lógica de la imagen a registrar sobre la imagen de referencia.

Graficación del error. En este caso el error medio cuadrático entre los conjuntos consistentes de características es de 1.4233 pixel (figura 54), el cual aún no se alcanza a apreciar a simple vista. En esta figura se observa que las 21 características consistentes originalmente obtenidas (triángulos verdes), una vez transformadas (asteriscos azules) se trasladaron hacia dentro de los círculos rojos que representan las características correspondientes en la imagen de referencia.

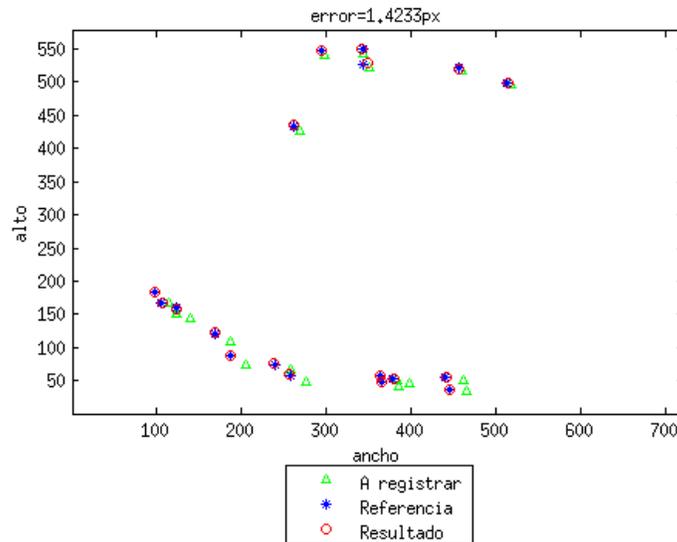


Figura 54 Graficación del error.

Vemos que en esta segunda parte de la prueba, se obtuvo un error de registro mayor que en la primera parte. Sin embargo, en la figura 53 no se aprecia aún ninguna distorsión o falta de coincidencia entre las redes vasculares de ambas imágenes, a pesar de que la traslación y rotación requeridas fue mayor.

5.2 Experimento 2 (Bidireccional)

Para este par de imágenes de retina se obtuvo el registro de forma bidireccional. A continuación se muestran las características del registro para las dos pruebas, tanto en la forma óptica-fluoresceínica (2a) como en la fluoresceínica-óptica (2b); cada prueba contiene las imágenes resultantes de cada etapa del registro de imágenes. De forma general se puede observar que la traslación y rotación de las dos pruebas es diferente, lo que se puede observar en las correspondientes gráficas del error, con valores de 1.4263px y 2.0928px respectivamente.

5.2.1 Experimento 2a

Tipo: Sin microaneurismas

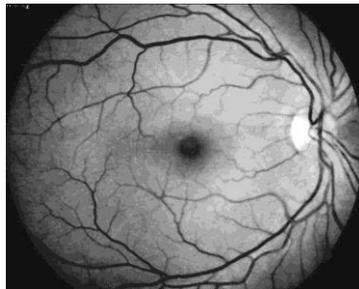
Categoría: o-f (óptica-fluorescénica)

A registrar/nombre: Imagen óptica / 2a_c2

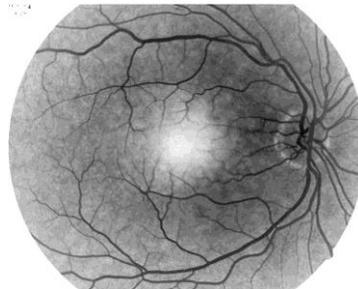
De referencia/nombre: Imagen fluorescénica / 2b_c2

Preprocesamiento

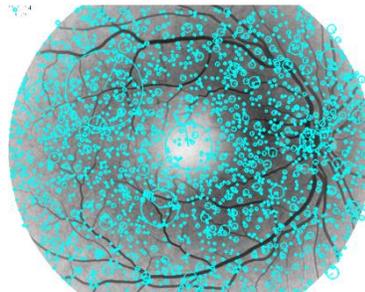
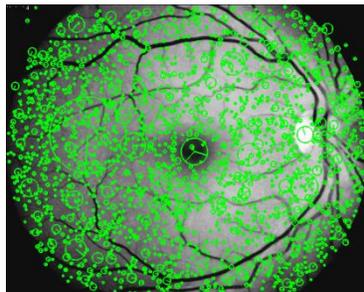
A registrar



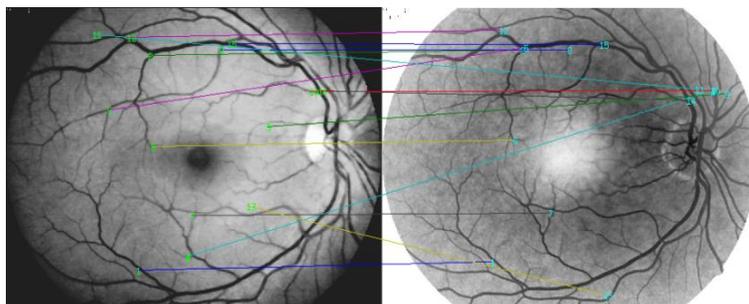
De referencia



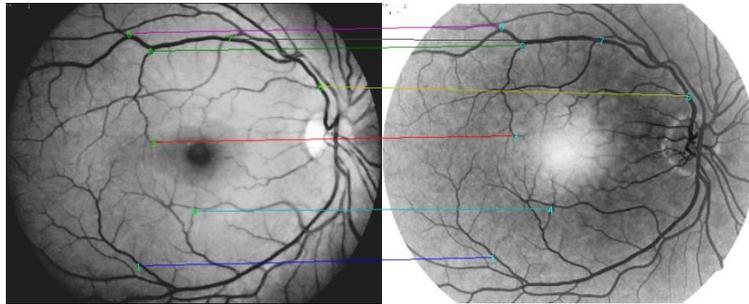
Selección de características



Correspondencia

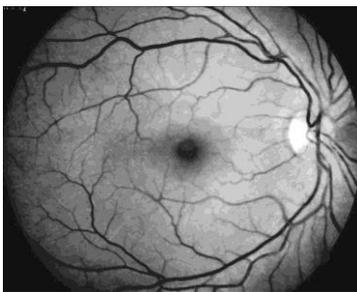


Correspondencia Euclidiana.

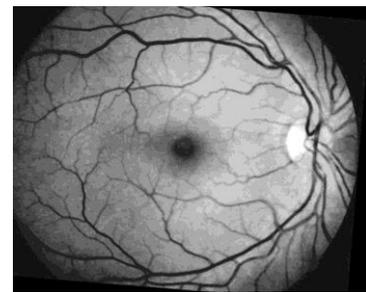


Correspondencia de Hough

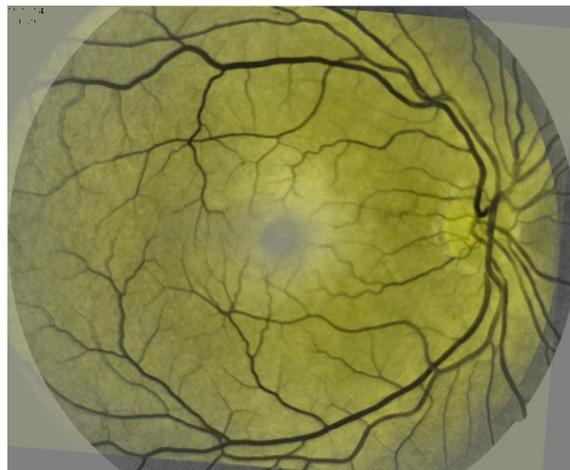
Ajuste – Transformación



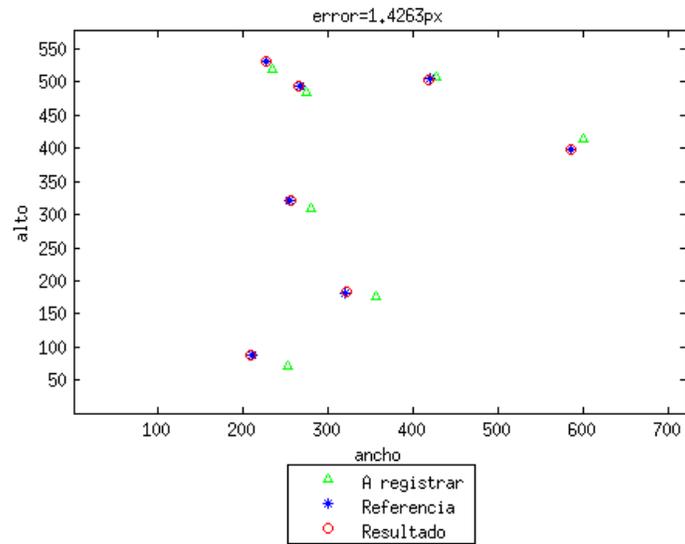
$$* \begin{bmatrix} 1.0020 & -0.0821 & -2.8289 \\ 0.0861 & 0.9922 & -33.3705 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} =$$



Unión



Error



5.2.2 Experimento 2b

Tipo: Sin microaneurismas

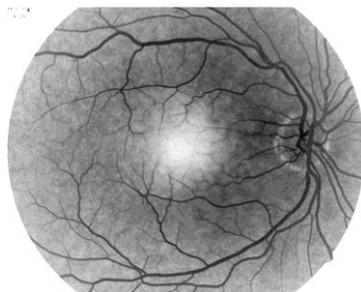
Categoría: f-o (fluoresceínica-óptica)

A registrar/nombre: Imagen fluoresceínica / 2b_c2

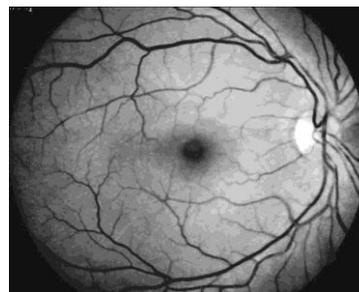
De referencia/nombre: Imagen óptica / 2a_c2

Preprocesamiento

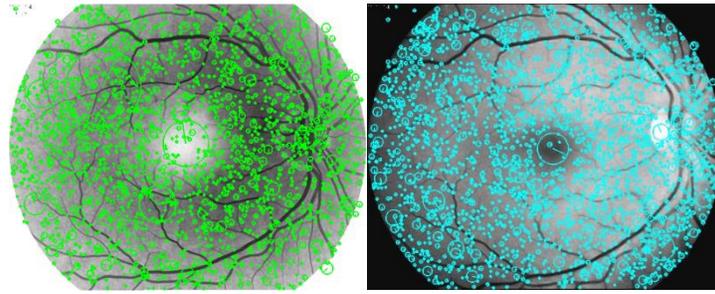
A registrar



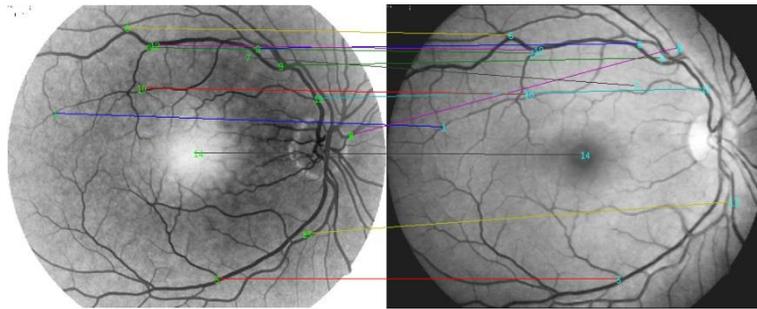
De referencia



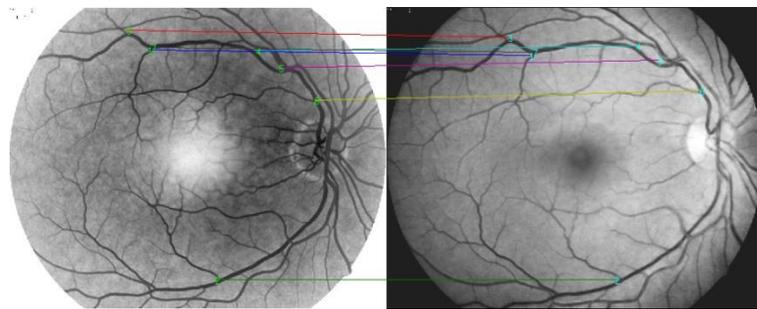
Selección de características



Correspondencia

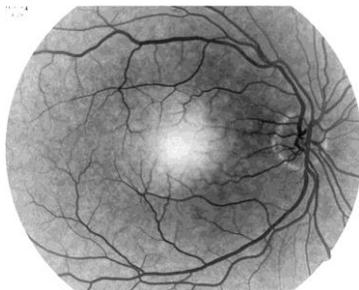


Correspondencia Euclidiana

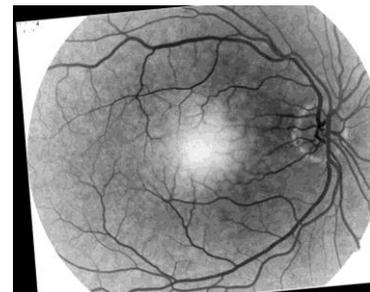


Correspondencia de Hough

Ajuste – Transformación



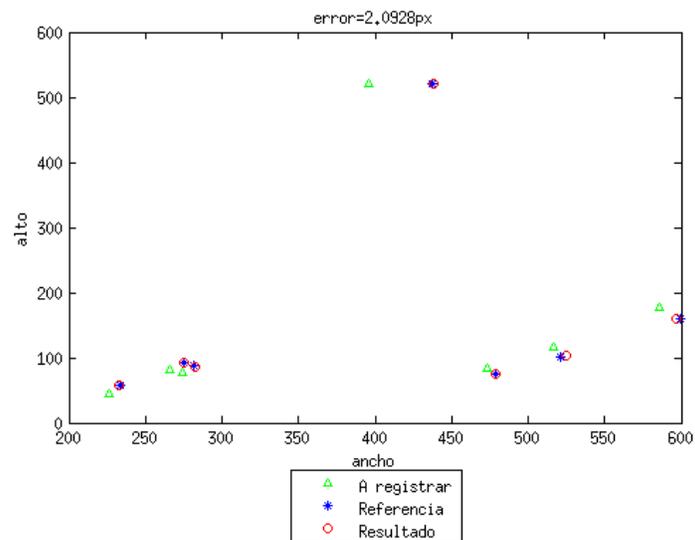
$$* \begin{bmatrix} 0.9848 & 0.0805 & 5.9298 \\ -0.0917 & 1.0068 & 33.5515 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} =$$



Unión



Error



5.3 Experimento 3 (Óptica – Fluorescénica)

Para esta prueba se obtuvo el registro óptica-fluorescénica, en el cual el par de imágenes de retina presenta microaneurismas, tal como se observa en la imagen de referencia; el registro se logró con un error de 1.3235px.

Experimento: 3a

Tipo: Con microaneurismas

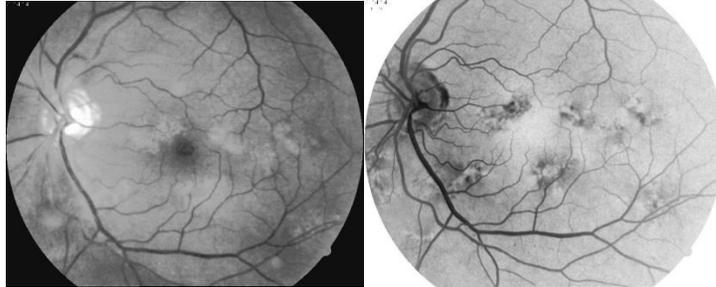
A registrar/nombre: Imagen óptica / 1a_c1

De referencia/nombre: Imagen fluoresceínica / 1b_c1

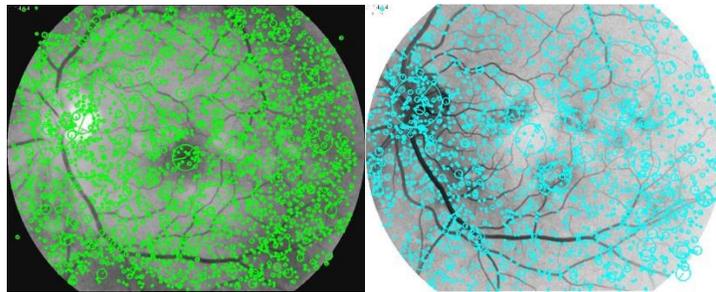
Preprocesamiento

A registrar

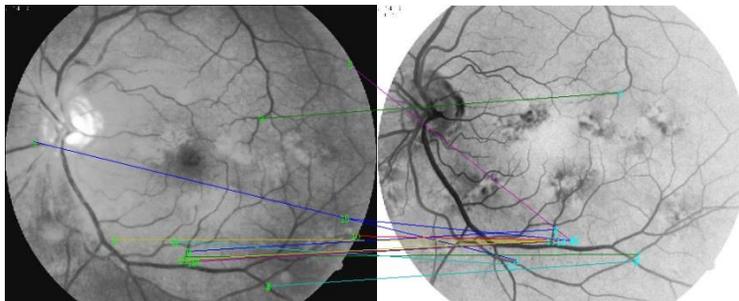
De referencia



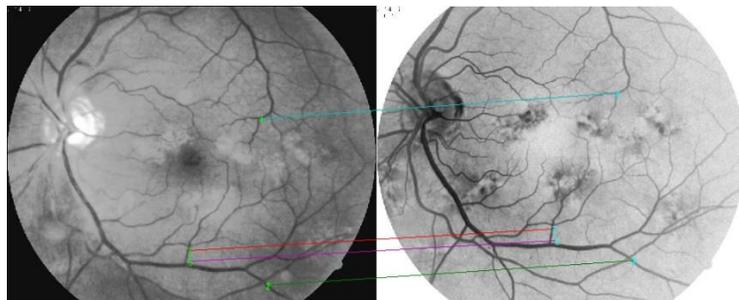
Selección de características



Correspondencia

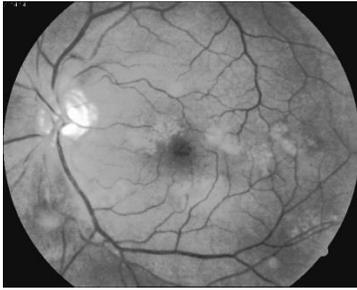


Correspondencia Euclidiana

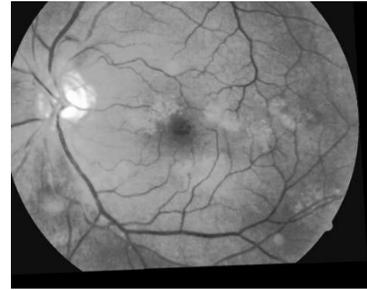


Correspondencia de Hough

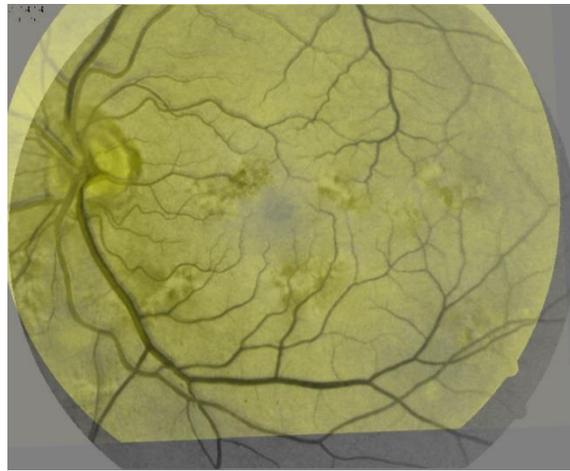
Ajuste – Transformación



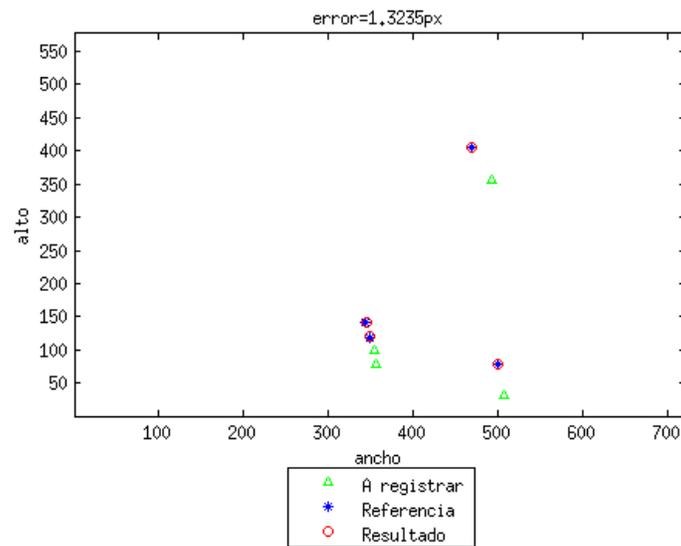
$$* \begin{bmatrix} 1.150292 & 0.1298 & -125.69 \\ 0.0321 & 1.0431 & -73.508 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} =$$



Resultado



Error



5.4 Experimento 4 (Fluorescénica – Óptica)

Esta prueba se realizó de forma fluorescénica-óptica y pertenece al grupo de imágenes con microaneurismas; la unión de las imágenes a registrar y de referencia resultó con un error de 2.4501px.

Experimento: 4a

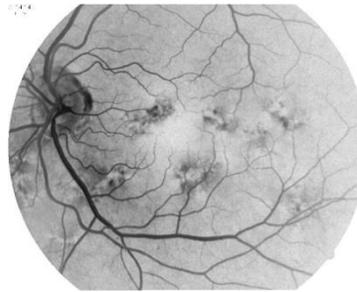
Tipo: Con microaneurismas

A registrar/nombre: Imagen óptica / 3b_c1

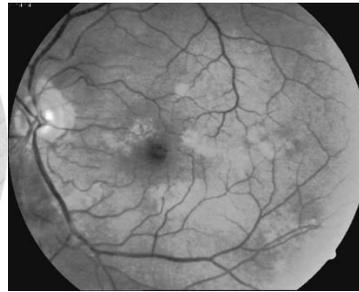
De referencia/nombre: Imagen fluorescénica / 3a_c1

Preprocesamiento

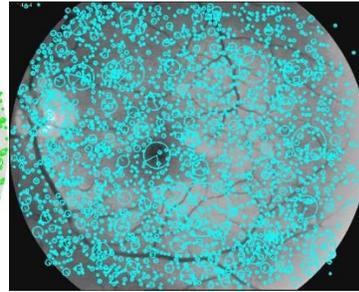
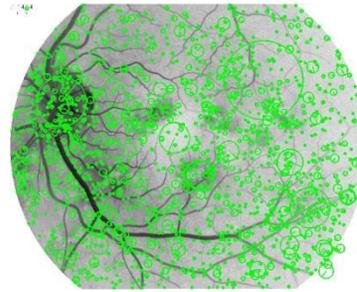
A registrar



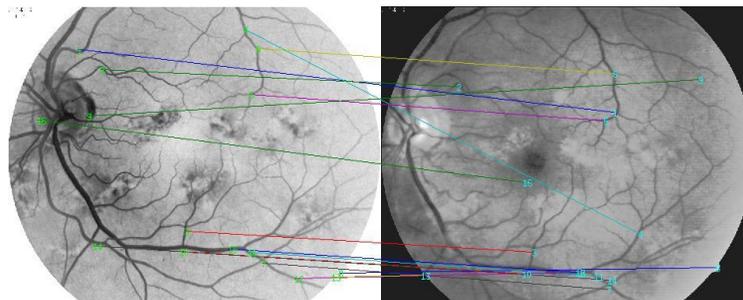
De referencia



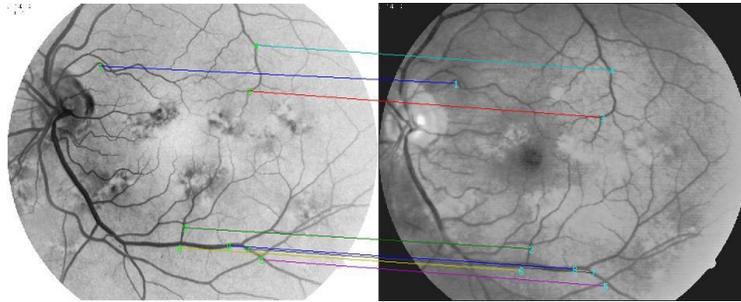
Selección de características



Correspondencia

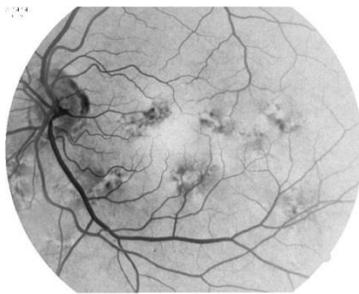


Correspondencia Euclidiana

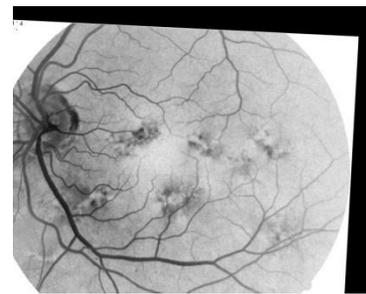


Correspondencia de Hough

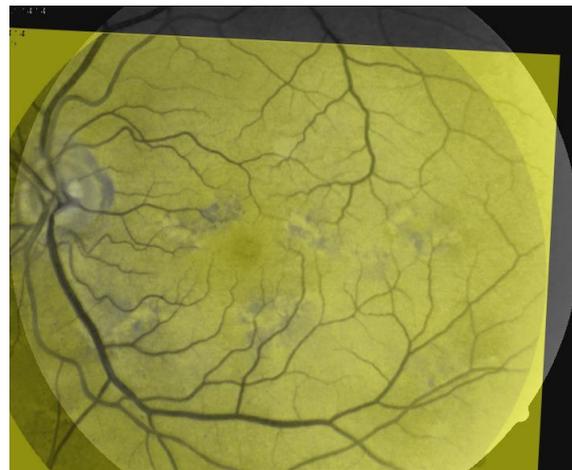
Ajuste – Transformación



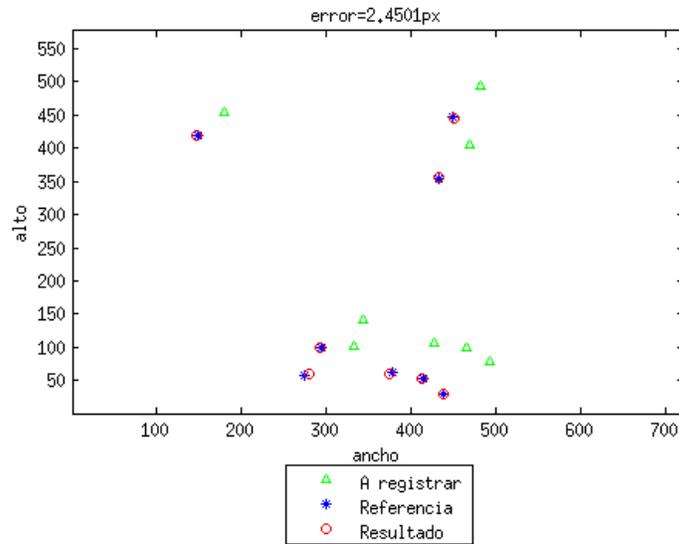
$$* \begin{bmatrix} 0.9985 & -0.0547 & -25.9549 \\ 0.0483 & 0.9989 & 25.9183 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix} =$$



Unión



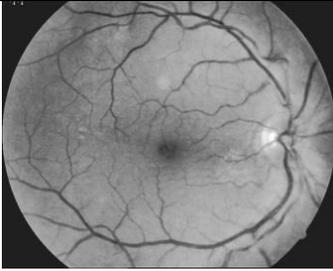
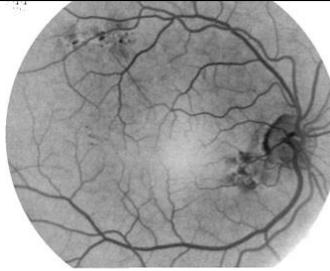
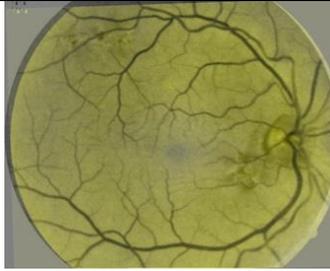
Error

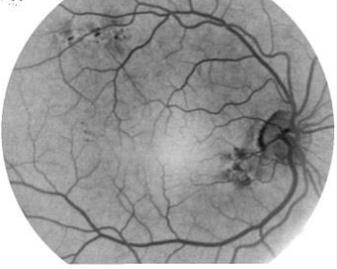
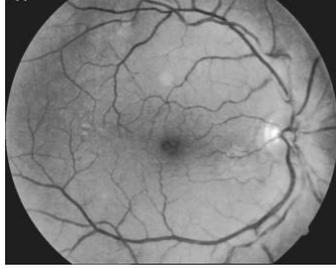
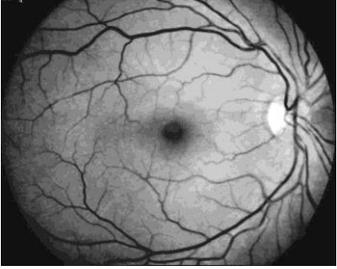
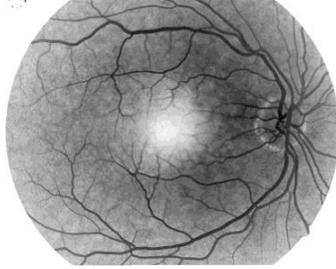
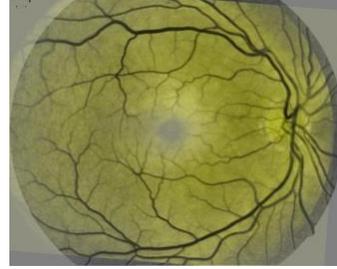
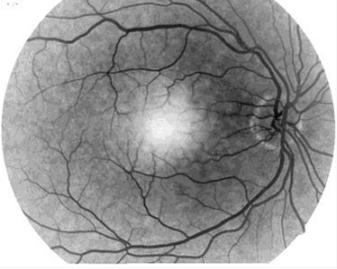
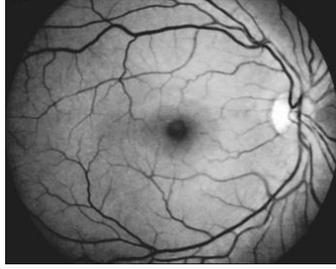
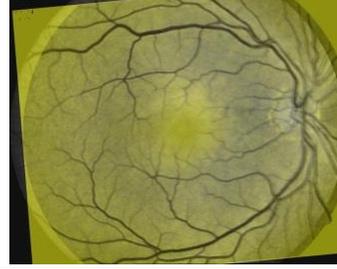
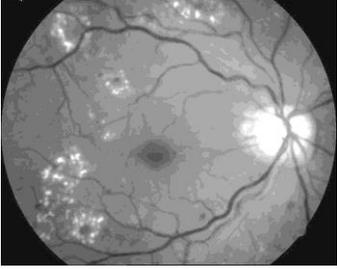
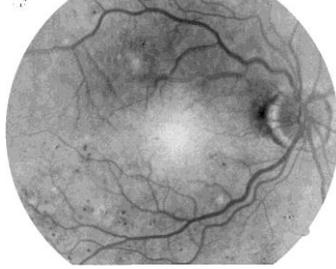
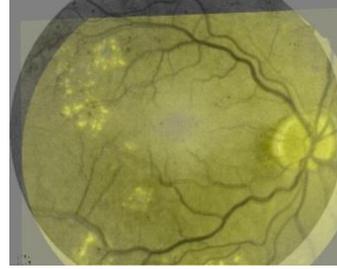
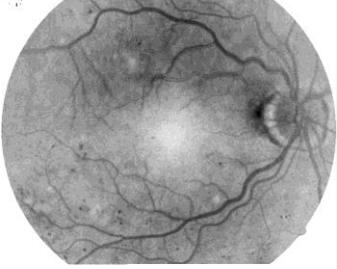
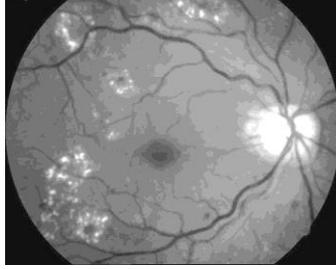
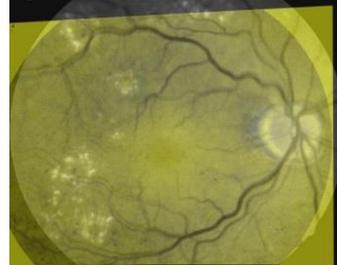


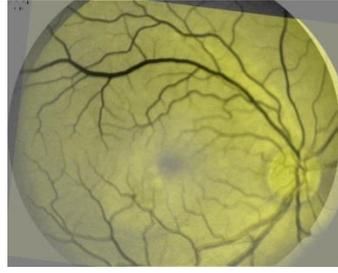
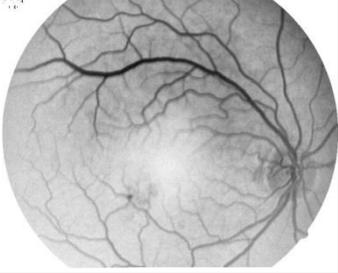
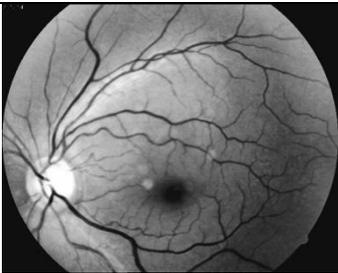
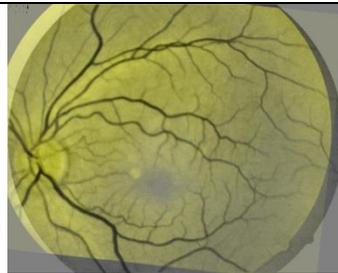
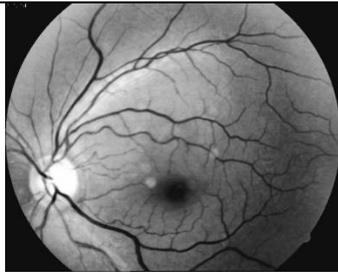
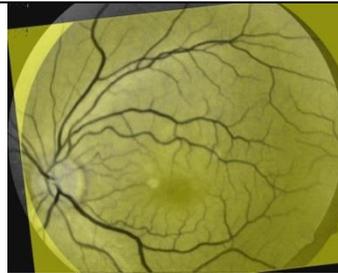
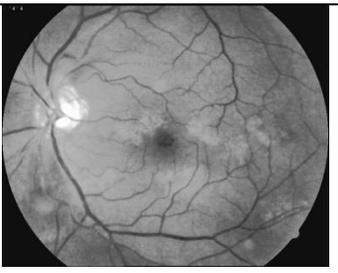
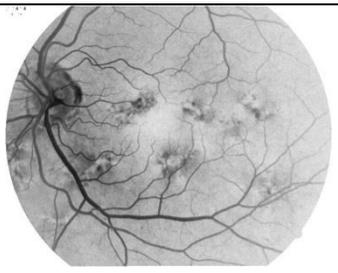
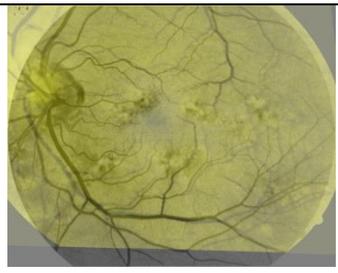
5.5 Tabla de resultados

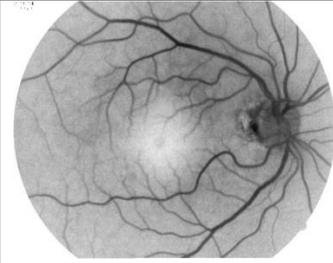
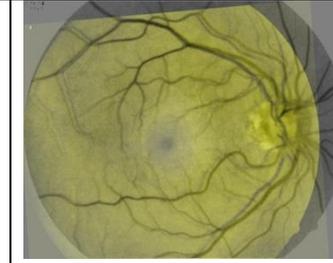
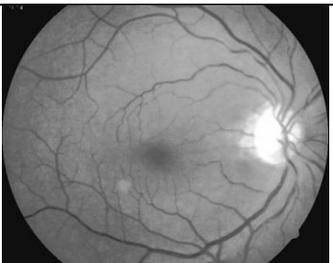
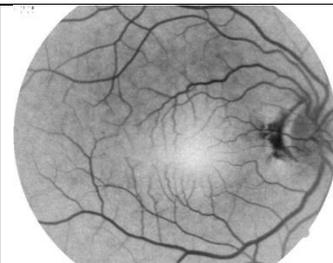
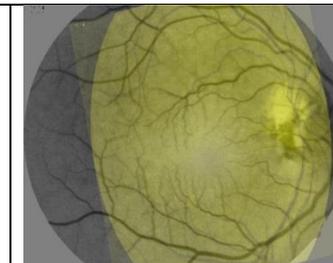
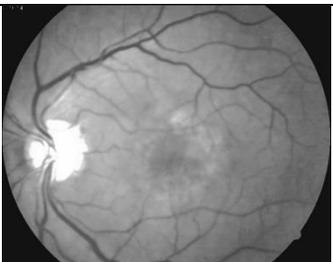
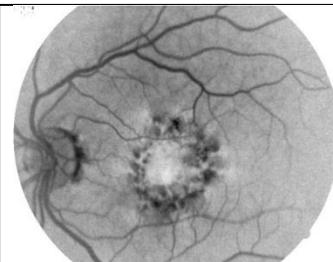
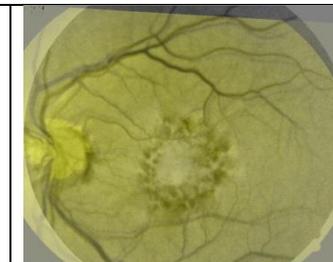
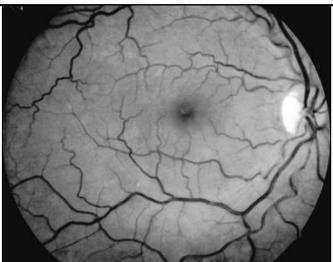
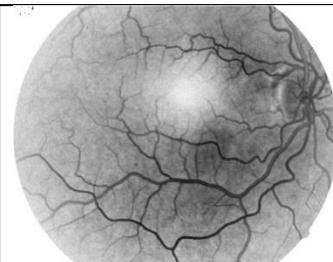
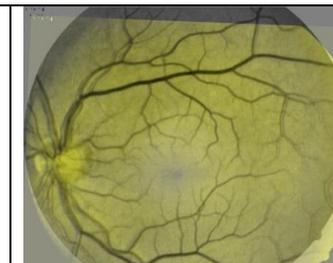
La tabla 3 contiene el conjunto de 26 pruebas con las que se logró el registro de imágenes de un total de 35 pares de imágenes de retina ópticas y fluoresceínicas disponibles, así como su correspondiente valor del error cuadrático medio; la primera parte corresponde al conjunto de pruebas bidireccionales, a continuación el conjunto de pruebas óptica-fluoresceínica y por último el conjunto de pruebas fluoresceínica-óptica.

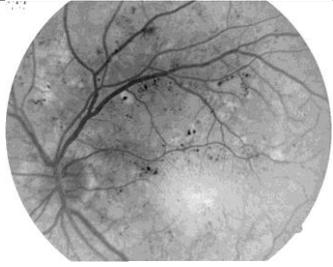
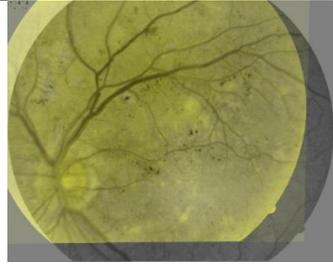
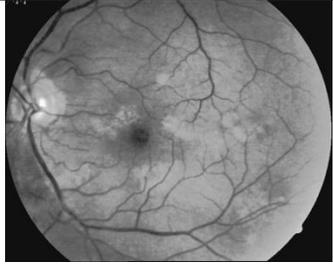
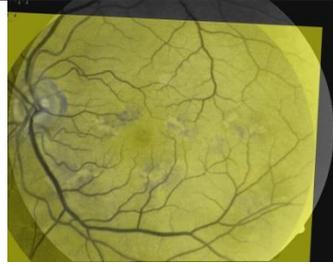
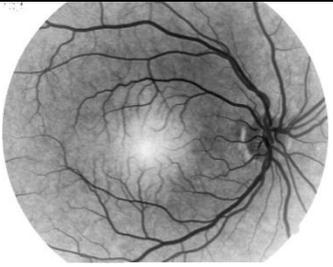
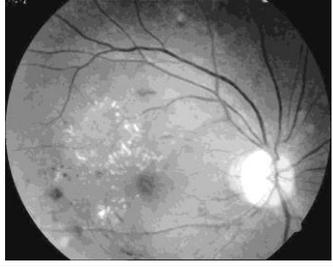
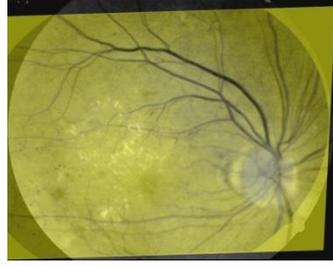
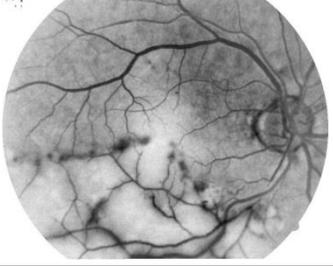
Tabla 3 Tabla de resultados.

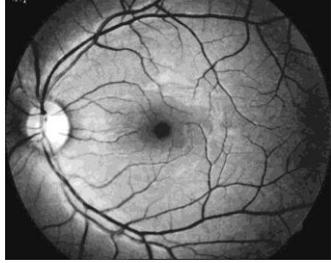
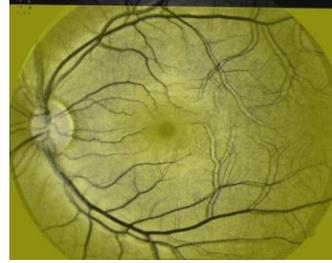
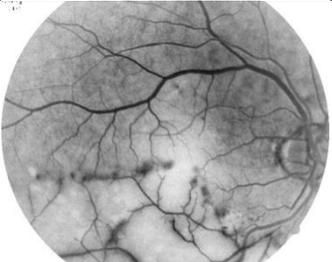
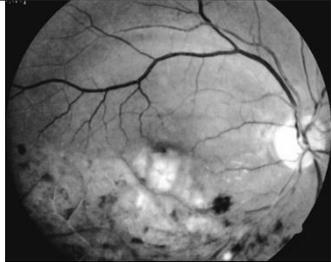
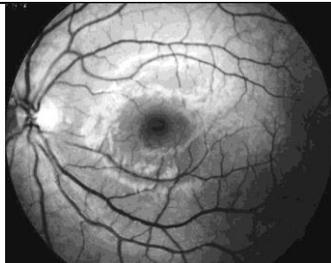
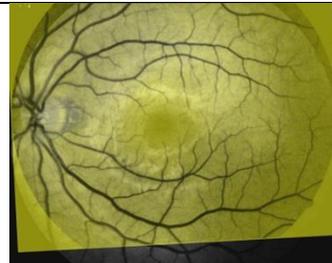
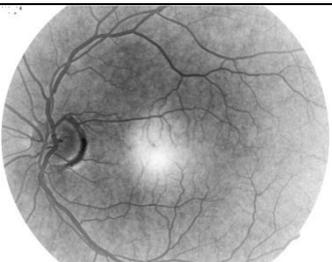
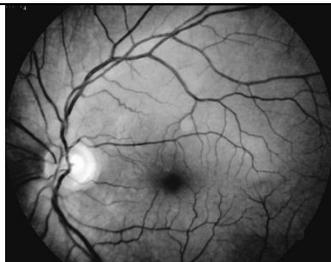
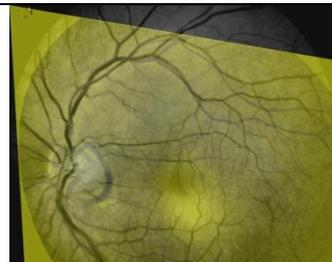
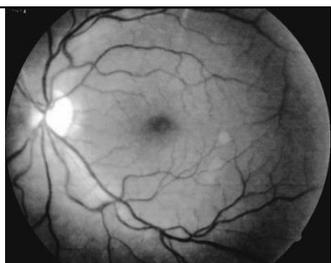
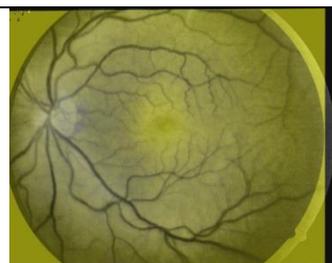
A registrar	De referencia	Unión	Error
Bidireccionales			
			0.9688px

			1.4233px
2_c1			
			1.4263px
2_c2			
			2.0928px
2_c2			
			0.5039px
5_c1			
			2.2331px
5_c1			

			1.2288px
			0.8980px
5_c2			
			2.4444px
			1.156px
6_c2			
Ópticas-Fluoresceínicas			
			1.5468px
1_c1			

			1.3336 e^{-13} px
4_c1			
			17.013px
7_c1			
			4.8165px
8_c1			
			0.2889px
14_c2			
			1.8688px
16_c1			

			0.5736px
19_c1			
Fluorescénicas- Ópticas			
			2.4501px
3_c1			
			41.290px
4_c2			
			7.4237px
6_c1			
			4.5023px
9_c1			

			7.3383px
9_c2			
			1.7242px
10_c1			
			0.7076px
11_c1			
			8.7602 e^{-14} px
12_c1			
			2.1066px
13_c2			

			0.3222px
14_c1			
			1.7854px
16_c2			
			3.517 e^{-13} px
17_c1			
			1.3655px
17_c2			
			4.2307px
25_c1			

A continuación se presentan algunas imágenes en las que no fue viable el registro; es importante mencionar que en nuestra opinión no se consiguió debido a que en la etapa de correspondencia de características no fue posible asociar 4 o más características en común,

por lo que las siguientes etapas no se pudieron realizar. Como puede observarse, la calidad de las imágenes deja mucho que desear.

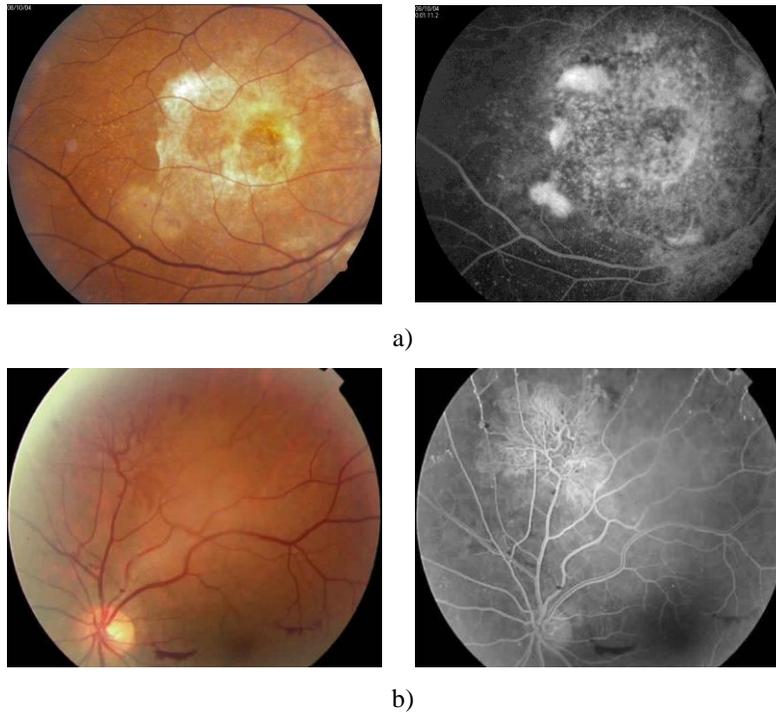


Figura 55 a), b) pares de imágenes de retina con las que no fue posible lograr el registro.

El motivo de por qué no se pudo hacer corresponder las características seleccionadas es porque después de aplicar el módulo de preprocesamiento las imágenes seguían siendo muy diferentes radiométricamente y de muy mala calidad, por lo que al final solo se consiguió corresponder una o dos características.

La tabla 4a) presenta los porcentajes y el error promedio obtenido del número de pares de imágenes en los que el registro fue viable y en los que no.

La tabla 4b) presenta el número, porcentaje y error promedio de los pares que presentaban microaneurismas y los que no.

La tabla 4c) presenta el número, porcentaje y error promedio en las tres diferentes formas en las que fue viable el registro.

Tabla 4 a), b), c) Resumen de los resultados de las pruebas.

	No. Pares	Porcentaje	Error Promedio (px)
El registro fue viable	26	74.28 %	2.48
El registro no fue viable	9	25.72 %	0
Total	35	100 %	2.48

a)

	No. Pares	Porcentaje	Error Promedio (px)
Con microaneurismas	12	46.15 %	1.14
Sin microaneurismas	14	53.85 %	1.34
Total	26	100 %	2.48

b)

	No. Pares	Porcentaje	Error Promedio(px)
Bidireccional	5	19.23 %	0.46
Óptica – Fluoresceínica	7	26.92 %	0.83
Fluoresceínica – Óptica	14	53.85 %	1.19
Total	26	100 %	2.48

c)

Al revisar los resultados de los experimentos realizados, observamos que el error promedio mayor es de 2.48 píxeles; de esto algo destacable es que en la mayoría de los casos este error fue más pequeño que ese valor, pero lo interesante tiene que ver con los casos en los que no fue viable el registro; en estos casos, aunque la cantidad de puntos clave fueron los necesarios para conseguir la matriz de transformación, la distribución de los puntos clave no fue adecuada; con esto nos referimos a que los puntos clave correspondían solo a una porción de la imagen, mientras que en los otros casos estaban distribuidos sobre toda la imagen.

Es oportuno hacer notar que en el análisis de retinas humanas, es muy común sacar varias imágenes de fondo de ojo de ambos ojos de los pacientes, precisamente para evitar el análisis (ya sea automático o manual) de imágenes que tengan poca calidad.

Capítulo VI

**Conclusiones y
trabajo futuro**

En este capítulo presentamos las conclusiones a las que hemos llegado después de revisar los experimentos realizados, así como los trabajos futuros en los que proponemos variantes en la metodología que sin dudas nos permitirá obtener nuevos resultados.

6.1 Conclusiones

El éxito del registro de imágenes de retina que nos compete, de acuerdo al caso de estudio que ha dado lugar al mismo, se debe en gran medida al módulo de preprocesamiento propuesto para que ambas imágenes incrementen su similitud a la vista, es decir, que visualmente tengan un mayor parecido. Es necesario aclarar que con este módulo no se consiguió registrar todas las parejas de imágenes de retina disponibles para las pruebas, debido a la mala calidad de algunas de las imágenes ópticas o fluoresceínicas disponibles, fundamentalmente debido a que en general las imágenes no fueron tomadas bajo condiciones estrictas de control.

La metodología propuesta fue capaz de conseguir el registro de imágenes de retina con ciertas características en común, como son el que se haya podido distinguir la red vascular de la retina y que los tonos alrededor de la misma aparezcan homogéneos. Algo importante de mencionar es que el proceso de registro utilizado permitió que el registro de imágenes haya sido viable en otros tipos de registro como es el caso del registro multitemporal y el registro multivista, obviamente con otras parejas de imágenes diferentes a las de retina, de los cuales no se mencionaron detalles ni se dieron resultados por no formar parte del contenido de este trabajo.

Con base en los objetivos específicos o particulares podemos brindar las siguientes conclusiones:

- La elección de operaciones que conforman el módulo de procesamiento fue determinante para la viabilidad del registro de imágenes digitales multimodales.
- El registro de imágenes multimodales fue viable con la metodología propuesta, tanto para el caso de estudio seleccionado con imágenes de retina, como para registros de tipo multitemporal y multivista con otras parejas de imágenes diferentes.
- Se implementó una serie de algoritmos de registro para resolver el caso de estudio, de los cuales podemos destacar los basados en plantillas y los basados en selección de puntos de interés; estos últimos presentaron ventajas significativas que facilitaron e hicieron posibles los registros logrados.
- Se diseñó y desarrolló una interfaz gráfica que permite manipular confiablemente la metodología propuesta.

- Se validó el desempeño de la implementación propuesta con resultados satisfactorios, a través de un vasto conjunto de pruebas que incluyeron las parejas de imágenes de retina ópticas y fluoresceínicas disponibles.

Después de revisar las imágenes resultantes y el registro multimodal que se realizó, podemos observar lo siguiente:

- La imagen de referencia debe tener una red vascular bien definida y los tonos de gris en el fondo alrededor de ésta deben aparecer homogéneos.
- La distribución de los puntos clave es relevante, ya que si la mayoría de los puntos clave que se obtienen están agrupados en una sola zona de la imagen, no hay seguridad de que la transformación sea adecuada para lograr el registro.
- La viabilidad del registro está condicionada a la cantidad de puntos de interés y su distribución respecto al tamaño de la imagen; con base en lo que aparece en la literatura abierta, sabemos que es necesario un número de características mayor que cuatro para garantizar el éxito del registro; en caso contrario el registro no será viable.
- Aunque el número de puntos clave necesarios para poder conseguir el registro es 4, observamos que entre 8 y 16 puntos clave proporcionaron suficiente información complementaria de la imagen para obtener una adecuada transformación y por lo tanto un registro exitoso. Esto no es general, pues dependerá del tipo de imágenes con las que se desea realizar el registro, ya que unas imágenes tienen más detalles que otras.

La forma en que se validó el funcionamiento de la metodología propuesta fue mediante el cálculo del error medio cuadrático existente entre el conjunto consistente de características de la imagen de referencia y el conjunto consistente de características de la imagen a registrar operado con la matriz de transformación, la que fue obtenida después del proceso de ajuste; si consideramos este error en términos de regresión, podemos interpretarlo como cuán distantes están los puntos clave respecto a una recta de regresión formada de acuerdo a las observaciones y predicciones; en términos de la imagen nos indica que un pixel transformado está a un error cuadrático dado del pixel correspondiente en la imagen de referencia.

6.2 Aportaciones

La principal aportación de este trabajo es la de una nueva metodología que permitió realizar el registro de imágenes de retina ópticas y fluoresceínicas, en particular, con la introducción de una primera etapa de preprocesamiento que incrementó la similitud de ambas imágenes, que hizo posible la realización de las siguientes etapas requeridas para el registro. Sin la inclusión de la etapa de preprocesamiento propuesta, no fue posible nunca lograr el registro de las parejas de imágenes ópticas y fluoresceínicas de una misma retina disponibles.

6.3 Trabajo futuro

Posibles trabajos que pueden ser realizados en el futuro son:

- Considerar nuevas opciones para el módulo de preprocesamiento con la finalidad de evaluar nuevos resultados con las mismas parejas de imágenes utilizadas.
- Probar con otras formas de selección de características, por ejemplo SURFT (Speed Up Robust Feature) que se traduce como: transformación robusta acelerada de características.
- Considerar otras formas de correspondencia de características, por ejemplo RANSAC. (Random Sample Consensus) que se traduce como: consenso de muestreo aleatorio.
- Ejecutar un conjunto de pruebas que incluyan cambio de perspectiva, cambio de escala y un registro consecutivo.
- Explorar nuevas transformaciones que operen a través de una perspectiva tridimensional.
- Explorar otras opciones de medición del error.

Referencias

- Ahmed, M. *Image Processing: Theory, Algorithms and Architecture*. Michigan: McGraw-Hill, (1995).
- Brown, G. L. «Survey of Image Registration Techniques.» *ACM Computing Surveys* 24, n° 4 (1992): 325-376.
- Burger, W y Burge, M J. *Digital Image Processing*. New York: Springer Science, (2008).
- Canny, J. «A computational approach to edge detection.» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n° 8 (1986): 679-698.
- Gonzalez, R y Woods, R. *Digital Image Processing*. Tercera edición. Prentice-Hall, (2008).
- Harris, C y Stephens, M. «Proc of the 4th Alvey Vision Conference.» *Plessey Research Roke Manor*, (1988): 147-151.
- Jin, F y Feng, D. «Image Registration Algorithm Using Mexican Hat Function-Based Operator and Grouped Feature Matching Strategy.» *Open Access PLOS ONE* 9 (2014).
- Thomas George B. «Cálculo, varias variables (undécima edición)». Massachusetts Institute of Technology. Pearson Educación de México, S.A. de C.V (2006).
- Liu, J, Akas, A, Obaidi, A y Moravec, A. «A comparative study of different corner detection methods.» *Computational Intelligence in Robotics and Automation*, (2009).
- Leithold, L. «El Cálculo (séptima edición) ». Oxford University Press México, S.A. de C.V. (1998).
- Lowe, D G. «Distinctive image features from scale-invariant key points.» *International Journal of Computer Vision* 60, n° 2 (2004): 91-110.
- Marr, D C y Hildreth, E. «Theory of Edge Detection.» *Proceedings of the Royal Society of London* 207, n° 1167 (1980): 187-217.
- Saxena, S y Kumar, R. «A Survey of Recent and Classical Image Registration Methods.» *Image Processing and Pattern Recognition*, n° 4 (2014): 167-176.
- Sharma, K y Goyal, A. «Classification Based Survey of Image Registration Methods.» *International Conference on Computer and Communication Technology*, (2013).

Referencias.

Vedaldi, A y Fulkerson, B. *An Open and Portable Library*. 2008. <http://www.vlfeat.org/> (último acceso: 1 de Noviembre de 2014).

Vishwakarma, H, y Katiyar, S K. «Comparative study of edge Detection algorithm on the remote sensing images using matlab.» *International Journal of Advances in Engineering Research* 2, nº 6 (2011).

Wiley, J. *Encyclopedia of Computer Science and Engineering*. New Jersey, (2009).

Zitová, B y Flusser, J. «Image registration methods: A survey.» *Image and Vision Computing*, nº 21 (2003): 977-1000.