



UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES
DEPARTAMENTO DE INVESTIGACIÓN EN FÍSICA

Programa de Doctorado en Ciencias (Física)

**ALGORITMO PARA AUTOENFOQUE Y FUSIÓN
SELECTIVA PARA EL ANÁLISIS DE IMÁGENES
MICRO Y MACROSCÓPICAS**

Tesis

Que para cubrir parcialmente los requisitos necesarios para obtener el grado de
Doctor en Ciencias (FÍSICA)

Presenta:

ALMA ROCÍO CABAZOS MARÍN

**Hermosillo, Sonora, México.
Junio del 2018**

Universidad de Sonora

Repositorio Institucional UNISON



**"El saber de mis hijos
hará mi grandeza"**



Excepto si se señala otra cosa, la licencia del ítem se describe como openAccess



UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES
Departamento de Investigación en Física
Programa de Posgrado en Ciencias (Física)

A QUIEN CORRESPONDA:

La que suscribe, Coordinadora del Programa de Posgrado en Ciencias (Física) de la Universidad de Sonora, hace constar que el **DR. JOSUÉ ÁLVAREZ BORREGO**, fungió como Director de Tesis de Doctorado "*Algoritmo para autoenfoco y fusión selectiva para el análisis de imágenes micro y macroscópicas*", que para obtener el grado de Doctor en Ciencias (Física) sustentó M.I. Alma Rocío Cabazos Marín, el día 22 de junio de 2018.

A petición del interesado y para los usos legales que estime convenientes, se extiende la presente en la ciudad de Hermosillo, Sonora, a los veintidós días del mes de junio del año dos mil dieciocho.

ATENTAMENTE
"EL SABER DE MIS HIJOS HARÁ MI GRANDEZA"



El saber de mis hijos
hará mi grandeza
Posgrado en
Ciencias (Física)
Departamento
de Investigación
en Física

DRA. SUSANA ÁLVAREZ GARCÍA
COORDINADORA DEL PROGRAMA DE
POSGRADO EN CIENCIAS (FÍSICA)

SAG/smlg.

Tesis defendida por
ALMA ROCÍO CABAZOS MARÍN
y aprobada por el siguiente comité

Dr. Josué Álvarez Borrego
Departamento de Óptica, CICESE

Dr. Alejandro Apolinar Iribe
Departamento de Física, UNISON

Dr. Francisco Javier Cuevas de la Rosa
Centro de Investigaciones en Óptica, CIO

Dr. Jorge Alberto Gaspar Armenta
Departamento de Investigación en Física,
UNISON

Dr. Eduardo Tepichín Rodríguez
Instituto Nacional de Astrofísica, Óptica y
Electrónica, INAOE

Dr. Felipe Ramos Mendieta
Coordinador del Programa de Doctorado
en Ciencias (Física), UNISON

Índice

Contenido	Página
Dedicatoria	I
Agradecimientos	II
Resumen	III
Abstract	IV
Lista de figuras	V
Lista de tablas	VI
Productos obtenidos	VII
Introducción	1
Capítulo 1. Antecedentes	5
1.1. Estado del arte	5
1.2. Procesamiento digital de imágenes, PDI	9
1.2.1. Adquisición de la imagen digital	10
1.2.2. Mejora de la imagen digital	13
1.3. Autoenfoco y fusión	14
1.4. Análisis de algoritmos de autoenfoco	14
1.4.1. Energía de la imagen Laplaciana	17
1.4.2. Tenengrad	18
1.4.3. Tenengrad absoluto	18
1.4.4. Varianza	19
1.4.5. Varianza normalizada	19
1.4.6. Vollath's F4	19
1.4.7. Vollath's F5	20
1.4.8. Filtro Gaussiano	20
1.4.9. Histograma Logarítmico	20
1.4.10. Histograma ponderado	21

1.4.11. Autoenfoco por correlación de Pearson y Fourier	21
1.5. Análisis de algoritmos para la fusión	23
1.5.1. Promedio simple (<i>SA</i>)	24
1.5.2. Máximo simple (<i>SM</i>)	25
1.5.3. Análisis de componente principal (<i>PCA</i>)	25
1.5.4. Transformada de coseno discreta (<i>DCT</i>)	26
1.5.5. Razón de la pirámide de paso-bajo (<i>ROLP</i>)	27
1.5.6. Transformada discreta de <i>wavelets</i> (<i>DWT</i>)	28
1.6. Métricas para la evaluación de calidad de la imagen	30
1.6.1. Métrica Wang and Bovik	31
1.6.2. Métrica Piella and Heijmans	32
1.6.3. Métrica Multi-image.	33
Capítulo 2. Metodología	35
2.1. Adquisición de la pila de imágenes	36
2.2. Determinar el vector de la imagen por escaneo en espiral	39
2.3. Espectro del vector de la imagen mediante transformada de Fourier	41
2.4. Determinar la imagen mejor enfocada con CNL	41
2.5. Fusión de imágenes aplicando el filtro parabólico	44
Capítulo 3. Resultados y discusión	48
3.1. Resultados de la obtención de la imagen mejor enfocada	48
3.2. Evaluación de calidad de la fusión usando las métricas Q , Q_{Wt} y Q_E	53
3.3. Discusión	59
Capítulo 4. Conclusiones	61
Referencias	62
Anexo	65

Dedicatoria

A mis hijos Luz Adriana, Jonathan y a mi esposo Juan Tapia.

Agradecimientos

A Dios y a mi familia por su amor y apoyo incondicional.

Al Dr. Josué Álvarez Borrego por su acertada dirección, tutoría y apoyo en la realización este trabajo.

A la coordinación del departamento del Departamento de Investigación en Física de la Universidad de Sonora, DIFUS, y a su personal administrativo, en particular al coordinador Dr. Felipe Ramos Mendieta y a Sandra León González por el apoyo oportuno en cada una de las etapas de este trabajo de tesis.

A los miembros del comité de seguimiento de tesis, Dr. Ángel Coronel Beltrán, Dr. Alejandro García Juárez y Dr. Julio Saucedo Morales, por sus aportaciones, tutorías y recomendaciones durante el desarrollo de este trabajo.

A los miembros del comité de revisión de tesis, Dr. Francisco Javier Cuevas de la Rosa (CIO), Dr. Eduardo Tepichín Rodríguez (INAOEP), Dr. Jorge Alberto Gaspar Armenta (DIFUS) y Dr. Alejandro Apolinar Iribe (DFUNISON), por sus acertadas sugerencias y observaciones al presente documento.

Al Consejo Nacional de Ciencia y Tecnología, CONACYT por su apoyo con la beca otorgada para la realización de esta investigación.

Mi agradecimiento también para el Centro de Investigación Científica y de Educación Superior de Ensenada, CICESE, por el apoyo brindado con su infraestructura y equipo de cómputo, en particular agradezco a Isaac Fimbres Valenzuela por su apoyo en el centro de cómputo del departamento de Óptica y a Julián Javier Delgado Jiménez del departamento de Telemática por su apoyo en cómputo paralelo.

Y a todas las personas, instituciones y entidades que contribuyeron a esta investigación.

Resumen de la tesis de **Alma Rocío Cabazos Marín** presentada como requisito parcial para la obtención del grado de **Doctor en Ciencias (Física)**.

**ALGORITMO PARA AUTOENFOQUE Y FUSIÓN SELECTIVA
PARA EL ANÁLISIS DE IMÁGENES MICRO Y MACROSCÓPICAS**

Resumen aprobado por:

Dr. Josué Álvarez Borrego
Director

En esta tesis se desarrolló un nuevo algoritmo de autoenfoco y fusión selectiva de imágenes, AFS, como una contribución al procesamiento digital de imágenes. El método aquí propuesto optimiza los procesos del autoenfoco y la fusión de imágenes capturadas por sistemas ópticos como el microscopio, cámara digital y telescopio, como resultado se obtienen imágenes de alta calidad en un tiempo óptimo de procesamiento.

El método de autoenfoco selecciona la imagen mejor enfocada de un conjunto de imágenes capturadas a diferentes distancias del objeto. Se define un vector para cada imagen en el conjunto, el vector contiene una muestra de los elementos de la imagen, los cuales son seleccionados mediante el barrido en espiral de la imagen. Posteriormente, se calcula el espectro de cada vector aplicando la transformada de Fourier. Luego, se realiza una correlación no lineal del espectro del vector de cada una de las imágenes del conjunto con el espectro de la imagen de referencia. Comparando los índices de correlación se determina una medida para determinar cual es la imagen mejor enfocada.

Una vez definida la imagen con mejor enfoque se lleva a cabo la fusión, la cual utiliza un subconjunto del conjunto total de imágenes, se seleccionan aquellas que tienen un valor de medición de enfoque cercano al valor de mejor enfoque. Para determinar los elementos relevantes de las imágenes los cuales se incluirán en la fusión, se utiliza el filtro parabólico.

La evaluación de resultados se realizó utilizando la métrica de multi-imagen, una medida de calidad que representa el porcentaje de mejora de la imagen fusionada comparada con la de mejor enfoque y el resto de las imágenes que integran la fusión. La comparación se realizó con otros métodos de fusión como las “*wavelets*”, como resultado el método de AFS propuesto obtuvo los índices de mayor calidad. Se concluye que la AFS mejora la calidad de las imágenes en un tiempo óptimo de procesamiento.

Palabras clave: *enfoco automático; fusión; correlación; evaluación de calidad de imagen; métrica; procesamiento digital de imágenes.*

Abstract of the thesis presented by **Alma Rocío Cabazos Marín** as a partial requirement for the degree of **Doctor of Science in Physics**.

ALGORITHM FOR SELF-FOCUS AND SELECTIVE FUSION FOR THE ANALYSIS OF MICRO AND MACROSCOPIC IMAGES

Abstract approved by:

Dr. Josué Álvarez Borrego
Director

In this thesis a new algorithm of autofocus and selective fusion of images, AFS, was developed as a contribution to the digital processing of images. The method proposed here optimizes the processes of autofocusing and the fusion of images captured by optical systems such as the microscope, digital camera and telescope, as a result high quality images are obtained in an optimal processing time.

The autofocus method selects the best-focused image of a set of images captured at different distances from the object. A vector is defined for each image in the set; the vector contains a sample of elements of the image, which are selected by the spiral scan of the image. Subsequently, the spectrum of each vector is calculated by applying the Fourier transform.

Then, a non-linear correlation of the spectrum of the vector of each of the images of the set with the spectrum of the reference image is made. By comparing the correlation indexes, a measure is determined to determine which is the best-focused image.

Once the image with the best focus is defined, the fusion is carried out, which uses a subset of the total set of images, those that have a focus measurement value close to the best focus value are selected. To determine the relevant elements of the images that will be included in the fusion, the parabolic filter is used.

The evaluation of results was carried out using the multiple-image metric, a measure of quality that represents the percentage of improvement of the fused image compared with that of better focus and the rest of the images that make up the merger. The comparison was made with other fusion methods such as “wavelets”. As a result the proposed AFS method obtained the highest quality indices. It is concluded that the AFS improves the quality of the images in an optimal processing time.

Keywords: *Autofocus; fusion; correlation; image quality evaluation; metrics; digital image processing.*

Lista de figuras

Fig.		Página
1	Descripción del sistema.	3
2	Etapas básicas del procesamiento digital de imágenes.	10
3	Descripción de los elementos de una imagen.	11
4	Sistema de captura.	12
5	Descripción del proceso de adquisición de la imagen digital.	12
6	Proceso de autoenfoque.	15
7	Enfoque de la imagen, captada a diferentes distancias.	16
8	Comparación del contraste entre las bandas de píxeles de una sección de la imagen desenfocada y de la IME, respectivamente.	16
9	Fusión de imágenes basada en Wavelets.	29
10	Metodología AFS.	35
11	Microscopio <i>Leica DM RXA2</i> con platina motorizada para X/Y y Z desplazamientos.	36
12	Conjunto de imágenes, donde $f(i, j)_w$ es una imagen digital capturada a una distancia focal Δz .	37
13	Conjuntos de imágenes utilizadas para las pruebas de autoenfoque y fusión.	38
14	Barrido en espiral de una imagen de 1616×1216 .	40
15	Identificación de la IME y la Imagen de Referencia en el conjunto.	42
16	Cada renglón representa tres distintos conjuntos de imágenes. En la columna (a) están las Imágenes de referencia $f(i, j)_I$ que son las más desenfocadas, mientras que en la columna (b) se muestran las imágenes mejor enfocadas $f(i, j)_{IME}$.	43
17	Metodología para determinar la IME, con CNL.	44
18	Las imágenes que serán fusionadas $f(i, j)_{IME}$ son un subconjunto del total de imágenes del conjunto original, aquellas con índice cercano a la IME.	45

Fig.		Página
19	Representación gráfica del filtro parabólico.	45
20	Comparación de las imágenes de referencia (izquierda) y la IME (derecha) de tres diferentes conjuntos de imágenes.	50
21	Se muestra en el eje x los índices de las imágenes y su medida de de-correlación d en el eje y . Correspondientes al conjunto de Tricodina. La IME será aquella que tiene el valor máximo de de-correlación.	51
22	El gráfico muestra en el eje x el nombre del algoritmo, en el eje y su correspondiente tiempo de ejecución en segundos, para el conjunto de 60 imágenes de Tricodina.	52
23	El gráfico muestra en el eje x el nombre del algoritmo, en el eje y su correspondiente tiempo de ejecución en segundos, para el conjunto de 60 imágenes de Tricodina.	52
24	El valor máximo del índice de de-correlación se utiliza para determinar el rango de imágenes a incluir en la fusión. El gráfico muestra en el eje X los índices de las imágenes y en el eje Y su medida de de-correlación, d , para el conjunto de Tricodina.	53
25	Detalle de un acercamiento de la IME obtenida por el autoenfoco y la imagen que resulta de la fusión.	54
26	Imágenes resultado de la fusión usando el algoritmo AFS.	55
27	Detalle de una sección de la imagen para comparar la calidad de fusión de las Wavelets con el método AFS .	56
28	Tiempo de ejecución en segundos, comparación de calidad de fusión DWT y AFS.	59

Lista de tablas

Tabla		Página
1	Características principales de los algoritmos de autoenfoco.	23
2	Características principales de los algoritmos de fusión.	30
3	Índice de la IME seleccionada para cada conjunto por cada algoritmo de prueba.	51
4	Comparación de las medidas de calidad Q , Q_{Wt} y Q_E de AFS con los otros algoritmos de fusión aquí analizados. Evaluación correspondientes empleando el conjunto de imágenes de Limbo Solar.	57
5	Comparación de las medidas de calidad Q , Q_{Wt} y Q_E de AFS con los otros algoritmos de fusión aquí analizados. Evaluación correspondientes empleando el conjunto de imágenes de Tricodina	58

Productos obtenidos

Publicación en revista arbitrada:

Cabazos-Marín, Alma Rocío & Álvarez-Borrego, Josué. (2018). “*Automatic focus and fusion image algorithm using nonlinear correlation: image quality evaluation.*”; *Optik* 164 10.1016/j.ijleo.2018.02.101; OPTIK ©2018 Elsevier, p. 224-242; <https://authors.elsevier.com/c/1Wj3Q6wQfGw6q>

Congreso Nacional e Internacional:

Alma Rocío Cabazos-Marín, Josué Álvarez-Borrego, Ángel Coronel-Beltrán. (2015). “*Quality assessment of autofocus and image fusion.*”; *MOPM-2015-OP-20*; Mexican Optics and Photonics Meeting (MOPM 2015). León, Guanajuato, México 9-11 Septiembre 2015.

Congreso Internacional:

Alma Rocío Cabazos-Marín, Josué Álvarez-Borrego, and Ángel Coronel-Beltrán. (2014). “*Autofocus and Fusion using nonlinear Correlation.*”; *AIP Conference Proceedings* 1618, 147 (2014); doi: 10.1063/1.4897702; International Conference of Computational Methods in Sciences and Engineering (IMMSE 2014). Atenas, Grecia, 04-07 abril 2014. <https://aip.scitation.org/doi/abs/10.1063/1.4897702>

Introducción

Hoy en día, la mayoría de los dispositivos de captura de imágenes cuentan con funciones automatizadas para el enfoque y la fusión. Estas funciones facilitan y minimizan los procesos manuales de ajuste de los dispositivos digitales durante la captura de imágenes y garantizan la obtención de imágenes de alta calidad.

El autoenfoco y la fusión de imágenes desempeñan un papel importante en el procesamiento digital de imágenes, donde el objetivo es obtener una imagen lo más semejante posible a la escena real y que además destaque las características o elementos relevantes en la imagen que permitan su posterior análisis.

Los diferentes métodos de autoenfoco buscan obtener el mejor plano focal de un sujeto o escena de manera automática, mientras que los métodos de fusión consisten en tomar elementos de dos o más imágenes e integrarlas en una sola, se obtiene así una nueva imagen fusionada que combina estratégicamente los elementos seleccionados de cada imagen.

En esta tesis se proponen nuevos métodos de autoenfoco y de fusión, ambos trabajan en el dominio de frecuencias, es decir, el procesamiento se basa en el análisis espectral de la imagen.

En esta propuesta el algoritmo de autoenfoco aplica la transformada de Fourier para determinar la medida de enfoque de una imagen. Esta medida de enfoque es calculada para cada imagen de un conjunto de imágenes obtenidas por el método de barrido en Z, es decir, realiza diversas capturas de un mismo objeto variando la distancia z entre el objeto y la lente, ya sea el microscopio, telescopio o cámara. El método aplica la correlación no lineal para determinar cuál es la Imagen Mejor Enfocada (IME) mediante la comparación de las medidas de enfoque de las diferentes imágenes del conjunto.

Este algoritmo aplica la transformada de Fourier para obtener el espectro de cada imagen en el subconjunto de fusión, a cada espectro se le aplica un filtro parabólico para realzar las

características relevantes de cada imagen, posteriormente se seleccionan las mejores características de cada imagen de este subconjunto y se integran en una nueva imagen, esta nueva imagen es el resultado de la fusión.

Con este nuevo método de fusión se obtiene una imagen de mayor calidad que la imagen mejor enfocada, a diferencia de los métodos clásicos que solo fusionan dos imágenes, el algoritmo aquí propuesto tiene la capacidad de fusionar más de dos imágenes, el número de imágenes a fusionar está determinado por un rango de valores cercanos a la mejor medida de calidad de enfoque, correspondiente a la imagen mejor enfocada. Así mismo, se logra realizar esta tarea en un menor tiempo de procesamiento que el de los métodos clásicos.

En esta tesis, el algoritmo de autoenfoco mejora el tiempo de lectura de la imagen de entrada mediante el método de escaneo en espiral de la imagen, esto es, se toman las coordenadas generadas por la ecuación de la espiral y se da lectura sólo a los píxeles de la imagen que pasan a través de dichas coordenadas, los valores muestra se almacenan en un vector con suficiente información para procesar el autoenfoco.

Esta tesis inicia con los *Antecedentes*, una breve descripción de los principales conceptos utilizados, como el autoenfoco, la fusión y las métricas utilizadas para la evaluación de la calidad de la imagen, capítulo 1. La descripción del proceso y la *Metodología* se presenta en el capítulo 2. La evaluación, comparación de los *Resultados* y discusiones son presentados en el capítulo 3. Finalmente, las *Conclusiones* de esta tesis se presentan en el capítulo 4 y al final se listan las principales *Referencias*, capítulo 5 y los *Anexos* capítulo 6.

Planteamiento del problema

En esta tesis se propone un método de enfoque automatizado, *autoenfoco*, y un método de *fusión* que mejore la calidad de una imagen.

El nuevo método de fusión, integra en una sola imagen una selección de imágenes de un conjunto, capturadas a diferente distancia z del objeto, de esta manera se logra mejorar la calidad de la imagen mejor enfocada. Así mismo, se hace la propuesta del desarrollo de un algoritmo computacional que implemente nuevos métodos de autoenfoco y fusión selectivo, AFS, para el análisis de imágenes micro y macroscópicas, **Fig. 1**.

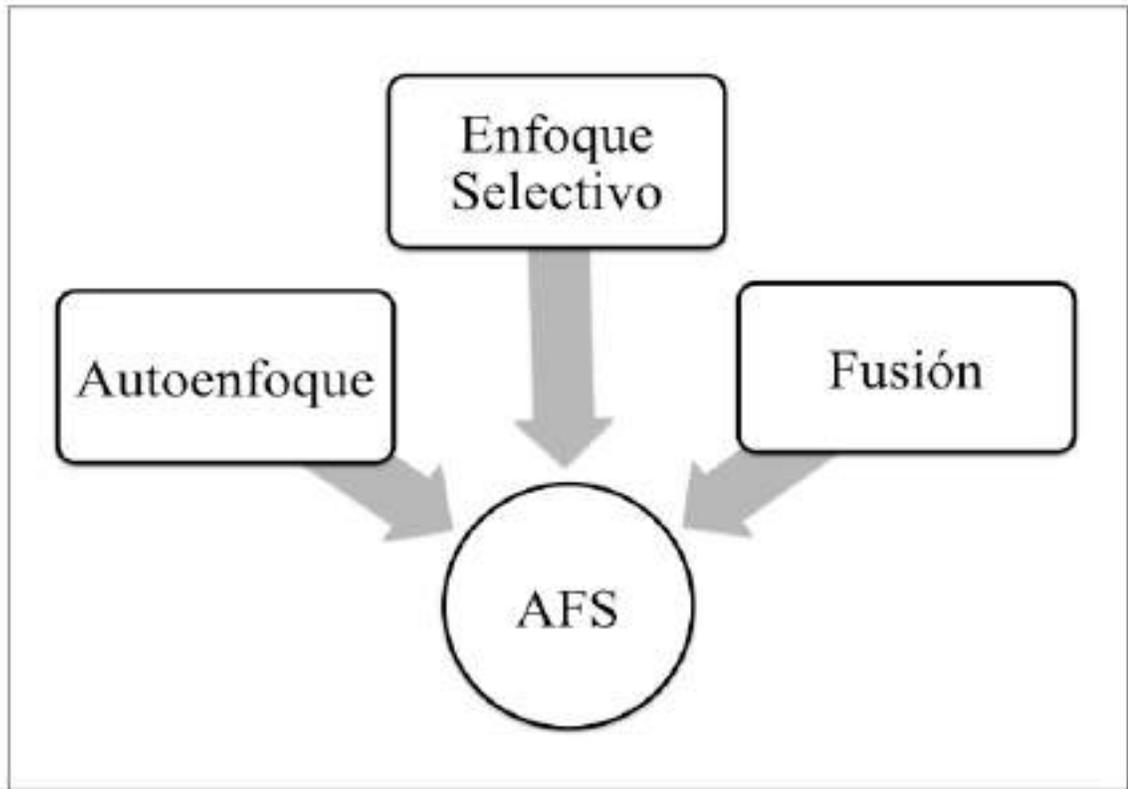


Fig. 1 Descripción del sistema.

Objetivo general

Desarrollar un algoritmo computacional que implemente nuevos métodos para el autoenfoco y fusión selectivo para el análisis de imágenes micro y macroscópicas.

Objetivos específicos

- Analizar diversos algoritmos de autoenfoco y fusión de imágenes.
- Desarrollar un nuevo algoritmo de Autoenfoco.
- Desarrollar un nuevo algoritmo de Fusión.
- Implementar el algoritmo AFS propuesto.
- Evaluar los resultados.

Hipótesis

Se espera que:

- Esta tesis contribuya al procesamiento digital de imágenes automatizado y que sus resultados faciliten el análisis de imágenes micro y macroscópicas.
- Los resultados de los algoritmos de autoenfoco y fusión sean imágenes de alta calidad que satisfagan las métricas de evaluación y conserven los elementos importantes de información de la imagen para facilitar su posterior procesamiento o análisis.
- El algoritmo sea óptimo en ejecución, rendimiento, robustez, economice tiempo y recursos computacionales.

Alcances y limitaciones

La tesis plantea innovar los métodos de enfoque y fusión con la propuesta de nuevos algoritmos, los métodos implementan conceptos de transformada de Fourier, correlación no-lineal (CNL), escaneo espiral y filtro parabólico.

La implementación realiza la demostración de los algoritmos mediante el uso de varios conjuntos de imágenes digitales pre-capturadas con el dispositivo de carga acoplada (CCD, *Charge-Coupled Device*) del microscopio, del telescopio o de la cámara digital. El desarrollo de la solución emplea el lenguaje y bibliotecas de MATLAB.

Justificación

El procesamiento digital de imágenes es una herramienta importante para la investigación, ya que el análisis, clasificación e identificación de elementos visuales son objeto de estudio, estos procesos solo pueden llevarse a cabo con imágenes nítidas que representen fielmente la escena u objeto que capturan del mundo real.

Por lo anterior, se puede comprender que parte importante del procesamiento de imágenes es contar con una imagen correctamente enfocada y de ser posible mejorar estas características de enfoque mediante la fusión de imágenes.

Capítulo 1

Antecedentes

1.1 Estado del arte

Históricamente las técnicas de Procesamiento Digital de Imágenes (PDI) tuvieron sus inicios entre los años 1960 y 1970. Las primeras aplicaciones PDI evolucionaron en conjunto con el desarrollo de nuevas tecnologías en las áreas de la computación, a inicios de 1960 las computadoras ya eran suficientemente poderosas para soportar tareas significativas de PDI [González y Woods, 2008]. En 1960, Leica patentó los primeros sistemas de autoenfoco. En 1976 presentó durante el Photokina la cámara Leitz Correfot, una cámara que utilizaba un sistema de confirmación de enfoque parecido a los sistemas de detección de fase que se emplean en la actualidad y en 1978 presentan una cámara réflex de objetivo único (SLR, *Single Lens Reflex*) con autofoco completamente operativo [Leica-Microsystems, 2013]. En 1964 las fotografías de la luna fueron transmitidas por el *Ranger 7* y procesadas para corregir varios tipos de distorsión inherentes a la imagen de la cámara de televisión a bordo [González y Woods, 2008].

En el aspecto teórico diferentes autores han contribuido con diversos métodos para determinar una medida de enfoque. En 1968 en su artículo “*Focusing*”, B.K.P. Horn, [Horn, 1968] describe un método de enfoque automático de un vidisector (TVC). Tenenbaum en 1970, propuso en calcular la magnitud de los gradientes espaciales utilizando el operador Sobel [Tenenbaum, 1970], esta medida realiza la convolución de la imagen en las direcciones vertical y horizontal, y deriva dos valores que representan las magnitudes del gradiente en la región de interés de la imagen. Los primeros enfoques computarizados, han incorporado criterios de enfoque como una fase inicial de búsqueda de

enfoco [Jarvis, 1976], en la que estos criterios consisten en una búsqueda caliente/frío, una puntuación que aumenta o disminuye y cambia la dirección, sube o baja, hacia el mejor enfoque. Groen en 1985 propone métodos basados en la varianza de la imagen, considera las variaciones de las intensidades de los píxeles y emplea una función para amplificar las diferencias con respecto a la media de intensidad de la imagen [Groen, 1985]. En 1988 Vollath propone las medidas de enfoque de Vollath F4 y F5 estas medidas se basan en la función de autocorrelación, la varianza y la desviación estándar.

Eric Krotkov, en su trabajo sustentado por la fuerza aérea, DARPA y ARMY [Krotkov, 1987] presenta soluciones a dos problemas relacionados con el autoenfoco, el primero es determinar mejor la posición del motor de enfoque que proporciona el enfoque más nítido en un punto de objeto a una distancia desconocida; y el segundo cómo calcular la distancia a un punto de un objeto bien enfocado. En su artículo emplea una función de criterio para medir la nitidez del enfoque, emplea una técnica de búsqueda de Fibonacci para localizar de forma óptima el modo de la función criterio. Así también describe que el desenfoque puede verse como una atenuación de altas frecuencias espaciales y manifiesta que un método basado en maximizar la magnitud del gradiente de intensidad resulta superior a otros métodos.

En su trabajo, Santos [Santos, et al, 1997] describe una evaluación sistemática de diversas funciones de autoenfoco empleadas para estudios de citometría analítica de imagen fluorescente de núcleos contrastados. Santos A., considera que el enfoque es el primer paso en el análisis de imágenes microscópicas. Fueron evaluadas treinta funciones empleando procedimientos cualitativos y cuantitativos. Sus resultados muestran que las funciones basadas en medidas de correlación tienen el mejor desempeño para este tipo de imágenes.

En el trabajo de Subbarao [Subbarao, 1992] se describe un método para seleccionar la medida de enfoque óptima con respecto al ruido de nivel de grises de un conjunto dado de medidas de enfoque. El método se basa en dos métricas para estimar la sensibilidad al ruido

de las diferentes medidas de enfoque. La primera métrica, la Medida de Incertidumbre de Autofoco, establece la relación entre el ruido de nivel de gris y el error resultante en la posición de la lente para el enfoque automático. La segunda métrica, *Autofocusing Root-Mean-Square Error*, es una métrica mejorada estrechamente relacionada con la anterior. La selección de la medida óptima de enfoque de un conjunto dado implica el cálculo de todas las medidas de enfoque en el conjunto.

Mario Bueno-Ibarra presenta un nuevo algoritmo de autoenfoco [Bueno-Ibarra, 2005a] basado en la transformada de Fourier unidimensional y en la correlación de Pearson. Usando un conjunto de varias imágenes a diferentes distancias Z de una muestra biológica. El algoritmo utiliza la transformada de Fourier para obtener y extraer el contenido de frecuencia de imagen en un patrón de vector previamente especificado y encontrarlo en cada imagen capturada; compara estos vectores de frecuencia con el vector de frecuencia de una imagen de referencia, y se determina la mejor imagen en el foco a través de la correlación de Pearson. El autor propone como trabajo futuro se realicen estudios para diseñar y probar nuevos patrones de procesos de escaneo, basados en la transformación de Fourier, para su incorporación en el algoritmo propuesto.

Megan J. Russell [Russell, 2007], evalúa los algoritmos de autoenfoco para microscopia aplicada a la tuberculosis, probando tres diferentes medidas de enfoque, los métodos, energía de la imagen Laplaciana [Subbarao, 1992], la varianza del histograma logarítmico [Forero, 2004] y el primer orden derivado de Gauss [Geusebroek, 2000]. Concluye que la energía de la imagen Laplaciana medida de enfoque brinda buenos resultados si se utiliza en diapositivas con contenido de media y baja densidad. Sugiere desarrollar una medida, basada en la textura de la imagen, para indicar el nivel de densidad del contenido en una serie de imágenes para el juicio automatizado de la calidad probable de la posición de enfoque óptimo. Otro trabajo relacionado es el de Mateos, que realiza una evaluación comparativa de algoritmos de autofocus para un sistema en tiempo real para la detección automática de *Mycobacterium tuberculosis* [Mateos-Pérez, et al, 2012].

En 2009 Zeder, propone una función de transformación llamada histograma ponderado y la medida de enfoque está dada por la suma de todos los valores de gris transformados [Zeder, 2009].

Finalmente, por mencionar una aplicación de interés general, un equipo de investigación de Ishikawa, Japón desarrolló una guía óptica para vehículos terrestres autónomos basada en una cámara de infrarrojos y la técnica de enfoque para detectar los objetos a evadir y establecer la trayectoria de la unidad [Akiyama, et al, 2013]. El procesamiento de imagen de enfoque automático de su trabajo emplea la técnica de transformada de onda de Daubechies para determinar la mejor posición de enfoque, la cual se transforma en la distancia del objeto. Realizaron los experimentos con un coche real en la carretera, para evaluar la capacidad de detección de la unidad de enfoque automático infrarrojo.

En cuanto a la fusión, históricamente los campos en los que mayormente se aplica es en medicina, en los sistemas de información geográfica y en la astronomía, usualmente la fusión combina la información suministrada por distintos sensores para formar una imagen con determinadas características mejoradas [González y Woods, 2008]. La fusión puede crearse a partir de varias imágenes de un mismo formato o de múltiples formatos de imagen, por ejemplo, la imagen por resonancia magnética (MRI, *Magnetic Resonance Imaging*), tomografía por emisión de positrones (PET, *Positron Emission Tomography*), tomografía por emisión de fotón único (SPECT, *Single Photon Emission Computed Tomography*), y la tomografía computarizada (CT, *Computed Tomography*), [González y Woods, 2008]. En la actualidad, las técnicas de fusión a nivel de píxel disponen de un gran número de metodologías y algoritmos. Las más populares son las basadas en operaciones aritméticas como la transformada Promedio simple [Naidu, 2008], Máximo simple [Rani, 2013], y las basadas en la transformada Análisis de Componentes Principales, (PCA, *Principal Component Analysis*), [Naidu, 2008] y [Desale, 2013]. La Transformada discreta de coseno, (DCT, *Discrete Cosine Transform*), [Naidu, 2012][Desale, 2013]. La relación de la pirámide de paso bajo (ROLP, *Ratio of Low-Pass Pyramid*) [Toet, 1989], y la fusión basada en Wavelets [Pajares, 2004], entre otras técnicas.

Los primeros campos de aplicaciones del autoenfoco y fusión de imágenes fueron en la áreas de investigación médica, militar, exploración espacial y terrestre. Actualmente las aplicaciones se extienden a un amplia gama de áreas de investigación, como la industria, aeronáutica, automotriz, seguridad, mediciones ambientales, astronomía, microscopía, entre otras; así también se aplica a dispositivos de uso personal o doméstico.

Las tecnologías de autoenfoco y fusión de imágenes continúan su evolución concurrentemente con los avances en la miniaturización de componentes, almacenamiento masivo y sistemas de visualización, los cuales son indispensables para el PDI. Así mismo, se desarrollan nuevos materiales, teorías y métodos gracias a las aportaciones de las áreas de las ciencias, principalmente de la física óptica.

1.2 Procesamiento digital de imágenes, PDI

El estudio e investigación de nuevos métodos de autoenfoco y fusión de imágenes pertenece al campo de investigación del procesamiento digital de imágenes, PDI.

El PDI es un conjunto de técnicas que se aplican para mejorar la calidad de las imágenes, tiene como objetivo extraer información útil de una imagen con la finalidad de realizar alguna tarea específica ya sea de transformación, restauración o mejoramiento para aplicar algún tipo de análisis de la imagen [González y Woods, 2008].

El análisis de imágenes es la extracción de propiedades y características de la imagen que permiten diversas tareas, como la búsqueda de información específica en una imagen, reconocimiento de patrones, identificación de objetos, clasificación, conteo, ordenamiento, visualización, entre otros [González y Woods, 2008].

Algunas de las etapas del PDI, básicamente son: adquisición de la imagen, mejora de la imagen, extracción de información e interpretación, **Fig. 2**.

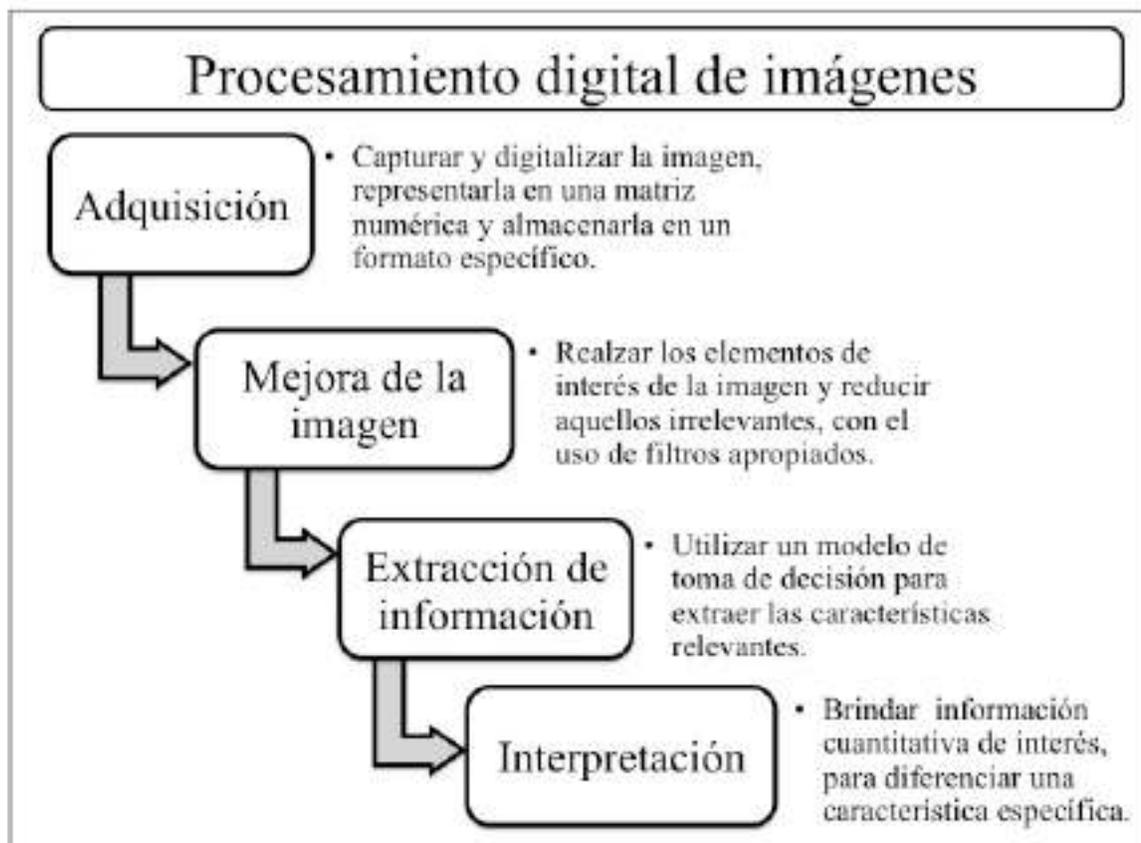


Fig. 2 Etapas básicas del procesamiento digital de imágenes.

Los métodos de autoenfoco y fusión tienen como objetivo obtener como resultado imágenes de alta calidad, requeridas para diferentes aplicaciones, ya sea para la interpretación humana, el análisis de datos de la imagen, la percepción de máquina autónoma o visión artificial, entre otras.

1.2.1 Adquisición de la imagen digital

La adquisición de imágenes puede ser tan simple como realizar la lectura de una imagen que ya está en forma digital en algún dispositivo de almacenamiento, también se puede capturar mediante un dispositivo óptico digital o adquirirse analógicamente mediante algún sensor de iluminación o de energía electromagnética, requiriéndose en este último caso su posterior digitalización, dado que las etapas posteriores requieren de imágenes digitales.

Una imagen digital está compuesta de un número finito de elementos, cada uno de los cuales tiene una ubicación particular y un valor, a estos elementos se les llama píxeles; el pixel es el término más utilizado para designar a los elementos de una imagen digital.

Una imagen de dimensiones $M \times N$ puede definirse como una función bidimensional, $M \times N$ donde x e y son coordenadas espaciales, la amplitud f es la intensidad o nivel de gris de la imagen en ese punto, donde los valores de f son todas cantidades finitas y discretas, **Fig. 3**.

Para procesar la imagen por medio de la computadora es preciso someter a la $f(x,y)$ a un proceso de discretización tanto en las coordenadas como en la intensidad, a este proceso se le denomina digitalización, **Fig. 4**, una imagen digital puede considerarse como una matriz $M \times N$ cuyos índices de fila y columna identifican un punto de la imagen, denominado elemento de la imagen o pixel, y el valor del correspondiente elemento de la matriz indica el nivel de gris o de color de ese punto, **Fig. 5**.

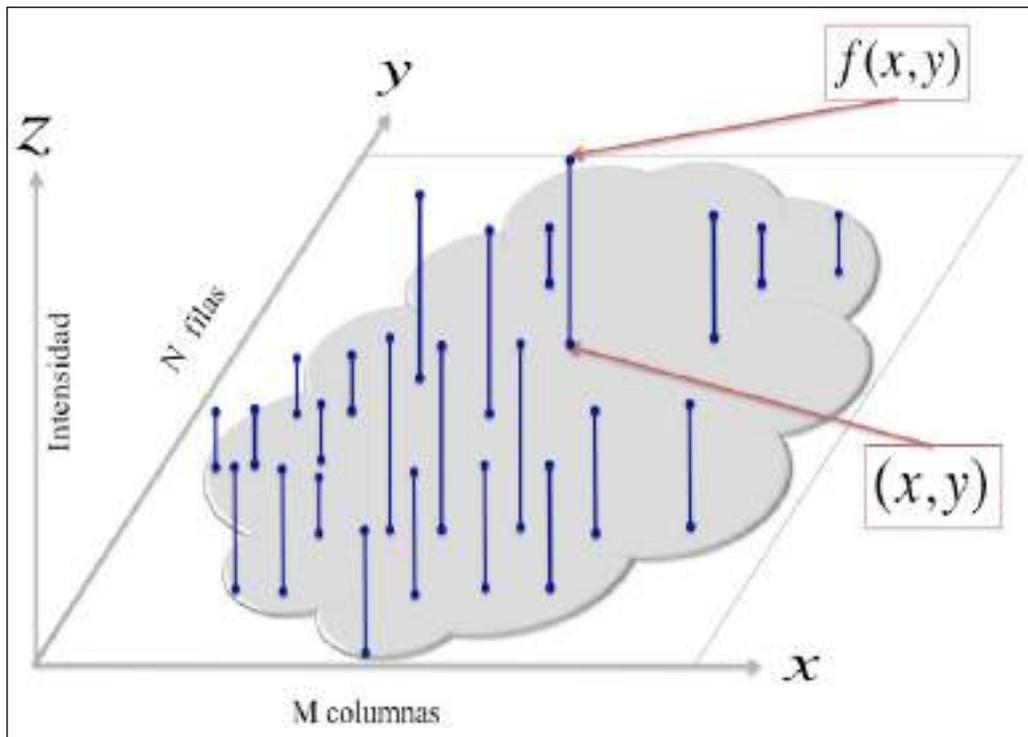


Fig. 3 Descripción de los elementos de una imagen.

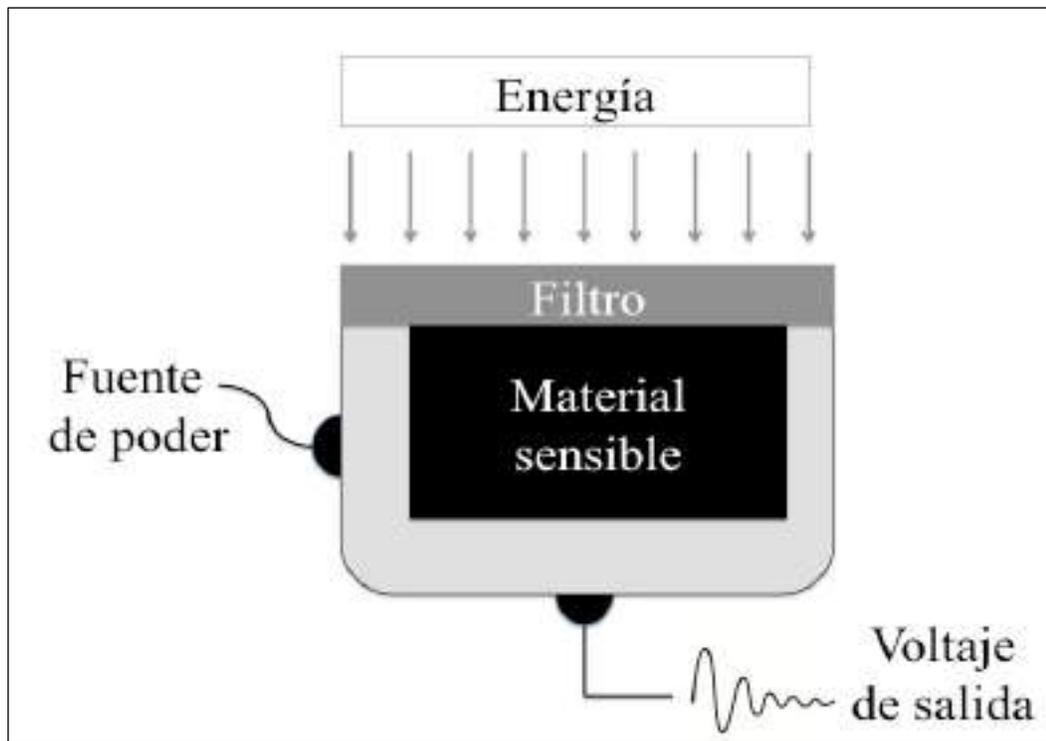


Fig. 4 Sistema de captura.

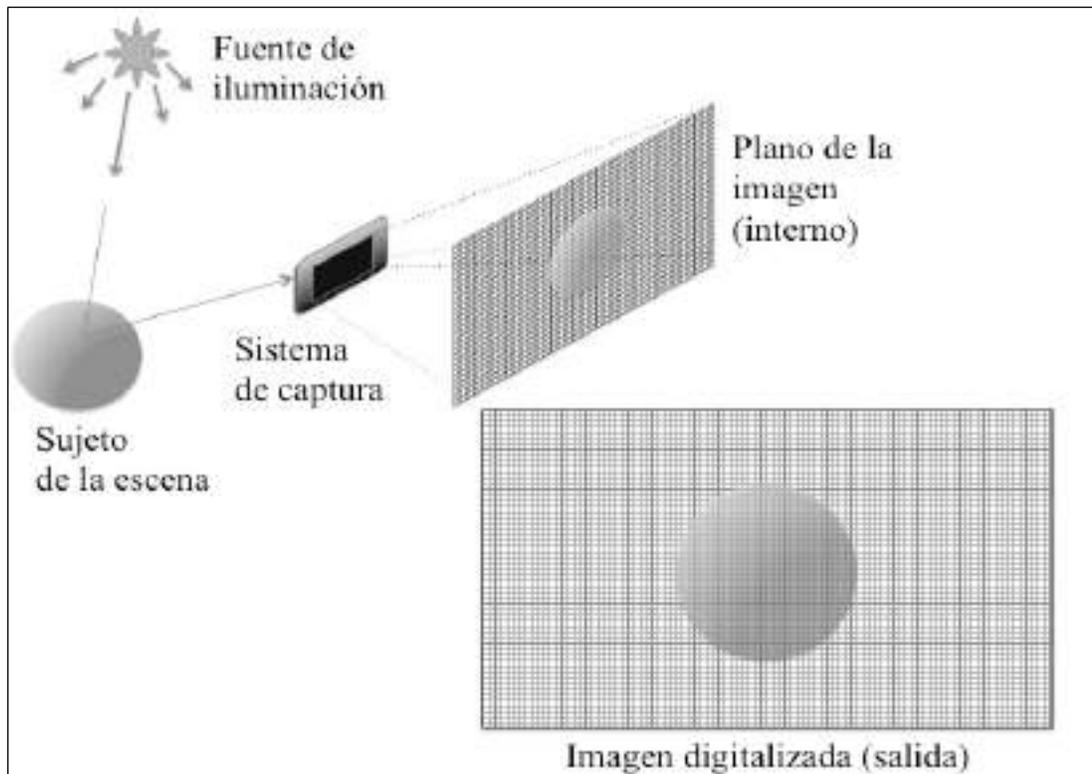


Fig. 5 Descripción del proceso de adquisición de la imagen digital.

La imagen digital puede requerir de un pre-procesamiento que también se lleva a cabo en esa etapa de adquisición, que puede consistir en el redimensionamiento y escalamiento de la imagen [González y Woods, 2008].

1.2.2 Mejora de la imagen digital

Existen diversas técnicas para el mejoramiento de la imagen, resaltar aspectos deseados, eliminar ruido, mejorar contraste, dependiendo del dominio problema serán las técnicas que se apliquen. Una de las técnicas consiste en hacer una transformación de la imagen de manera que se acentúen o disminuyan ciertas características de la imagen, para ello se utilizan los filtros.

Existen básicamente dos tipos de filtrado, en dominio espacial y en el dominio de la frecuencia. Los filtros en el dominio espacial realizan operaciones directamente con los píxeles de la imagen, mientras que los filtros en el dominio de las frecuencias emplean el espectro de la imagen con la representación global de los tonos contra número de frecuencia de éstos.

Básicamente, hay tres tipos de filtros:

- **Filtros de paso bajo**, conserva las bajas frecuencias y atenúa o elimina las altas frecuencias.
- **Filtros de paso alto**, conserva las altas frecuencias y atenúa o elimina las bajas frecuencias.
- **Filtros de banda**, atenúa las altas y bajas frecuencias, se conservan las frecuencias que se encuentren en una banda determinada.

Los filtros en el dominio espacial suelen emplear máscaras representadas en matrices con valores específicos con las que realizan las operaciones con los píxeles directamente. Algunos filtros espaciales son Roberts, Sobel, Laplaciano Gaussiano, por nombrar algunos de los se describen en la siguiente sección. Algunos filtros en el dominio de la frecuencia más comunes son el filtro Ideal, el filtro Butterworth, filtro Gaussiano entre otros. En esta tesis se utiliza un filtro parabólico que es un tipo de filtro pasa altas.

1.3 Autoenfoco y fusión

La aplicación de técnicas de autoenfoco y fusión asumen un papel relevante dentro del procesamiento de imágenes. El autoenfoco tiene como finalidad la obtención del mejor plano focal entre un conjunto de imágenes capturadas a diferentes distancias focales.

La fusión de imágenes permite mostrar las características visuales más representativas que se encuentran a diferentes planos focales y fusionarlas en una sola imagen de mayor calidad.

Los algoritmos de autoenfoco y fusión deben cumplir las siguientes características:

- Ser óptimos y eficientes,
- procesar grandes cantidades de información en el menor tiempo posible y
- obtener una imagen de alta calidad que realce las principales características del objeto en la imagen para su posterior análisis, identificación o clasificación.

1.4 Análisis de algoritmos de autoenfoco

Se refiere a la capacidad de una cámara o sistema óptico para enfocar de forma automatizada [Álvarez-Borrego, et al, 2008]. El objetivo del enfoque es encontrar el mejor plano focal de una imagen por métodos clásicos o automatizados.

El método clásico para enfocar es el ajuste manual, el cual se realiza mediante el desplazamiento del objetivo a través del eje óptico, se comparan visualmente las imágenes y se selecciona la mejor imagen enfocada de acuerdo al criterio personal de percepción visual; esta selección visual puede ser relativa debido a imperfecciones particulares que pueden existir en la visión de cada persona.

Un sistema de autoenfoco está compuesto por sensores, una lente, una cámara o CCD que percibe al sujeto, un sistema de control o computadora que calcula el nivel de enfoque actual, un motor para mover el sistema de lentes para mejorar la distancia de enfoque; las acciones que realizan estos elementos se repiten cíclicamente hasta que se encuentra el mejor enfoque para capturar la imagen, **Fig. 6**.

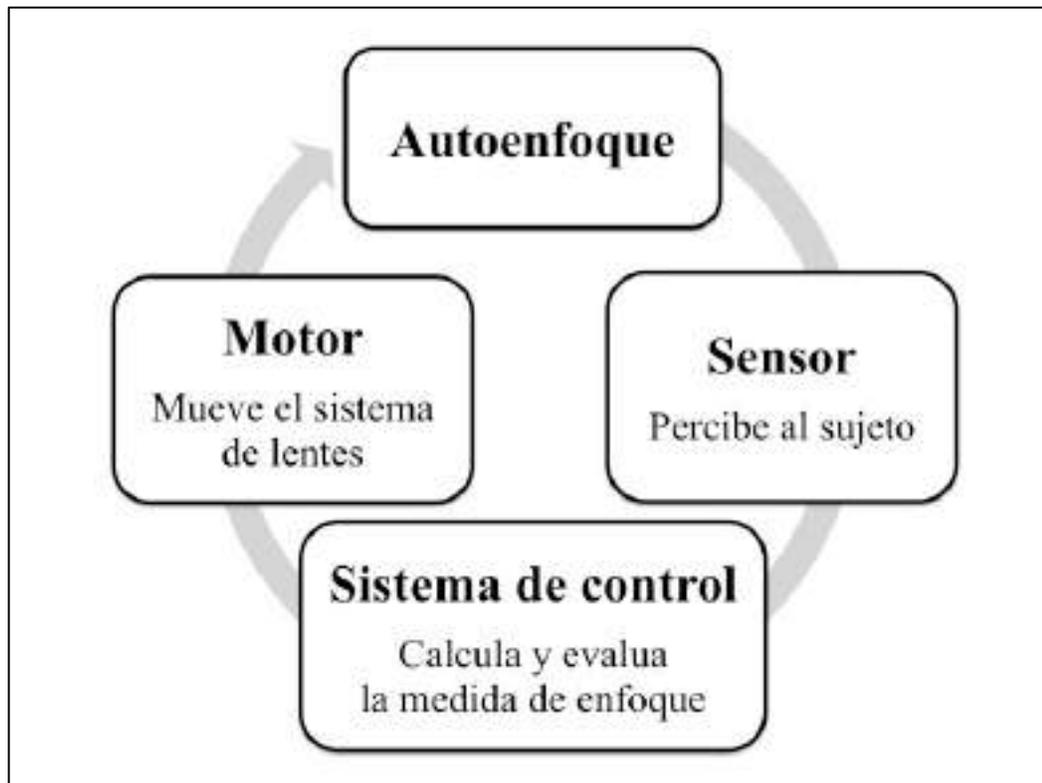


Fig. 6 Proceso de autoenfoco.

Los métodos de enfoque automático pueden ser activos, pasivos o híbridos. Un sistema activo clásico emite ondas sonoras o de luz (ultrasonido, microondas, infrarrojos) para detectar la distancia del sujeto desde la cámara. Los sistemas activos típicos que utilizan sistemas de infrarrojos pueden utilizar una variedad de técnicas para juzgar la distancia como la triangulación, la cantidad de luz infrarroja reflejada por el sujeto o el lapso de tiempo [Brown, 2000].

El autoenfoco pasivo puede lograrse mediante detección de fase o medición de contraste. La detección de fase se basa en la triangulación de la distancia del sujeto, se utilizan dos sensores que evalúan la luz que pasa a través de dos lados extremos de la lente a través del mismo objetivo. La distancia está determinada por la diferencia de las imágenes captadas por ambos sensores, **Fig. 7**, [Mansurov, 2012].

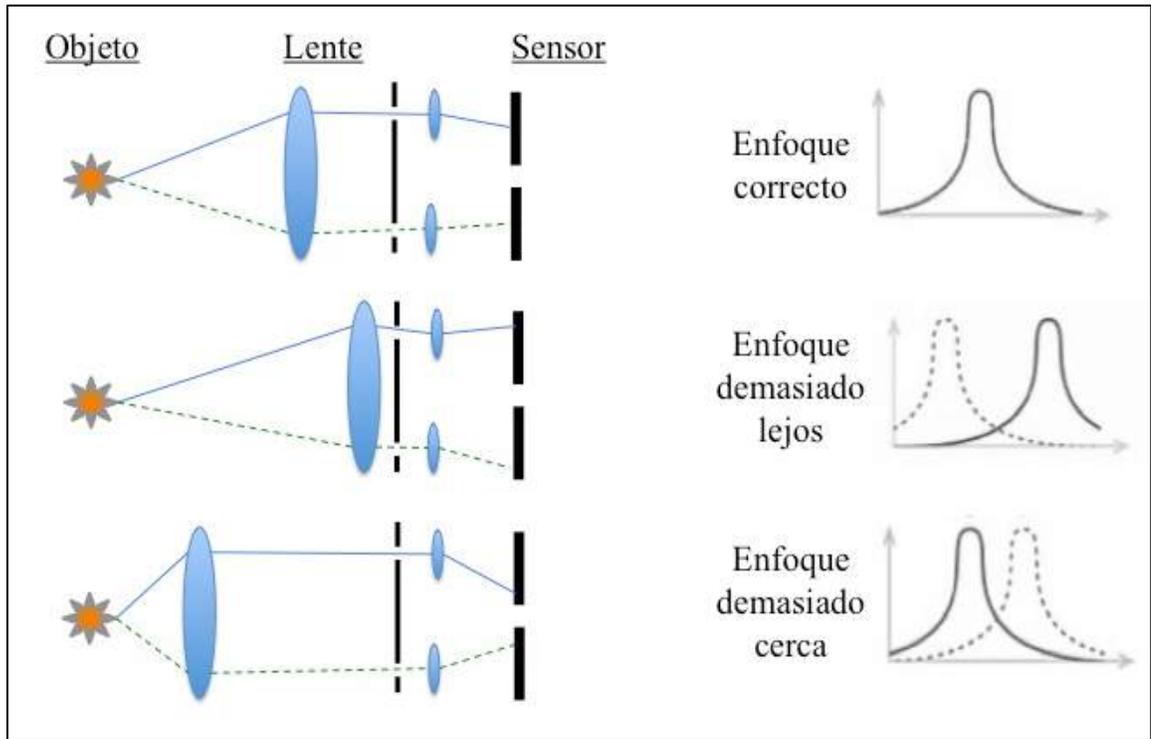


Fig. 7 Enfoque de la imagen, captada a diferentes distancias.

Para calcular el contraste de la imagen se comparan los valores de intensidad de cada pixel en una banda de píxeles proporcionada por el CCD. **Fig. 8.**

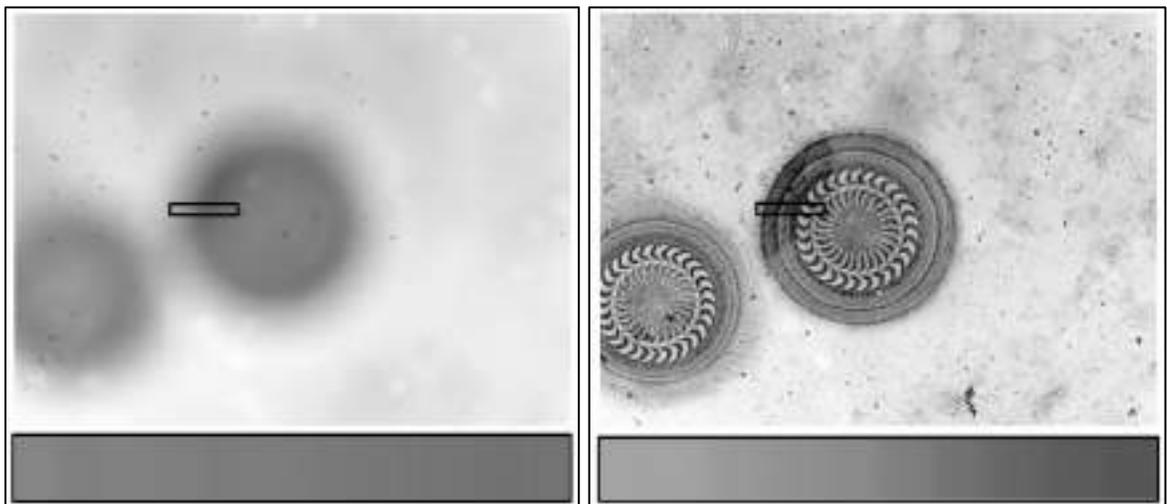


Fig. 8 Comparación del contraste entre las bandas de píxeles de una sección de la imagen desenfocada y de la IME, respectivamente.

Existen diversos algoritmos de autoenfoco para obtener el mejor enfoque, se han realizado diversos estudios comparativos que aplican autoenfoco a diferentes áreas de aplicación y diferentes objetivos. Para este estudio se seleccionó un conjunto de algoritmos para su análisis, se consideran aquellos utilizados más popularmente para el procesamiento de imágenes [Groen, 1985], [Krotkov, 1988], [Bueno-Ibarra, 2005c], [Russell, 2007] [Mateos-Pérez, et al, 2012].

Algunos algoritmos que controlan de forma automatizada estos procesos se basan en varios aspectos teóricos del autoenfoco tradicional. En esta tesis se analizaron diez modelos comunes de autoenfoco: la energía de la imagen Laplaciana [Subbarao, 1992], Tenengrad y Tenengrad absoluto [Tenenbaum, 1970], varianza y varianza normalizada [Forero, et al, 2004], Vollath's F4 y Vollath's F5 [Vollath, 1988], filtro gaussiano [Groen, 1985], histograma logarítmico [Hilsenstein, 2005] e histograma ponderado [Zeder, 2009], y el autoenfoco por correlación de Pearson y Fourier [Bueno-Ibarra, 2005c].

A continuación se describen los algoritmos de autoenfoco con los cuales se realizaron las pruebas para el análisis.

1.4.1 Energía de la imagen Laplaciana

Esta métrica hace uso de la convolución de la imagen con la máscara discreta de Laplace [Subbarao, 1992], utiliza el siguiente *Kernel* de convolución bidimensional:

$$a(i, j) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$F_L = \sum_{i,j} [g(i-1, j) + g(i+1, j) + g(i, j-1) + g(i, j+1) - 4g(i, j)]^2, \quad (1)$$

donde $g(i, j)$ es el nivel de gris del pixel (i, j) en la imagen g .

1.4.2 Tenengrad

Este algoritmo convoluciona una imagen con operadores Sobel y suma el cuadrado de todas las magnitudes mayores que un umbral [Tenenbaum, 1970] [Schlag, et al, 1983]. La máscara de convolución está dada por:

$$S = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, S' = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad (2)$$

donde S y S' son las máscaras de convolución de Sobel. La función de medida Tenengrad está dada por:

$$F_T = \sum_{i,j} [g(i,j) \otimes S]^2 + [g(i,j) \otimes S']^2, \quad (3)$$

donde \otimes representa el operador de convolución, $g(i,j)$ es el nivel de gris en el pixel (i,j) de la imagen g .

1.4.3 Tenengrad absoluto

Esta función de medida es similar a la anterior, pero en este caso se toma el valor absoluto de los coeficientes del gradiente con el fin de reducir el costo computacional [Jarvis, 1676]. La función Tenengrad absoluto F_{AT} está dada por:

$$F_T = \sum_{i,j} |g(i,j) \otimes S| + |g(i,j) \otimes S'|, \quad (4)$$

donde \otimes representa el operador de convolución, $g(i,j)$ es el nivel de gris en el pixel (i,j) de la imagen g , S es la máscara de convolución Sobel y S' su respectiva transpuesta.

1.4.4 Varianza

Esta medición calcula la variación de intensidades de los píxeles y amplifica las grandes diferencias de la media de intensidad de la imagen [Groen, 1985]. La función de la varianza F_{VAR} está dada por:

$$F_{VAR} = \frac{1}{MN} \sum_{i,j} [g(i,j) - \bar{g}]^2, \quad (5)$$

donde M y N son la anchura y altura respectivamente de la imagen en píxeles, $g(i,j)$ es el nivel de gris en el píxel (i,j) en la imagen g y \bar{g} es el promedio de píxeles en la imagen.

1.4.5 Varianza normalizada

Se basa en la ecuación anterior pero además se normaliza utilizando el brillo medio de la imagen [Groen, 1985]. La función de medida de varianza F_{NVAR} está dada por:

$$F_{NVAR} = \frac{1}{MN\bar{g}} \sum_{i,j} [g(i,j) - \bar{g}]^2, \quad (6)$$

donde M y N son la anchura y altura respectivamente de la imagen en píxeles, $g(i,j)$ es el nivel de gris en el píxel (i,j) en la imagen g y \bar{g} es el promedio de píxeles en la imagen.

1.4.6 Vollath's F4

Este algoritmo utiliza una función que se basa en la autocorrelación de la imagen, es robusto contra el ruido y eficiente computacionalmente, caracterizado por tener un rango de trabajo particularmente amplio [Vollath, 1988]. La función F_{VLT4} de medición de Vollath F4 está dada por:

$$F_{VLT4} = \sum_{i=1}^{M-1} \sum_{j=1}^N g(i,j) \cdot g(i+1,j) - \sum_{i=1}^{M-2} \sum_{j=1}^N g(i,j) \cdot g(i+2,j), \quad (7)$$

donde M y N son la anchura y altura respectivamente de la imagen en píxeles, $g(i,j)$ es el nivel de gris en el píxel (i,j) en la imagen g .

1.4.7 Vollath's F5

Este método es similar al Vollath F4, pero este suprime las frecuencias altas, se basa en la función de desviación estándar y es capaz de lograr la máxima nitidez [Vollath, 1988]. La función F_{VLT5} de medición de Vollath F5 está dada por:

$$F_{VLT5} = \sum_{i=1}^{M-1} \sum_{j=1}^N g(i, j) \cdot g(i+1, j) - MN\bar{g}^2, \quad (8)$$

donde M y N son la anchura y altura respectivamente de la imagen en pixeles, $g(i, j)$ es el nivel de gris en el pixel (i, j) en la imagen g .

1.4.8 Filtro Gaussiano

Esta función se basa en un filtro de degradado; se deriva de la convolución de la imagen con una derivada Gaussiana de primer orden [Geusebroek, 2000]. La función F_{GSS} de filtro de Gauss está dada por:

$$F_{GSS} = \frac{1}{MN} \sum_{i,j} [g(i, j) \otimes G_i(i, j, \sigma)]^2 + [g(i, j) \otimes G_j(i, j, \sigma)]^2, \quad (9)$$

donde \otimes representa el operador de convolución, $g(i, j)$ es el nivel de gris del pixel (i, j) en la imagen g , $G_i(i, j, \sigma)$ y $G_j(i, j, \sigma)$ son la derivada Gaussiana de primer orden con escala de σ en la derivada con dirección i y j . El valor de σ depende de los elementos de la imagen; para estructuras tipo barra este valor puede ser determinado por $\sigma \approx \left(\frac{d}{2}\right)\sqrt{3}$

donde d representa el grosor de la barra.

1.4.9 Histograma Logarítmico

Este método que utiliza el brillo de la imagen a través de una transformación logarítmica del histograma de niveles de gris de la imagen [Hilsenstein, 2005]. La función F_{LHS} de medición del histograma logarítmico viene dada por:

$$F_{LHS} = \sum (l - E_{\log}(l))^2 \cdot \log(p_l), \quad (10)$$

donde

$$E_{\log}(l) = \sum_l l \cdot \log(p_l), \quad (11)$$

y l es el nivel de gris en el histograma y p_l es la probabilidad para cada nivel de gris.

1.4.10 Histograma ponderado

Este procedimiento se basa en un histograma de la imagen ponderada sin introducir un umbral constante. La suma de todos los valores de grises transformados se utiliza como medida de enfoque, con la cual se determina la imagen mejor enfocada [Zeder, 2009].

$$F_{HIP} = \sum_l \sqrt[5]{h(l)} \cdot l^5 \cdot 10^{-15}. \quad (12)$$

El procedimiento multiplica la raíz quinta del número de píxeles de cada nivel de gris $h(l)$ por la quinta potencia de este nivel de gris l multiplicado por 10 a la potencia -15.

1.4.11 Autoenfoco por correlación de Pearson y Fourier

El algoritmo procesa una cantidad Q de vectores V_q para cada imagen $f(i, j)_k$ del conjunto de imágenes, cada vector procesado es identificado por el subíndice $q = 1, \dots, Q$, estos vectores V_q se distribuyen en la imagen de forma equidistante, de manera que forman un patrón de exploración en cada imagen analizada.

La ecuación que construye el único vector para cada imagen puede expresarse en la forma:

$$FSV_K = |H_1(f)|_k^2 \cup |H_2(f)|_k^2 \cup \dots \cup |H_q(f)|_k^2, \quad (13)$$

donde $|H_1(f)|_k^2, |H_2(f)|_k^2, \dots, |H_q(f)|_k^2$ se obtienen calculando el espectro de potencia de Fourier para cada vector V_q , estas magnitudes contienen las frecuencias altas y bajas de estos vectores y \cup denota la unión o concatenación de vectores.

Cada vector FSV_k individual se compara mediante la correlación de Pearson con el vector FSV_1 de la imagen de referencia, la cual puede ser la primera imagen o la última del conjunto la cual está más fuera de foco. La mejor imagen enfocada f_{BF} se obtiene por medio del coeficiente de correlación de Pearson r . Se crea un vector $N(r_K)$ que contiene los valores correspondientes a los coeficientes de correlación de Pearson para cada imagen k -ésima del conjunto, luego se busca el valor mínimo del vector resultante:

$$f_{BF} \{f_K \text{ donde } MIN(N(r_K)), \quad (14)$$

donde $N(r_K)$ es el vector de coeficientes r_K de valores normalizados con respecto al valor máximo obtenido, calculado a partir de la comparación de cada imagen f_k del conjunto con la imagen de referencia. Por lo tanto, la mejor imagen enfocada comparada con la primera imagen, el valor del coeficiente de correlación de Pearson r , será cercana a 0.0, lo que significa que entre ambas imágenes no existe una fuerte correlación, es decir, no se parecen [Bueno-Ibarra, 2005c].

En la tabla siguiente, **Tabla 1**, se enuncian las principales características de los algoritmos de autoenfoco descritos anteriormente.

Tabla 1. Características principales de los algoritmos de autoenfoque.

	Mascara	Convolución	Espacial	Frecuencia	Media	Correlación	Filtro	T. Fourier
Energía de la imagen Laplaciana	✓	✓	✓					
Tenengrad	✓	✓	✓					
Tenengrad absoluto	✓	✓	✓					
Varianza			✓		✓			
Varianza normalizada			✓		✓			
Vollath's F4			✓			✓		
Vollath's F5			✓			✓		
Filtro Gaussiano		✓	✓		✓		✓	
Histograma Logarítmico			✓					
Histograma ponderado			✓					
Autoenfoque por correlación de Pearson y Fourier				✓		✓		

1.5 Análisis de algoritmos para la fusión

El objetivo de la fusión de imágenes es combinar las características visuales más representativas de dos o más imágenes. La fusión multi-foco se realiza con capturas de

imágenes, una misma escena en diferentes planos focales, las cuales se fusionan en una única imagen, que resulta en una nueva imagen de calidad superior. Los métodos de fusión deben preservar información válida y útil existente en la imagen de entrada y evitando cualquier distorsión en el proceso de fusión.

Los métodos de fusión se pueden clasificar en dos grupos: el método de fusión del dominio espacial (SD, *Spatial Domain*) y el método del dominio de transformación (TD, *Transform Domain*) [Baraiya, 2014]. En el método de dominio espacial se trata directamente con los píxeles de la imagen, mientras que en el método del dominio de transformación la imagen se transfiere al dominio de frecuencia usando, por ejemplo, la transformada de Fourier de la imagen. Algunos ejemplos de métodos SD son: el promedio simple, (SA, *Simple Average*) [Naidu, 2008]; simple máximo, (SM, *Simple Maximum*) [Rani, 2013]; análisis de componentes principales, (PCA, *Principal Components Analisis*) [Naidu, 2008] y [Desale, 2013]; entre otros. Algunos ejemplos de los métodos TD son la Transformada de Coseno Discreta, (DTC, *Discrete Cosine Transform*) [Ahmed, 1974] y [Vassil, 2004], la Transformada Discreta de Ondoletas, (DWT, *Discrete Wavelets Transform*) [Desale, 2013].

1.5.1 Promedio simple

Este método de fusión se considera el método más simple de fusión de imágenes [Naidu, 2008] y consiste en tomar la intensidad media de los píxeles correspondientes de cada imagen de entrada. La imagen de fusión resultante está dada por:

$$F_{SA}(i, j) = \frac{g_1(i, j) + g_2(i, j)}{2}, \quad (15)$$

donde g_1 y g_2 son las imágenes de entrada que están siendo fusionadas en F_{SA} y $g_1(i, j)$ y $g_2(i, j)$ son las intensidades del píxel (i, j) de cada imagen de tamaño $M \times N$, donde $0 \leq i \leq M - 1$ y $0 \leq j \leq N - 1$.

Este método de fusión hace un procesamiento espacial y podemos decir que no hay mejora en la calidad de la imagen debido a que el hecho de promediar los valores de los píxeles atenúa los contornos de la imagen.

1.5.2 Mximo simple

Este mtodo realiza la fusin seleccionando los pxeles de intensidad mxima correspondientes a ambas imgenes de entrada [Rani, 2013]. La imagen F_{SM} que resulta de la fusin est dada por:

$$F_{SM}(i, j) = \sum_{i=0}^{M-1} \sum_{j=0}^N \text{MAX}[g_1(i, j), g_2(i, j)], \quad (16)$$

donde $g_1(i, j)$ y $g_2(i, j)$ son las intensidades del pixel (i, j) de las imgenes de entrada.

1.5.3 Analisis de componente principal

El PCA es un mtodo que reduce los conjuntos de datos multidimensionales en dimensiones inferiores, es decir, calcula una descripcin compacta y ptima del conjunto de datos. El objetivo es reducir el nmero de variables del conjunto de datos conservando la mayor parte de la variabilidad original de los mismos. Los pesos para cada imagen fuente se determinan usando el vector propio correspondiente al valor propio ms grande de la matriz de covarianza de cada imagen fuente [Naidu, 2008].

Sea X un vector al azar d -dimensional, V una matriz de proyeccin ortonormal, tal que $Y = V^T X$ y la covarianza de Y , $\text{cov}(Y)$, es la diagonal e inversa de V , es decir, $V^{-1} = V^T$, usando algo de algebra se obtiene que:

$$V \text{cov}(Y) = V V^T \text{cov}(X) V = \text{cov}(X) V, \quad (17)$$

podemos escribir V como $V = [V_1, V_2, \dots, V_d]$ y que

$$\text{cov}(Y) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{d-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_d \end{bmatrix}, \quad (18)$$

sustituyendo en la Ec. (17) tenemos

$$[\lambda_1 V_1, \lambda_2 V_2, \dots, \lambda_d V_d] = [\text{cov}(X) V_1, \text{cov}(X) V_2, \dots, \text{cov}(X) V_d], \quad (19)$$

lo que puede escribirse como:

$$\lambda_i V_i = \text{cov}(X) V_i, \quad (20)$$

donde $i = 1, 2, \dots, d$ y V_i es un eigen-vector de $\text{cov}(X)$.

Las imágenes de entrada que corresponden a las imágenes a fusionar $g_1(i, j)$ y $g_2(i, j)$ se disponen en dos vectores columna $V(1)$ y $V(2)$. El vector resultante tiene una dimensión de $n \times 2$, donde n es la longitud de cada vector de imagen. Se calculan los componentes normalizados P_1 y P_2 usando la Ec. (19) a partir del eigen-vector obtenido:

$$P_1 = \frac{V(1)}{\sum V} \text{ y } P_2 = \frac{V(2)}{\sum V}. \quad (21)$$

La imagen fusionada F_{PCA} está dada por:

$$F_{PCA}(i, j) = P_1 g_1(i, j) + P_2 g_2(i, j), \quad (22)$$

donde P_1 y P_2 son los componentes normalizados, $g_1(i, j)$ y $g_2(i, j)$ las imágenes de entrada.

1.5.4 Transformada de coseno discreta

El método DCT utiliza una secuencia de puntos de datos finitos en términos de una suma de funciones coseno los cuales oscilan a diferentes frecuencias [Desale, 2013]. Los grandes coeficientes de DCT se concentran en la región de baja frecuencia; por lo tanto, se sabe que tiene propiedades de densidad de energía excelentes [Ahmed, 1974], [Naidu, 2012], [Vassil, 2004].

La transformada de coseno discreta 2D de una imagen $g(i, j)$ de tamaño $M \times N$ se define como:

$$G(k_1, k_2) = \alpha(k_1) \alpha(k_2), \quad (23)$$

esto es:

$$G(k_1, k_2) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i, j) \left[\cos\left(\frac{\pi(2i+1)k_1}{2M}\right) \cdot \cos\left(\frac{\pi(2j+1)k_2}{2N}\right) \right], \quad (24)$$

para $0 \leq i \leq M-1$ y $0 \leq j \leq N-1$

y $\alpha(k_1), \alpha(k_2)$ toman el valor correspondiente dadas las siguientes condiciones:

$$\alpha(k_1) = \begin{cases} \frac{1}{\sqrt{M}}, k_1 = 0 \\ \sqrt{\frac{2}{M}}, 1 \leq k_1 \leq M-1 \end{cases}, \quad (25)$$

$$\alpha(k_2) = \begin{cases} \frac{1}{\sqrt{N}}, k_2 = 0 \\ \sqrt{\frac{2}{N}}, 1 \leq k_2 \leq N-1 \end{cases}, \quad (26)$$

donde k_1 y k_2 son variables de frecuencia discretas.

Para llevar a cabo la fusión, la imagen se divide en bloques no superpuestos de tamaño $N \times N$. Los coeficientes DCT se calculan para cada bloque y se obtienen los coeficientes DCT fusionados. Posteriormente, se aplica transformada inversa del coseno discreto sobre los coeficientes fusionados para producir la imagen de cada bloque. El procedimiento se repite para cada bloque, reintegrándose para formar la nueva imagen fusionada [Naidu, 2012].

1.5.5 Razón de la pirámide de paso-bajo

La construcción de la razón de la pirámide de paso-bajo, ROLP, consiste primeramente en generar una pirámide Gaussiana G para la imagen de la fuente de entrada, se toma la relación de dos capas sucesivas. La pirámide ROLP se define por:

$$R_i = G_i / EXPAND[G_{i+1}] \quad \text{para } 0 \leq i \leq N-1, \quad (27)$$

$$y \quad R_N = G_N. \quad (28)$$

donde cada nivel R_i es la relación de dos niveles sucesivos en la pirámide Gaussiana G , y la pirámide en su nivel i es el nivel de la pirámide, N es el índice del nivel superior de la pirámide y $EXPAND$ es la operación de interpolación que se realiza como el inverso de la operación de la pirámide Laplaciana descrita por Burt and Adelson (1985), [Toet, 1989].

La fusión de imágenes ROLP, se basa en la identificación de patrones relevantes definidos por sus valores de contraste de luminancia locales. Las imágenes de entrada se

descomponen de conjuntos de manchas claras y oscuras en diferentes niveles de resolución. A partir de entonces, los valores de contraste absoluto se comparan. La fusión de imagen real se realiza seleccionando los patrones con el contraste máximo absoluto de luminancia. La imagen fusionada se reconstruye a partir del conjunto de patrones primitivos obtenidos, como resultado, se conservan detalles relevantes con un contraste de luminancia local relativamente alto en la imagen compuesta, [Toet, 1989].

1.5.6 Transformada discreta de *wavelets*

La DWT descompone la imagen en diferentes bandas de frecuencia baja-alta, alta-baja, alta-alta a diferentes escalas y la frecuencia baja-baja se divide de nuevo en esta misma escala, **Fig. 9**. La banda baja-baja contiene la información media de la imagen mientras las otras bandas contienen información relativa a la orientación espacial. Los valores absolutos más altos de los coeficientes wavelet en las bandas altas corresponden a características salientes tales como bordes o líneas [Pajares, 2004], [Deepali, 2011], [Deepak, 2012].

Las DWT son funciones oscilatorias de duración y energía finita con valor cero medio [Deepali, 2011].

Algunos de los métodos DWT que se utilizan comunmente para la fusión son las familias de *wavelets* más conocidas cómo son: *Haar*, *Daubechies*, *Coiflets*, *Symlets*, *Bi-ortogonal*, *Meyer*, *Sombrero Mexicano*, *Shannon* y *Morlet* [Mathworks, 2015].

La señal se descompone en versiones escaladas y desplazadas de la función madre elegida. Las *Wavelets* de base son funciones generadas por la transacción y la dilatación de la *Wavelet* madre [Naidu, 2008], se puede describir por:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right), (a, b \in R), a > 0, \quad (29)$$

donde el factor de dilatación es a y el factor de traslación es b [Naidu, 2008].

El esquema de fusión de imágenes, **Fig. 9**, describe el algoritmo de fusión basado en DWT. Las imágenes de entrada $g_1(i, j)$ y $g_2(i, j)$ se descomponen usando DWT, la aproximación y los coeficientes de ambas imágenes se combinan usando la regla de ϕ .

$$\left[\phi \left\{ DWT(g_1(i, j)), DWT(g_2(i, j)) \right\} \right], \quad (30)$$

donde la regla de fusión ϕ puede ser, por ejemplo, la media de los coeficientes de aproximación y elegir aquellos coeficientes con la mayor magnitud de cada sub-banda.

La imagen fusionada $F_g(i, j)$ puede obtenerse tomando la transformada inversa de la *Wavelet* discreta (IDWT, *Inverse Discrete Wavelet Transform*) como $F_g(i, j) = IDWT$.

En esta tesis, los métodos de fusión basada en la familia de *Wavelets*: Coiflets, Daubechies, Haar, Bi-ortogonal y Symlets, se aplicaron en las pruebas ya que se consideran un buen punto de referencia para la evaluación comparativa con el método de fusión, AFS, propuesto.

En la **Tabla 2** se resumen las principales características de los algoritmos de fusión arriba descritos.

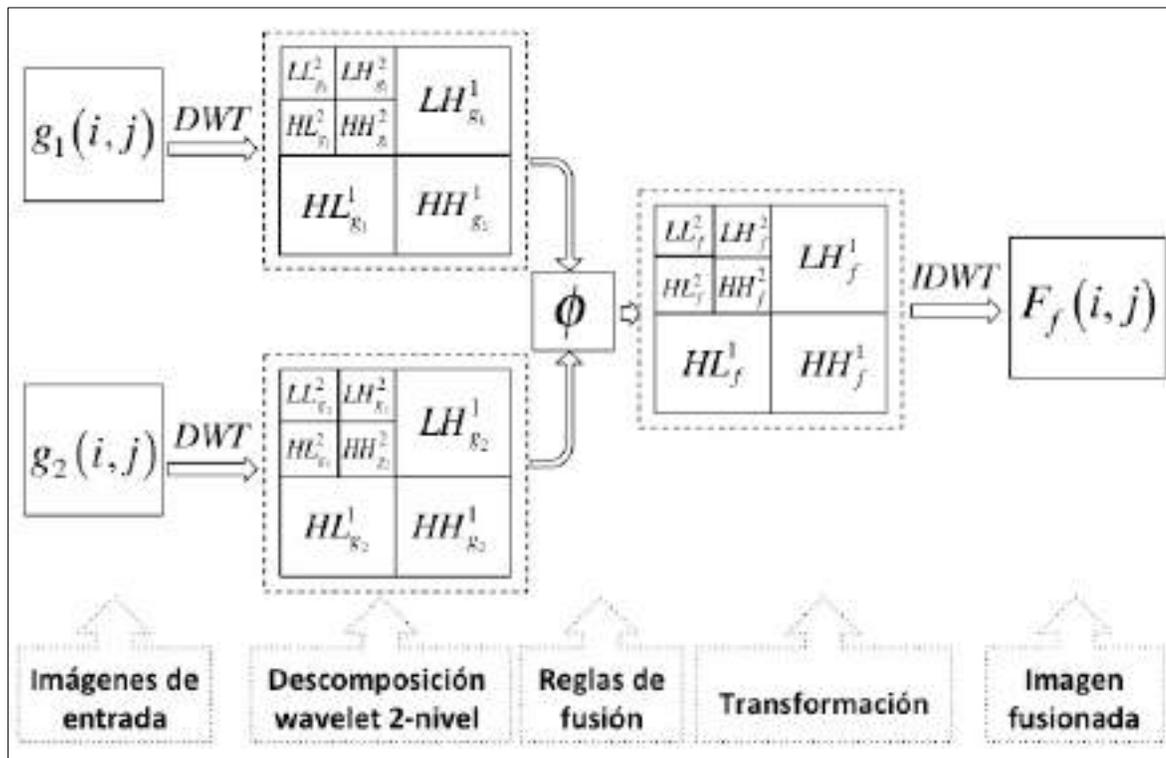


Fig. 9 Fusión de imágenes basada en *Wavelets*.

Tabla 2. Características principales de los algoritmos de fusión.

	Promedio	Valor máximo	Covarianza	Transformada Coseno	Pirámide Gaussiana	Espacial	Frecuencia
Promedio simple (SA)	✓					✓	
Máximo simple (SM)		✓				✓	
Análisis de componente principal (PCA)			✓			✓	
Transformada de coseno discreta (DCT)				✓			✓
Razón de la pirámide de paso-bajo (ROLP)					✓		✓
Transformada discreta de wavelets (DWT)							✓

1.6 Métricas para la evaluación de calidad de la imagen

Las métricas de calidad de la imagen, se utilizan para evaluar la calidad de la imagen resultante de algún PDI. En esta tesis se utilizan para evaluar los resultados de los métodos de autoenfoco y fusión, y así poder hacer una comparación entre los resultados obtenidos por diferentes métodos y el método que aquí se propone.

Las métricas permiten calcular una medida de calidad, éstas consideran diversos criterios que analizan los elementos de la imagen, como su estructura, iluminación, contraste, correlación, bordes, entre otros. Las métricas revisadas en esta tesis son la propuesta por Wang y Bovik [Wang y Bovik, 2002], su versión extendida por Piella y Heijmans [Piella y Heijmans, 2003] y ésta última readaptada por la métrica Multi-imagen [Bueno-Ibarra, 2005b], las cuales se definen a continuación.

1.6.1 Métrica de Wang y Bovik

La métrica de Wang y Bovik [Wang y Bovik, 2002], devuelve un valor de medida de calidad Q_0 que califica la calidad de la imagen fusionada. Tiene la capacidad de medir las distorsiones estructurales en las imágenes.

La imagen con mejor enfoque será aquella que posea el máximo valor de Q_0 .

Para obtener la media de una imagen g se calcula primero:

$$\bar{g} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N g(i, j), \quad (31)$$

después la varianza V^g de la imagen g que esta dado por:

$$V^g = \frac{1}{(MN)-1} \sum_{i=1}^M \sum_{j=1}^N [g(i, j) - \bar{g}]^2, \quad (32)$$

las Ecs. (31) y (32) se aplican tambien a la segunda imagen h , con la que se va a comparar, para obtener la covarianza entre ambas imágenes:

$$COV^{gh} = \frac{1}{(MN)-1} \sum_{i=1}^M \sum_{j=1}^N [(g(i, j) - \bar{g})(h(i, j) - \bar{h})], \quad (33)$$

donde $g(i, j)$ y $h(i, j)$ son dos imágenes, \bar{g} y \bar{h} son las medias de las imágenes $g(i, j)$ y $h(i, j)$ respectivamente, V^g y V^h son las varianzas correspondientes, COV^{gh} es la covarianza entre las imagenes $g(i, j)$ y $h(i, j)$. La métrica de calidad Q_0 está definido por:

$$Q_0 = \left[\frac{COV^{gh}}{\sqrt{V^g} \sqrt{V^h}} \right] \left[\frac{2\bar{g}\bar{h}}{\bar{g}^2 + \bar{h}^2} \right] \left[\frac{2\sqrt{V^g} \sqrt{V^h}}{V^g V^h} \right], \quad (34)$$

la cual es una medida de similitud entre $g(i, j)$ y $h(i, j)$ y toma los valores entre 1 y -1. El maximo valor de Q_0 se obtiene cuando las imágenes $g(i, j)$ y $h(i, j)$ son idénticas.

Tres factores multiplicativos se reconocen en Ec. (34). El primer factor es el coeficiente de correlación entre las imágenes $g(i, j)$ y $h(i, j)$ el cual toma valores en el intervalo $[-1, 1]$; El segundo factor es un tipo de distorsión media de la iluminación y toma valores

en el rango $[0,1]$; Finalmente, el tercer factor es una medida de la distorsión debida al contraste y tiene un rango de $[0,1]$.

1.6.2 Métrica de Piella y Heijmans

La métrica de Piella y Heijmans [Piella y Heijmans, 2003] es una extensión de la métrica de Wang y Bovik [Wang y Bovik, 2002] descrita en la sección anterior. Utiliza la métrica Q_0 de Wang y Bovik para definir tres nuevas métricas Q , Q_{wt} y Q_E , en términos de los cuales la calidad de la imagen obtenida de la fusión f^* , se compara con las imágenes $g(i, j)$ y $h(i, j)$ que conforman dicha fusión.

Las métricas para medir la calidad de la fusión Q , Q_{wt} , y Q_E están definidas por:

$$Q(g, h, f^*) = \frac{1}{|T|} \sum_{w_s \in T} \left[\lambda_g(w_s) Q_0(g, f^* | w_s) + \lambda_h(w_s) Q_0(h, f^* | w_s) \right], \quad (35)$$

$$Q_{wt}(g, h, f^*) = \frac{1}{|T|} c(w_s) \sum_{w_s \in T} \left[\lambda_g(w_s) Q_0(g, f^* | w_s) + \lambda_h(w_s) Q_0(h, f^* | w_s) \right], \quad (36)$$

$$Q_E(g, h, f^*) = Q_{wt}(g, h, f^*)^{1-\alpha} \cdot Q_{wt}(g^e, h^e, f^{e*})^\alpha, \quad (37)$$

donde g y h son dos imágenes y f^* es la imagen fusionada, $|T|$ es la cardinalidad de la familia de todas las ventanas desplazables T , λ_g es el peso local de $g(i, j)$ con respecto a $h(i, j)$, y λ_h es el peso local de $h(i, j)$ con respecto a $g(i, j)$; los parametros g^e, h^e, f^{e*} son las imágenes de los bordes correspondientes a $g(i, j), h(i, j), f^*$. El parámetro α toma valores entre 0 y 1, entre mas cercano a 1 sea el valor de α mayor es la importancia dada a la contribución de los bordes.

Las cantidades λ_g y λ_h están dadas por:

$$\lambda_g(w_s) = \frac{\gamma(g|w_s)}{\gamma(g|w_s) + \gamma(h|w_s)}, \quad (38)$$

$$\lambda_h(w_s) = \frac{\gamma(h|w_s)}{\gamma(g|w_s) + \gamma(h|w_s)}. \quad (39)$$

El parametro de contribución de borde λ^{ov} está dado por:

$$\lambda^{ov}(g, h|w_s) = MAX\left(\{\gamma(g|w_s), \gamma(h|w_s)\}\right), \quad (40)$$

donde λ^{ov} es el valor general prominente resultante de una comparación entre dos ventanas w_s de las imágenes $g(i, j)$ y $h(i, j)$.

1.6.3 Métrica Multi-imagen

La métrica multi-imagen de Mario Bueno-Ibarra [Bueno-Ibarra, 2005b] se basa en la métrica de Piella y Heijmans [Piella y Heijmans, 2003], a diferencia de la métrica de Piella, la métrica multi-imagen ofrece la capacidad de comparar más de dos imágenes para evaluar la calidad del resultado de la fusión de n -imágenes. Utiliza los índices Q , Q_{wt} y Q_E , para evaluar la calidad de la imagen f^* resultante de la fusión de dos o más imágenes. Esto significa que la imagen fusionada f^* se compara con cada una de las imágenes que se integraron en la fusión. Los índices son Q , Q_{wt} y Q_E , y se definen a continuación:

$$Q(f_{1,\dots,K}, f^*) = \frac{1}{|T|} \sum_{w_s \in T} \sum_{p=1}^K \lambda_{f_p}(w_s) Q_0(f_p, f^*|w_s), \quad (41)$$

$$Q_{wt}(f_{1,\dots,K}, f^*) = \sum_{w_s \in T} \sum_{p=1}^K c^*(w_s) \lambda_{f_p}(w_s) Q_0(f_p, f^*|w_s), \quad (42)$$

$$Q_E(f_{1,\dots,K}, f^*) = Q_{wt}(f_{1,\dots,K}, f^*)^{1-\alpha} \cdot Q_{wt}(f_{1,\dots,K}^e, f^{e*})^\alpha, \quad (43)$$

donde $f_{1,\dots,k}$ es el conjunto de K imágenes fusionadas en la imagen f^* y el peso local λ_{f_p} de las imágenes f_p . El valor de λ_{f_p} se obtiene por:

$$\lambda_{f_p}(w_s) = \frac{\gamma(f_p|w_s)}{\sum_{i=1}^K \gamma(f_i|w_s)}, \quad (44)$$

donde $\gamma(f_p|w_s)$ representa al conjunto $\gamma(f_1|w_s), \gamma(f_2|w_s), \dots, \gamma(f_K|w_s)$ de valores prominentes dentro de la ventana w_s , la contribución de borde $c^*(w_s)$ está dado por:

$$c^*(w_s) = \frac{\gamma^{ov}(f_{1,\dots,K}|w_s)}{\sum_{w_s \in T} \gamma^{ov}(f_{1,\dots,K}|w_s)}, \quad (45)$$

donde γ^{ov} es el valor prominente general, resultante de la comparación de dos o más ventanas desplazables w_s , y está dado por:

$$\lambda^{ov}(g, h|w_s) = MAX(\{\gamma(f_1|w_s), \gamma(f_2|w_s), \dots, \gamma(f_k|w_s)\}). \quad (46)$$

Capítulo 2

Metodología

En la actualidad, la mayoría de los sistemas de adquisición de imágenes tienen una función de enfoque automático. Uno de los objetivos de ésta tesis es crear un algoritmo que obtenga una imagen enfocada de alta calidad con un mínimo de procesos manuales con las características más similares posibles a la escena real y que también reúna información suficiente para su análisis.

La metodología propuesta en esta tesis, para el desarrollo de los nuevos algoritmos de autoenfoco y de fusión, se describe en el siguiente diagrama, **Fig. 10**.

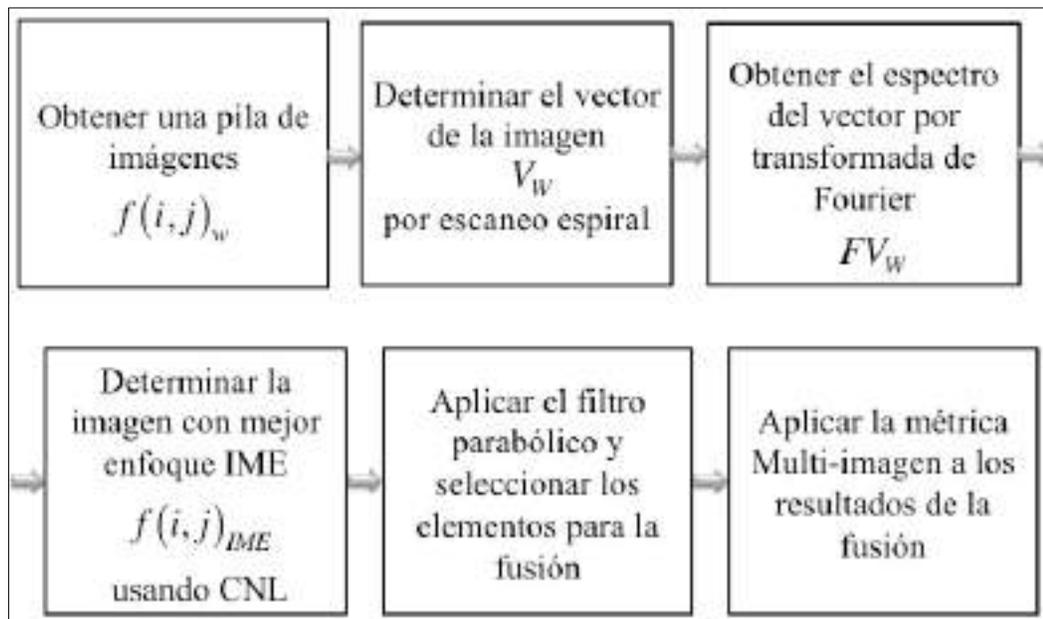


Fig. 10 Metodología AFS.

Básicamente, primero se obtiene un conjunto de imágenes digitalizadas mediante algún dispositivo de captura; después se representa cada imagen en un vector que contiene elementos de la imagen seleccionados por barrido en espiral, en seguida, se obtiene el espectro de vector que representa a cada imagen aplicando la transformada de Fourier, y después se aplica el criterio para identificar la mayor medida de enfoque que determina la IME.

Una vez que se tiene la IME se procede a identificar el subconjunto de imágenes que formarán la fusión, a este selecto subconjunto se les aplica el filtro parabólico que realiza los elementos relevantes de la imagen, éstos son seleccionados para integrarse en la creación de una nueva imagen, la imagen fusionada. Finalmente, se aplica una métrica para evaluar la calidad de la imagen resultante del proceso de la fusión.

2.1 Adquisición del conjunto de imágenes

El primer paso de la metodología de enfoque automático es obtener un conjunto de imágenes. Para esta tesis se utilizaron diferentes conjuntos, algunas obtenidas por microscopio, otras por cámara digital y otras mediante un telescopio.

Las imágenes obtenidas por microscopio fueron capturados mediante el CCD, modelo *Leica DC-300* integrada a un microscopio *Leica DM RXA2* con platina motorizada para desplazamientos, X/Y y Z , **Fig. 11**.



Fig. 11 Microscopio *Leica DM RXA2* con platina motorizada para X/Y y Z desplazamientos.

El método de captura de las imágenes mediante el microscopio consiste en establecer una posición fija de la platina en X/Y y un desplazamiento vertical en Z , **Fig. 12**, para así capturar una secuencia de imágenes de una misma muestra a diferentes distancias del objetivo. El conjunto de imágenes se compone de W imágenes f_1, f_2, \dots, f_w capturadas en incrementos de Δz en la dirección del eje Z del microscopio, por medio del CCD, cada imagen de tamaño $M \times N$, donde $f(i, j)_w$ es una imagen de tamaño $M \times N$ con píxeles (i, j) , para $i = 1, 2, \dots, M$; $j = 1, 2, \dots, N$; $w = 1, 2, \dots, W$.

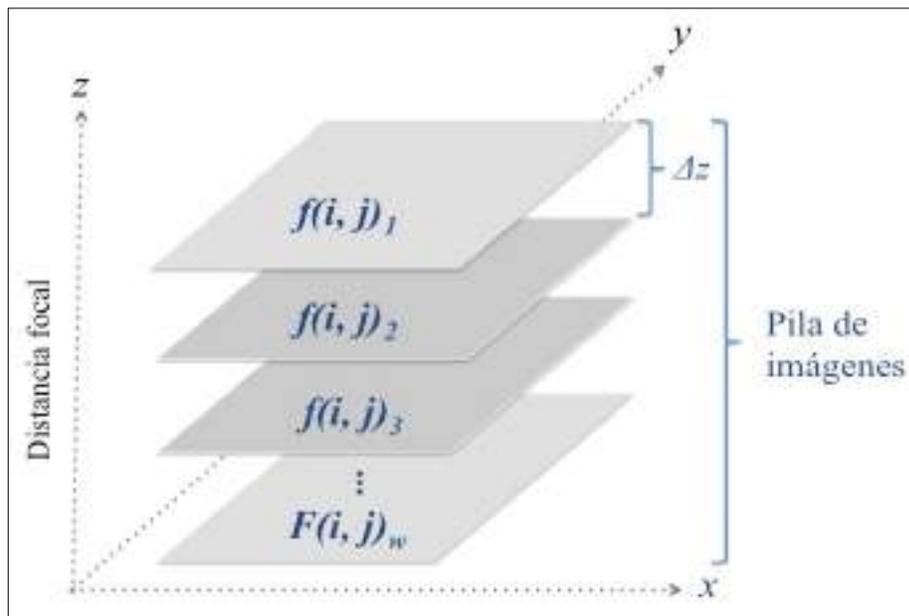


Fig. 12 Conjunto de imágenes, donde $f(i, j)_w$ es una imagen digital capturada a una distancia focal Δz .

Se utilizaron tres conjuntos de imágenes de muestras microscópicas de algunos microorganismos marinos:

- 60 imágenes a color (RGB) de una muestra de *Tricodina*, **Fig. 13(a)**,
- 100 imágenes en escala de grises, de una muestra de *Diatomea*, **Fig. 13(b)**,

estas imágenes tienen una resolución de 2088×1550 píxeles, en un formato de mapa de bits (BMP).

Así también, se utilizó un conjunto de imágenes obtenidas por telescopio, la cual consiste de:

- 10 imágenes del disco y el limbo solar, **Fig. 13 (c)**, las cuales fueron proporcionadas por el Observatorio Carl Sagan de UNISON [Saucedo-Morales, 2014]. Estas fueron capturadas con el Telescopio Solar Coronado 70mm f/7.5 (λ 656.3 nm, $\Delta\lambda$ 0.5 Å) con filtro H-alfa (6563 Å) usando una exposición de 1.8 milisegundos, y por una cámara Lumenera SKYNYx 2-2. La resolución de estas imágenes es de 1616×1216 píxeles en escala de grises, en un formato de mapa de bits (BMP).

Una prueba más se realizó con un conjunto de imágenes obtenidas por cámara digital *Canon PowerShot S5 IS* y sometidas a filtros de difuminado para acentuar el desenfoque:

- 12 imágenes de un auto de juguete suspendido, **Fig. 13(d)**, la resolución de estas imágenes es de 2048×1536 y se realizaron pruebas con la imagen a color y en escala de grises, con formato bmp.

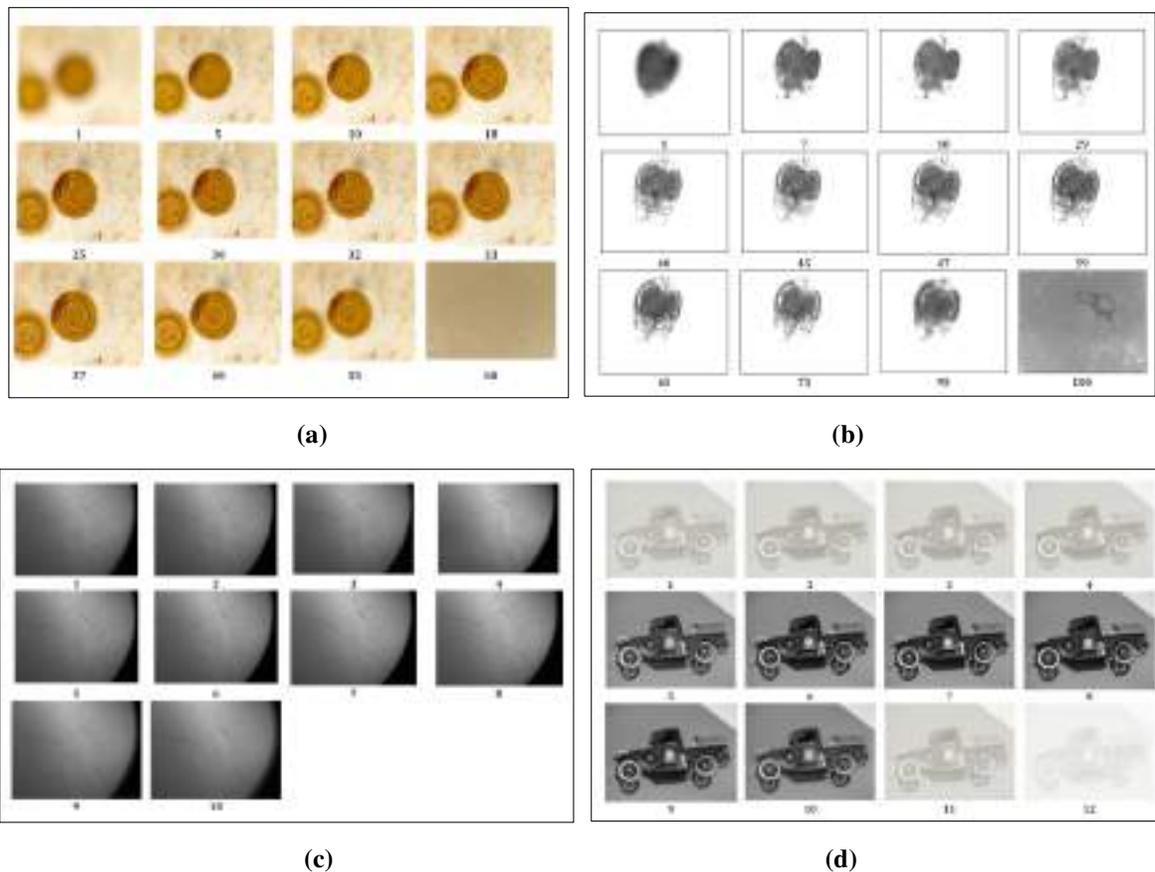


Fig. 13 Conjuntos de imágenes utilizadas para las pruebas de autoenfoco y fusión.

2.2 Determinar el vector de la imagen por barrido en espiral

En la actualidad las imágenes adquiridas mediante los nuevos sistemas de adquisición de imágenes poseen regularmente una alta resolución, lo cual es un atributo deseable ya que contribuye a la nitidez de la imagen. Sin embargo, el procesamiento de imágenes de alta resolución puede verse afectado en su eficiencia, sobre todo si se maneja una gran cantidad de imágenes con esta característica.

Con la finalidad de no sacrificar la eficiencia, debido a la alta complejidad en tiempo y espacio computacional que implica procesar grandes cantidades de imágenes de alta resolución, se propone implementar un nuevo método de lectura de la información de la imagen de entrada. El cual toma una muestra de información suficiente de la imagen ya sea para su análisis o evaluar algún criterio en la toma de decisiones de algún proceso.

En esta tesis se propone crear un vector representativo de cada imagen de entrada, el cual contenga una muestra representativa y suficiente de los elementos que componen la imagen. Esto se logra mediante un método de exploración o barrido en espiral de lectura para tomar una muestra de la imagen. Así, se optimiza el tiempo de ejecución y los requisitos de almacenamiento computacional.

El barrido en espiral consiste en tomar el dato de cada píxel (i, j) de la imagen $f(i, j)_w$ que coincida con las coordenadas generadas por la ecuación de la espiral y registrarlo en un vector V_w . Estas coordenadas (i, j) son generadas por las ecuaciones de los vectores la espiral (47) y (48), es decir, cada punto en los vectores X y Y dado por dichas ecuaciones representa un par de coordenadas de los píxeles (i, j) en la imagen, se toma el valor del píxel en dicha posición (i, j) y este valor se registra en el vector V_w para su posterior procesamiento. Esta exploración permite seleccionar un subconjunto de píxeles representativos que componen el vector V_w de la imagen $f(i, j)_w$.

El algoritmo de barrido espiral utiliza ecuaciones siguientes:

$$X = \frac{t}{n} \cdot \cos(t) + L \quad , \quad (47)$$

$$Y = \frac{t}{n} \cdot \text{sen}(t) + L \quad , \quad (48)$$

donde X y Y son vectores que determinan la posición (i, j) , respectivamente, en la imagen $f(i, j)_w$. El parámetro t determina la rotación de la espiral, n es el factor de escala que determina el tamaño de la espiral y L se relaciona con el tamaño de la imagen y tiene un valor de $\frac{l}{2}$, donde l es la longitud de la imagen.

El gráfico muestra el patrón de barrido, **Fig. 14**. En el eje x tenemos los valores de X dada por la Ec. (47) y en el eje y los valores de Y dada por la Ec. (48) éstos forman el vector V_w para cada una de las imágenes.

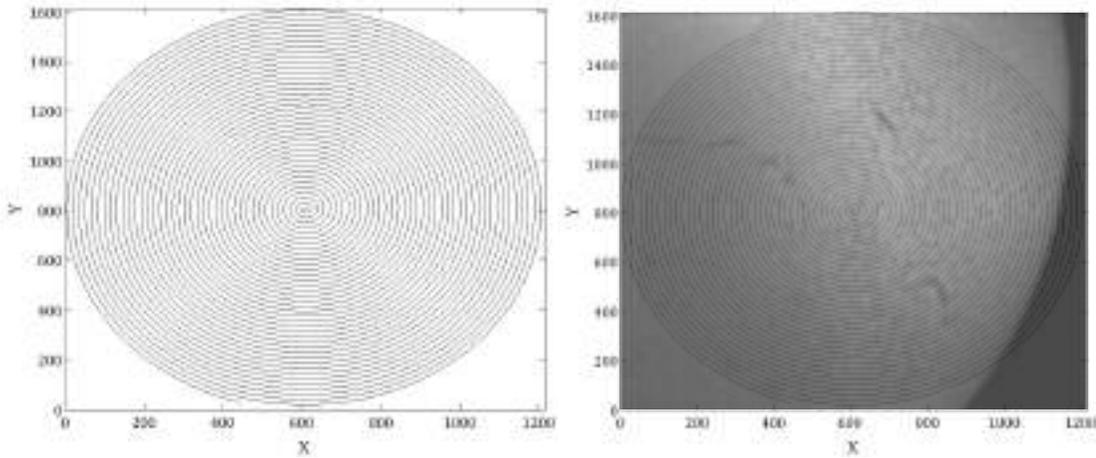


Fig. 14 Barrido en espiral de una imagen de 1616×1216 .

Los vectores V_1, V_2, \dots, V_w de las imágenes $f(i, j)_1, f(i, j)_2, \dots, f(i, j)_w$ del conjunto, están definidos por:

$$V_1(i) = f(X_i, Y_i)_1, V_2(i) = f(X_i, Y_i)_2, \dots, V_w(i) = f(X_i, Y_i)_w. \quad (49)$$

donde X_i y Y_i determinan qué pixel $(i, j)_w$ de la imagen $f(i, j)_w$ serán incluidos en el vector V_w . Se aplica una función de redondeo a los resultados de las Ecs. (47) y (48) correspondientes a X y Y , ya que los valores deben ser números enteros positivos pues se refieren a la posición de un pixel en la imagen.

2.3 Espectro del vector de la imagen mediante transformada de Fourier

Una vez calculados los vectores V_w para cada imagen del conjunto, el vector V_1 se toma como la imagen de referencia, que corresponde a la imagen menos enfocada.

El espectro FTV_w es obtenido para cada uno de los vectores V_w . A cada vector V_w se le aplica la transformada de Fourier con filtro no lineal, así el espectro es obtenido por:

$$FTV_w = |V(f)|_w^k e^{i\varphi} . \quad (50)$$

donde $|V(f)|_w$ es el módulo de la FFT de $V(f)_w$; $k = 0.3$ es el factor de no linealidad, y φ es la fase, FV_1, FV_2, \dots, FV_w contiene el espectro de vectores FV_w correspondientes a cada una de las imágenes $f(x, y)_w$.

2.4 Determinar la imagen mejor enfocada con CNL

Una vez que se tienen los FTV_w de las imágenes del conjunto, se hace la correlación no lineal entre el espectro de la imagen de referencia, FTV_1 , con el resto de los espectros FTV_2, \dots, FTV_w , se calcula la de-correlación d y se busca el valor máximo de d , éste valor máximo corresponde a la IME, **Fig. 15**.

La medida de enfoque se requiere para determinar la imagen mejor enfocada. En esta tesis, la medida de enfoque se obtiene mediante el cálculo de la CNL entre el espectro del vector FTV_1 de la Imagen de Referencia, IRef, que es la menos enfocada, con cada uno de los espectros de los vectores FTV_w de cada una de las imágenes en el conjunto. La CNL, se obtiene por:

$$C_{NL} = \max \left[FT^{-1} \left(|V(f)|_w^k e^{i\varphi} |V(f)|_1^k e^{-i\theta} \right) \right] . \quad (51)$$

donde $|V(f)|$ es el valor del módulo de la transformada de Fourier de la función $V(f)$; k es el factor de no linealidad que toma valores $0 > k > 1$; $w = 1, 2, \dots, W$; $i = \sqrt{-1}$; donde φ es la fase de la transformada de Fourier para cada imagen y θ es la fase de la imagen de referencia. FT^{-1} es la inversa de la transformada de Fourier.

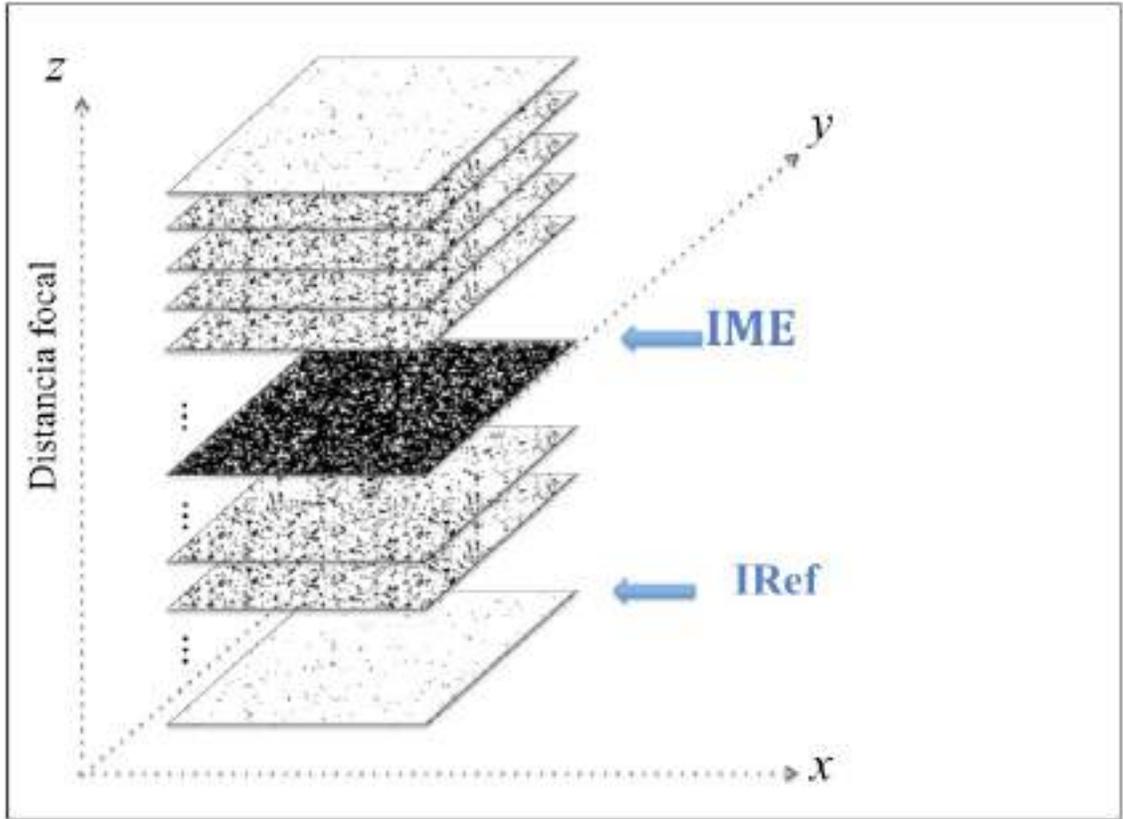


Fig. 15 Identificación de la IME y la Imagen de Referencia en el conjunto.

Si la CNL es cercana al valor 1 entonces hay una fuerte correlación, es decir, existen similitudes entre los vectores. Si el valor obtenido está cerca de 0, entonces no hay correlación entre los vectores, la intensidad de la asociación es baja.

La correlación será muy baja, es decir cercana a cero, cuando se realiza la comparación del espectro de la imagen de referencia FTV_1 , que es la imagen más desenfocada, **Fig. 16(a)**, con el espectro de la imagen mejor enfocada FTV_{IME} , **Fig. 16(b)**. Por lo tanto la imagen con mejor enfoque será la que posea el valor máximo de la diferencia entre 1 y $|C_{NL}|$, a lo que se le llama de-correlación. La de-correlación d , se define como:

$$d = 1 - |C_{NL}|, \quad (52)$$

y la IME estará determinada por el valor máximo de la de-correlación:

$$[MaxValor, i] = \max[d], \quad (53)$$

dónde $MaxValor$ es el máximo valor de d , y el índice i hace referencia a la imagen correspondiente a dicho valor máximo, la cual es la imagen mejor enfocada.

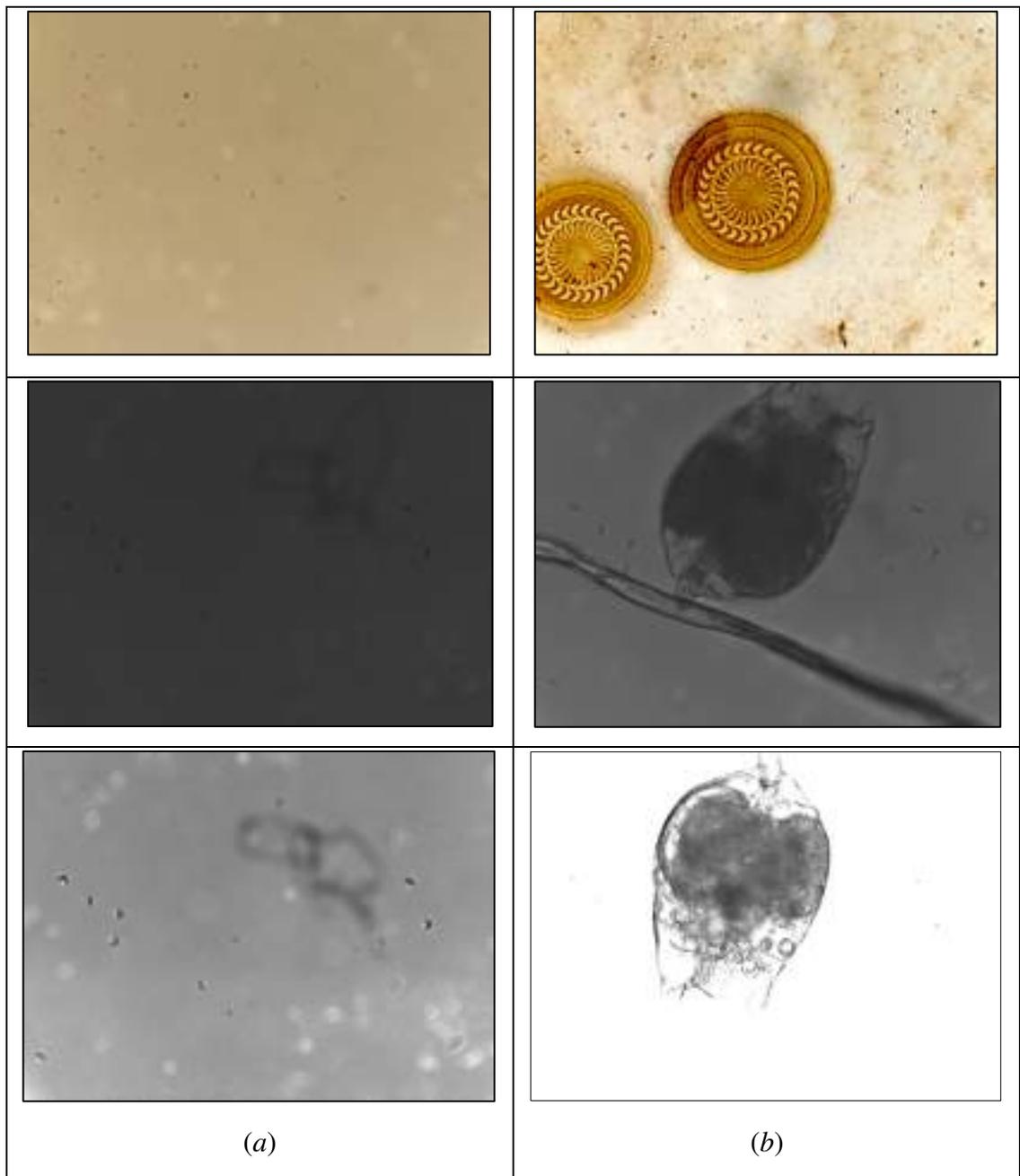


Fig. 16 Cada renglón representa tres distintos conjuntos de imágenes. En la columna (a) están las Imágenes de referencia $f(i, j)_1$ que son las más desenfocadas, mientras que en la columna (b) se muestran las imágenes mejor enfocadas $f(i, j)_{IME}$.

La metodología utilizada para determinar la IME con CNL se puede resumir en el diagrama de la **Fig. 17**.

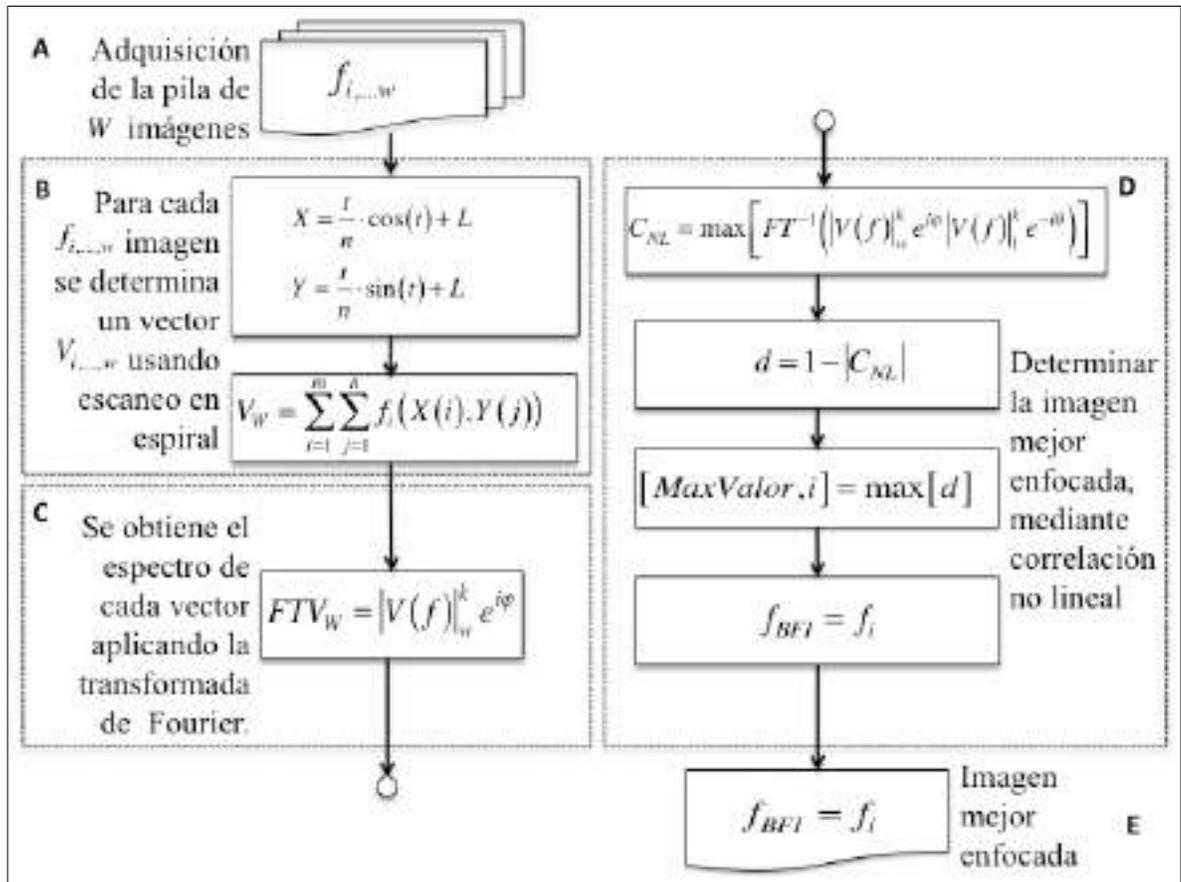


Fig. 17 Metodología para determinar la IME, con CNL.

2.5 Fusión de imágenes aplicando el filtro parabólico

La metodología para la fusión selecciona primero un subconjunto de imágenes, luego se obtienen los elementos relevantes de cada imagen de este subconjunto, se eligen los valores máximos resultantes de la multiplicación del espectro por el filtro parabólico para crear la fusión.

Una vez que ha identificado la imagen que tiene el mejor enfoque $f(i, j)_{IME}$, se procede a obtener una nueva imagen de mejor calidad mediante el proceso de fusión.

El subconjunto de imágenes que se fusionan son aquellas con el valor de medida de enfoque más cercano al del mejor enfoque. Sea $f(i, j)_{SME}$ el subconjunto de imágenes con mejor enfoque, que conforma la fusión, incluyendo a la IME, $f(i, j)_{IME}$, Fig.18.

Se aplica un filtro parabólico, pasa altas, que destaca los elementos con niveles más altos de intensidad de cada una de las imágenes del subconjunto que serán seleccionados para conformar la nueva imagen, **Fig. 19**.

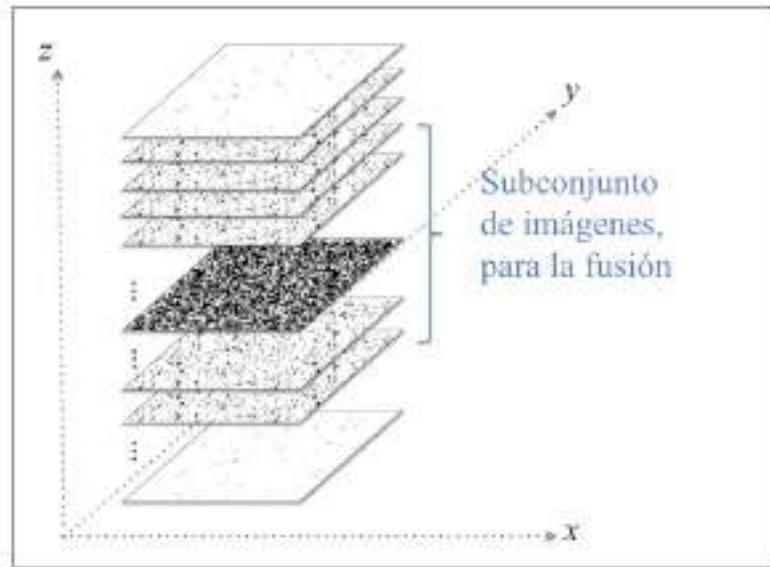


Fig. 18 Las imágenes que serán fusionadas $f(i, j)_{IME}$ son un subconjunto del total de imágenes del conjunto original, aquellas con índice cercano a la IME.

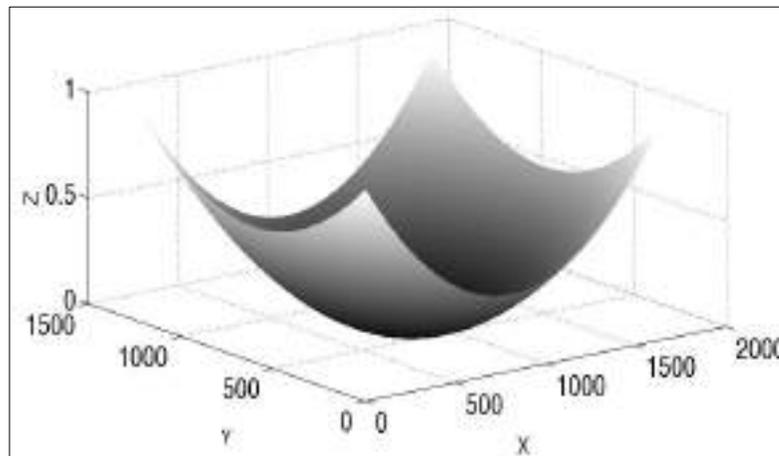


Fig. 19 Representación gráfica del filtro parabólico.

Las dimensiones del filtro parabólico están determinadas por el tamaño $M \times N$ de la imagen y dadas por:

$$c = \frac{(M-1)}{2}, r = \frac{(N-1)}{2}, \quad (54)$$

$$X = [-c : c], Y = [-r : r], \quad (55)$$

donde c y r son los valores de columna y fila respectivamente usados para calcular la matriz $X \times Y$ de filtro parabólico. El filtro parabólico fp está dado por:

$$S_1 = \frac{(2.2 \cdot \pi)}{c}, S_2 = \frac{(2.2 \cdot \pi)}{r}, \quad (56)$$

$$fp = \left| S_1 S_2 \left(0.01 (X^2 + Y^2) \right) \right|, \quad (57)$$

donde S_1 y S_2 son las constantes parabólicas y el valor de 0.01 normaliza los valores del filtro parabólico.

Posteriormente, se obtiene el espectro de la matriz de la imagen utilizando la transformada de Fourier bidimensional para cada imagen del subconjunto. La matriz del filtro parabólico se multiplica por el espectro de matriz de imagen, esto se aplica a cada imagen en el subconjunto, esto puede expresarse por

$$\rho(x, y)_i = \left[fp(x, y) \cdot FT2(f(x, y)_i) \right], \quad (58)$$

donde ρ es el producto de multiplicación, fp el filtro parabólico, $f(x, y)_i$ es el valor de píxel de la imagen f_i en las coordenadas (x, y) de la matriz, $i = 1, 2, \dots, \omega$ el número de la imagen y $FT2$ es la transformada de Fourier bidimensional.

Por último, se analiza cada resultado para seleccionar el valor máximo que se asignará a la nueva imagen fusionada y está dado por

$$f^*(x, y) = MAX \left[\rho(x, y)_i \right], \quad (59)$$

donde f^* es la nueva imagen fusionada que contiene los elementos de valor máximo para cada coordenada (x, y) de la imagen.

La fusión de imágenes comprende a la IME y a un subconjunto de imágenes del conjunto. Este subconjunto está conformado por una selección de imágenes las cuales

tienen una medida de enfoque pr3xima a la medida m3xima de enfoque de la IME. Para decidir qu3 im3genes se fusionar3n podemos calcular un rango de medidas de enfoque, dado por:

$$\text{rango} \left\{ \begin{array}{l} I_{SUP} = \max(d) \\ I_{INF} = \max(d) - \left[\frac{\max(d) - \text{median}(d)}{\text{mean}(d) \cdot W} \right] \end{array} \right. , \quad (60)$$

I_{SUP} determina el l3mite superior del rango; $\max(d)$ el valor m3ximo de la de-correlaci3n donde d es el valor de la de-correlaci3n; I_{INF} es el l3mite inferior; $\text{median}(d)$ es la funci3n que devuelve la mediana de los valores de de-correlaci3n, y $\text{mean}(d)$ es la funci3n que devuelve la media de los valores de de-correlaci3n, W es el n3mero de im3genes en el conjunto del cual se extrae el subconjunto.

Capítulo 3

Resultados y discusión

En el presente capítulo se muestran los resultados del autoenfoco y fusión del algoritmo AFS propuesto en esta tesis. Asimismo, se presentan las evaluaciones comparativas de los resultados del AFS con los algoritmos de autoenfoco y fusión descritos en el Capítulo 1. Para realizar esta evaluación se aplican las medidas de calidad Q , Q_{Wt} y Q_E de la métrica multi-imagen, Ecs. (41-43), utilizando un algoritmo computacional que implementa esta métrica.

Las pruebas se realizaron con diversos tipos de imágenes para comprobar la eficiencia y robustez del algoritmo. Se utilizaron diversos tamaños de imagen de $M \times N$ píxeles tales como 2088×1550 , 1616×1216 y 522×387 ; con diferentes modelos cromáticos, RGB (*Red, Green, Blue*) o de escala de grises. Además, se utilizaron diferentes formatos de imagen tales como BMP (*Bit Map Protocol*) y JPG (*Joint Photographic Experts Group*). Los conjuntos de imágenes fueron adquiridas por diferentes medios de captura, cámara digital *Canon PowerShot S5 IS*, microscopio *Leica DM RXA2* y telescopio *Coronado Solar Max*, como se detalla en el Capítulo 2.

3.1 Resultados de la obtención de la imagen mejor enfocada

El *método clásico o cualitativo* para verificar la mejor imagen enfocada es la comprobación visual, esta medida cualitativa que utiliza el observador, puede ser relativa porque la percepción de cada experimentador tiene diferencias en la agudeza visual y conceptuales para decidir la calidad de una imagen.

El *método formal o cuantitativo* consiste en aplicar las métricas o algoritmos computacionales diseñados para el análisis la imagen. Esta medida cuantitativa permite

hacer una evaluación mediante una función numérica, la cual considera determinados criterios que permiten calcular el porcentaje de calidad de una imagen con respecto a otras.

En la **Fig. 20** se comparan dos imágenes, una es la más fuera de foco del conjunto de imágenes y la otra es la mejor enfocada, en este caso el discernimiento visual es sencillo, sin embargo, cuando los niveles de enfoque son muy similares es difícil distinguir visualmente entre dos o más imágenes a la mejor enfocada.

Cuantitativamente esto es posible, realizando la comparación entre la medida de la de-correlación de cada espectro de la imagen contra los del resto de imágenes en el conjunto, hasta identificar el mayor valor de la d el cual corresponde a la mejor enfocada o IME. La gráfica, **Fig. 21**, compara las medidas de de-correlación, d , de la IME con las del resto de las imágenes en el conjunto.

Para estas pruebas se implementaron los algoritmos de autoenfoco: energía de la imagen Laplaciana [Subbarao, 1998], Tenengrad [Tenenbaum, 1970] [Schlag, et al, 1983], Tenengrad absoluto [Jarvis, 1676], varianza [Groen, 1985], varianza normalizada [Groen, 1985], Vollath's F4 [Vollath, 1988], Vollath's F5 [Vollath, 1988], filtro Gaussiano [Geusebroek, 2000], histograma logarítmico [Hilsenstein, 2005], histograma ponderado [Zeder, 2009], autoenfoco por correlación de Pearson y Fourier [Bueno-Ibarra, 2005a], descritos anteriormente (Sección 1.4, Capítulo 1).

Las pruebas realizadas consideran como criterios:

- la selección de imagen mejor enfocada y
- tiempo de ejecución.

En los resultados se obtuvo que el algoritmo AFS identifica eficientemente la IME, y en menor tiempo de ejecución que el resto de algoritmos. La IME seleccionada por el algoritmo resulta con la mejor evaluación de calidad por la métrica multi-imagen. La **Tabla 3** muestra los resultados en cuanto a la IME seleccionada por los algoritmos de autoenfoco analizados y el algoritmo AFS propuesto en esta tesis. En la gráfica de la **Fig. 22**, se muestran los tiempos promedio de ejecución de los mismos, donde puede apreciarse que AFS requiere menor tiempo de ejecución comparado con los otros algoritmos, aún cuando el número de imágenes analizadas varíe, **Fig. 23**.

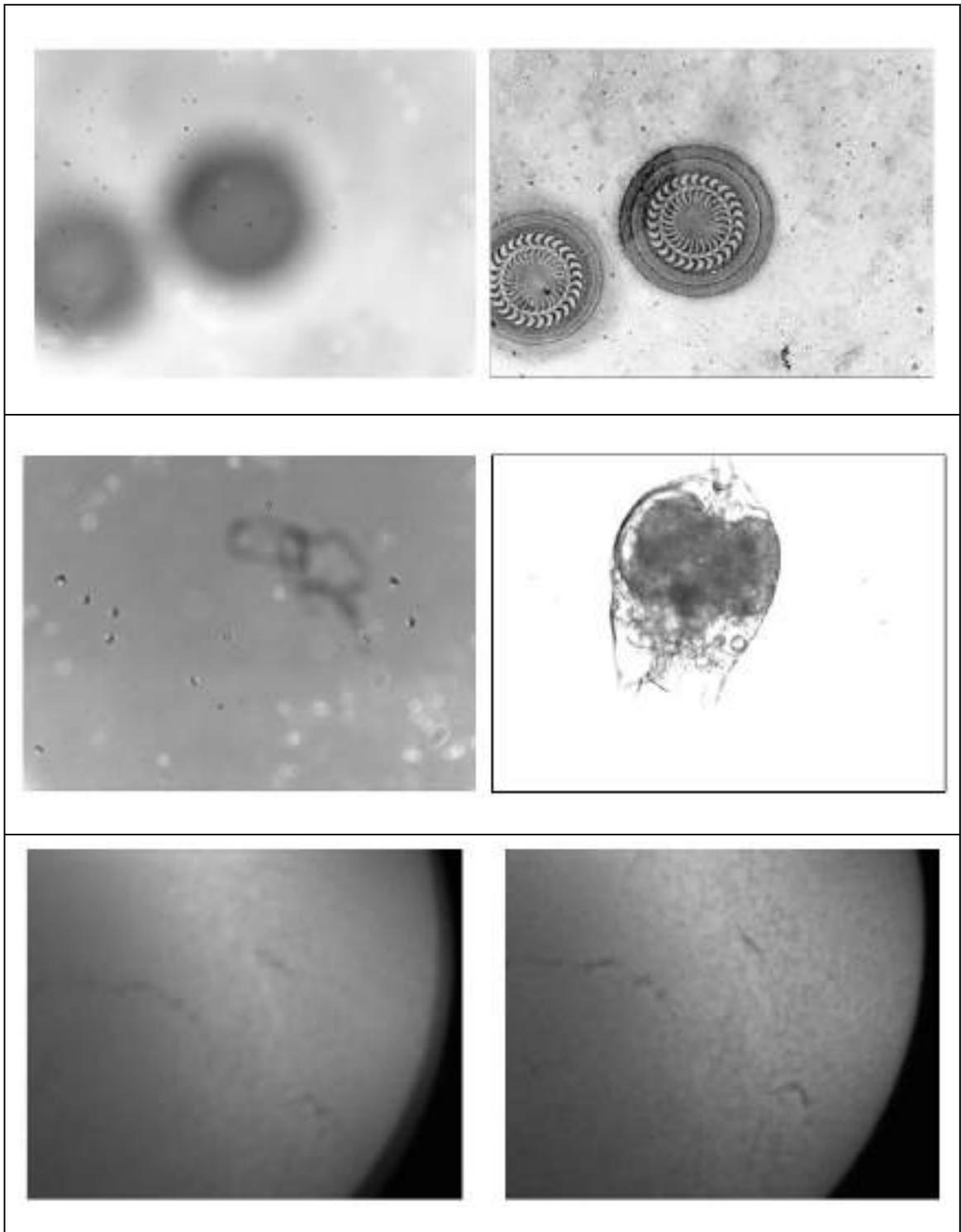


Fig. 20 Comparación de las imágenes de referencia (izquierda) y la IME (derecha) de tres diferentes conjuntos de imágenes.

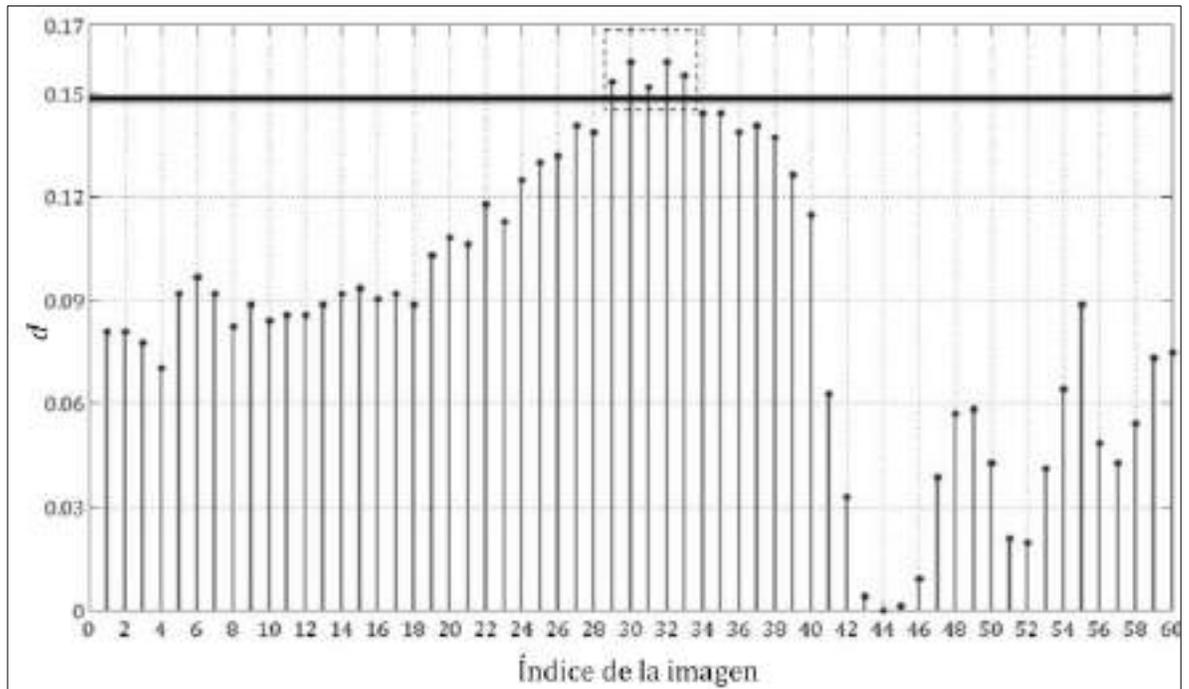


Fig. 21 Se muestra en el eje X los índices de las imágenes y su medida de de-correlación d en el eje Y . Correspondientes al conjunto de *Tricodina*. La IME será aquella que tiene el valor máximo de de-correlación.

Tabla 3. Índice de la IME seleccionada para cada conjunto por cada algoritmo de prueba.

Algoritmo	Tricodina RGB 2088×1550	Diatomea E. de Gris 2088×1550	Sol E. de Gris 1616×1216	Automóvil RGB 2048×1536
E. I. Laplaciana	34	48	5	6
Tenengrad	34	43	5	3
Tenengrad absoluto	33	44	4	4
Varianza	32	48	3	4
Varianza normalizada	32	48	3	5
Vollath's F4	33	43	4	5
Vollath's F5	32	44	3	5
Filtro Gaussiano	34	47	5	4
Histograma Logarítmico	25	39	6	6
Histograma ponderado	33	48	4	6
Autoenfoco por correlación de Pearson y Fourier	30	46	3	5
AFS	30	45	2	5

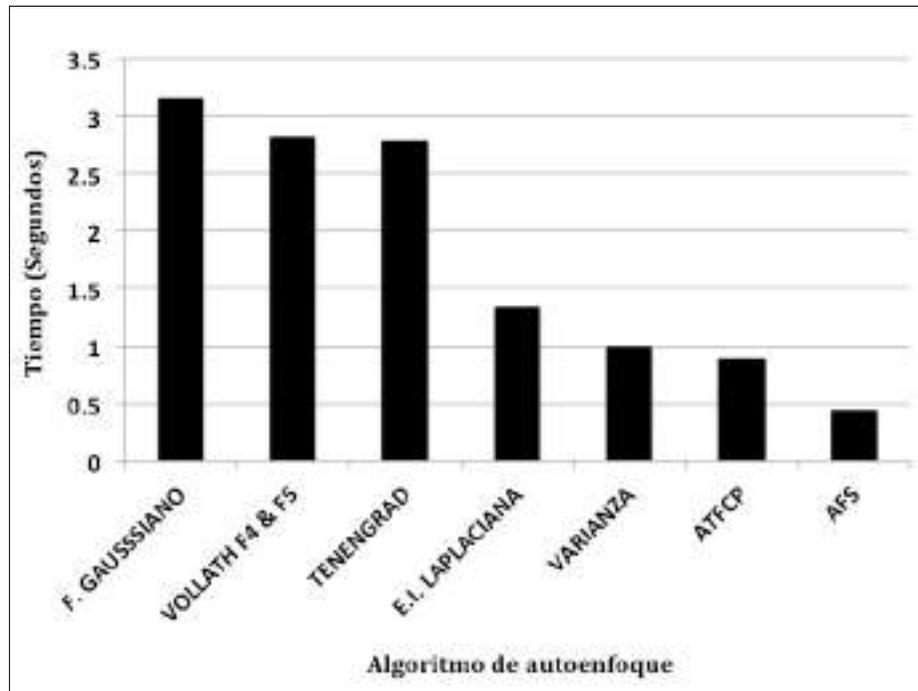


Fig. 22 El gráfico muestra en el eje x el nombre del algoritmo, en el eje y su correspondiente tiempo de ejecución en segundos, para el conjunto de 60 imágenes de *Tricodina*.

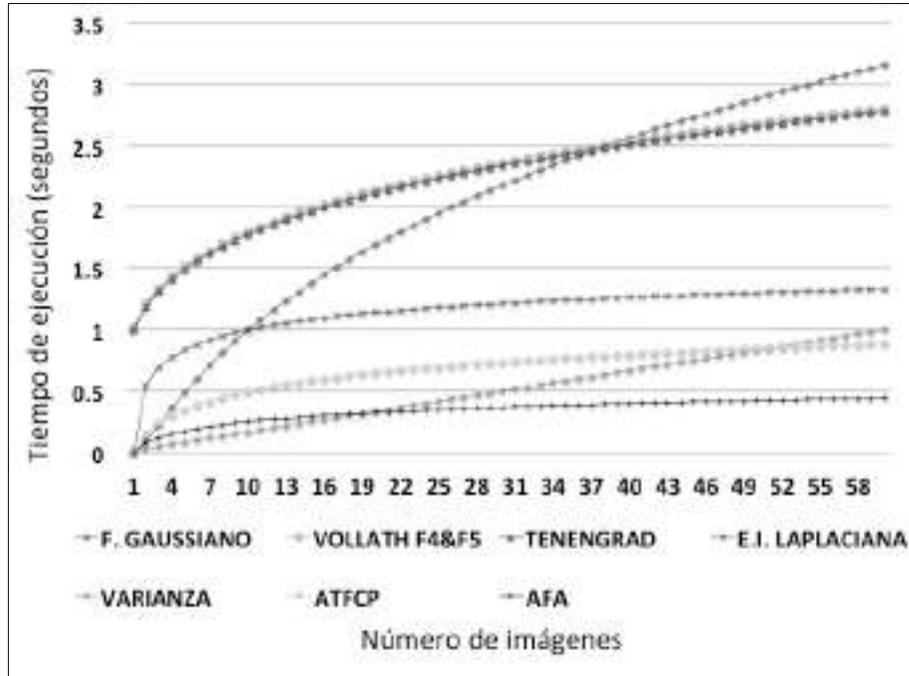


Fig. 23 El gráfico muestra en el eje x el nombre del algoritmo, en el eje y su correspondiente tiempo de ejecución en segundos, para el conjunto de 60 imágenes de *Tricodina*.

3.2 Evaluación de calidad de la fusión usando las métricas Q , Q_{wt} y Q_E

La calidad de la imagen de la fusión se evaluó utilizando la métrica de multi-imagen [Bueno-Ibarra, 2005b], que se describe en el capítulo 1, que calcula una medida de calidad que representa el porcentaje de mejora de la imagen resultante de la fusión comparada con el resto de las imágenes que conforman el subconjunto de la fusión.

Se realizaron diferentes pruebas al algoritmo de fusión propuesto. Una de ellas era comprobar si el número de elementos fusionados influye en el resultado, se encontró que fusionar una cantidad mayor o menor de imágenes sí modifica el resultado. Cada conjunto tiene característica diferentes por lo tanto requiere de una cantidad diferente de imágenes a fusionar, por lo que es importante hacer una selección de las imágenes que pertenecerán al subconjunto de fusión y así obtener mejores resultados. La **Fig. 24** muestra la selección de imágenes del conjunto *Tricodina* que se incluyen en la fusión.

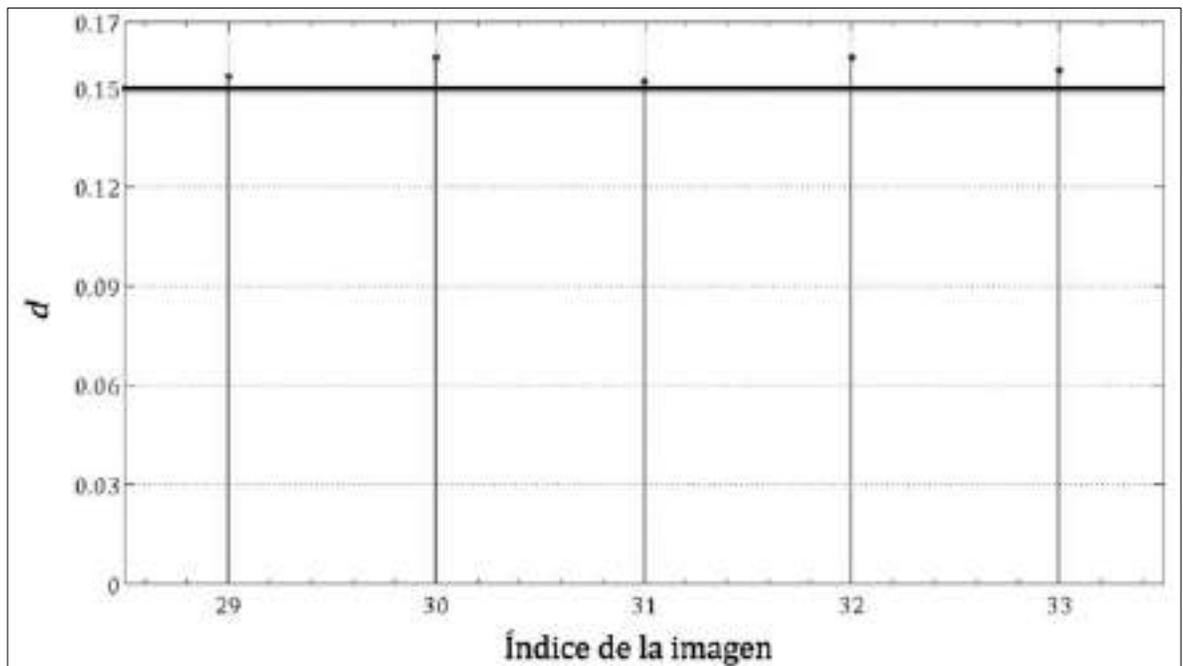


Fig 24 El valor máximo del índice de de-correlación se utiliza para determinar el rango de imágenes a incluir en la fusión. El gráfico muestra en el eje X los índices de las imágenes y en el eje Y su medida de de-correlación, d , para el conjunto de *Tricodina*.

Esto significa que la fusión con sólo dos imágenes no garantiza obtener la mejor fusión, ni tampoco fusionar una gran cantidad de imágenes garantiza que se obtenga algo de mejor calidad. Por esto, se en esta tesis se elaboró una fórmula Ec. (60), con la cual se puede determinar el número ideal de imágenes a incluir en el proceso de fusión. Esta fórmula se aplicó a los conjuntos de imágenes de prueba, una vez que se obtiene la IME, el valor máximo de de-correlación se utiliza para definir el rango de imágenes que serán parte de la fusión.

Las imágenes del proceso de la fusión que resultan del algoritmo AFS aquí propuesto en se muestran en las **Fig. 25** y la **Fig. 26**, la cantidad de elementos óptimos para la fusión se eligieron de acuerdo a la Ec. (60). La primera figura es del conjunto de sesenta imágenes de una muestra de *Tricodina* de las cuales cinco imágenes fueron fusionadas para obtener la mejor calidad de imagen, **Fig. 26(a)**. La segunda es del conjunto de diatomeas de cien imágenes de las cuales dieciséis fueron incluidas en la fusión **Fig. 26(b)**. La tercer figura es del conjunto de diez imágenes del limbo solar y sólo dos de ellas se eligieron para la fusión **Fig. 26(c)**. Y la cuarta es la imagen de fusión del automóvil de juguete, para el cual se fusionaron dos imágenes **Fig. 26(d)**.

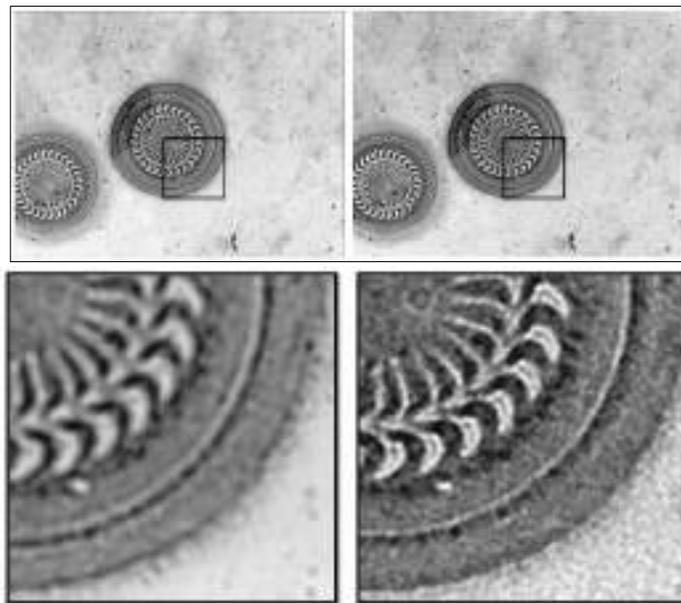


Fig. 25 Detalle de un acercamiento de la IME obtenida por el autoenfoque y la imagen que resulta de la fusión.

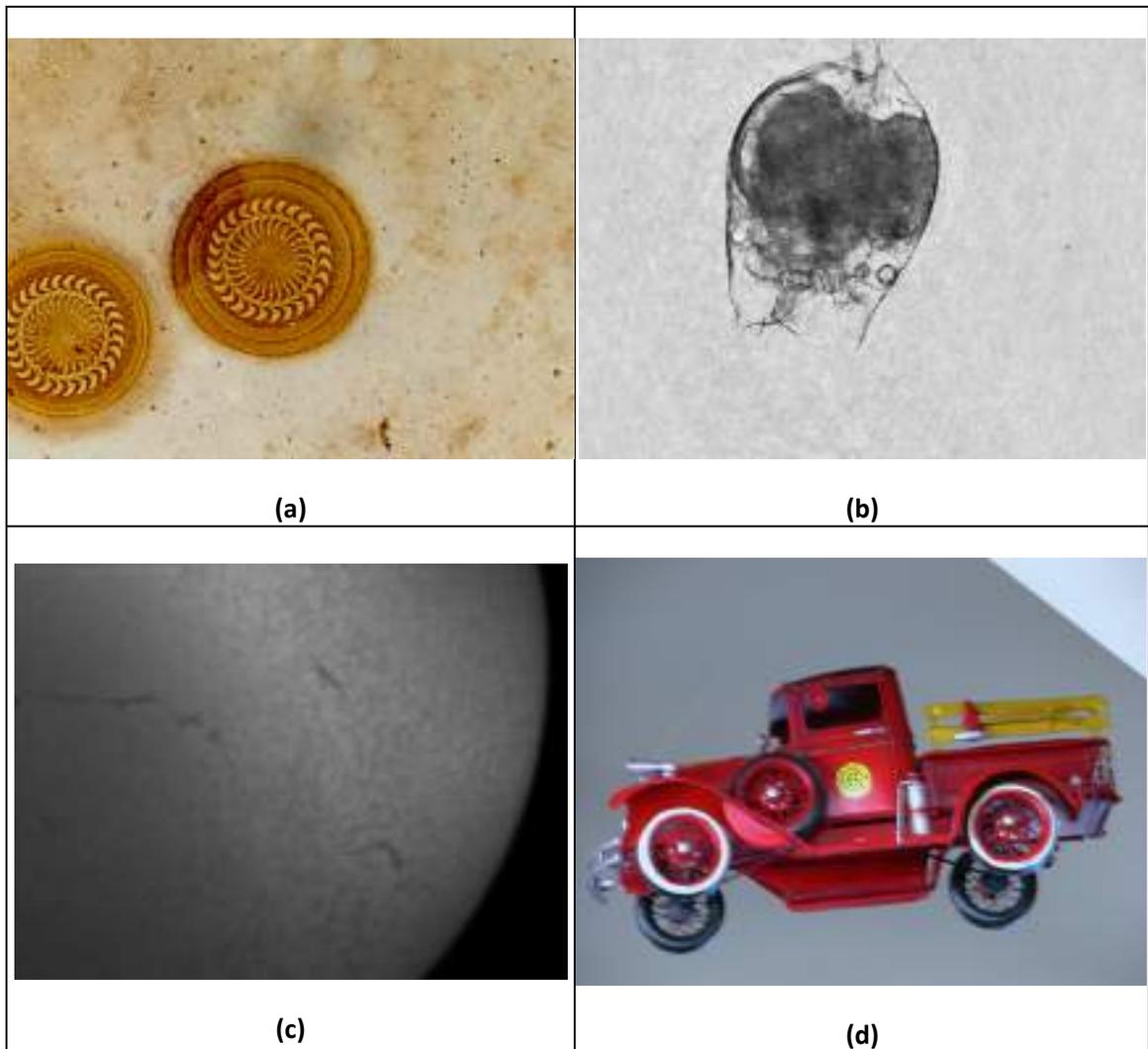


Fig. 26 Imágenes resultado de la fusión usando el algoritmo AFS.

Se analiza también, la eficiencia del algoritmo AFS comparando su eficiencia con las familias de wavelets más conocidas: *Haar*, *Daubechies*, *Coiflet*, *Symlet* y *Biortogonal*, [Mathworks, 2015b] con diferentes niveles de coeficientes, por su popularidad y funcionalidad. La fusión de la wavelet se implementó usando sus métodos: *max*, *min*, *max-min*, *mean* e *img1-img2*. Las imágenes resultantes de fusión de las imágenes que utilizan wavelets se muestran en la **Fig. 27**. Cada una de estas imágenes de fusión se evaluaron utilizando la métrica multi-imagen.

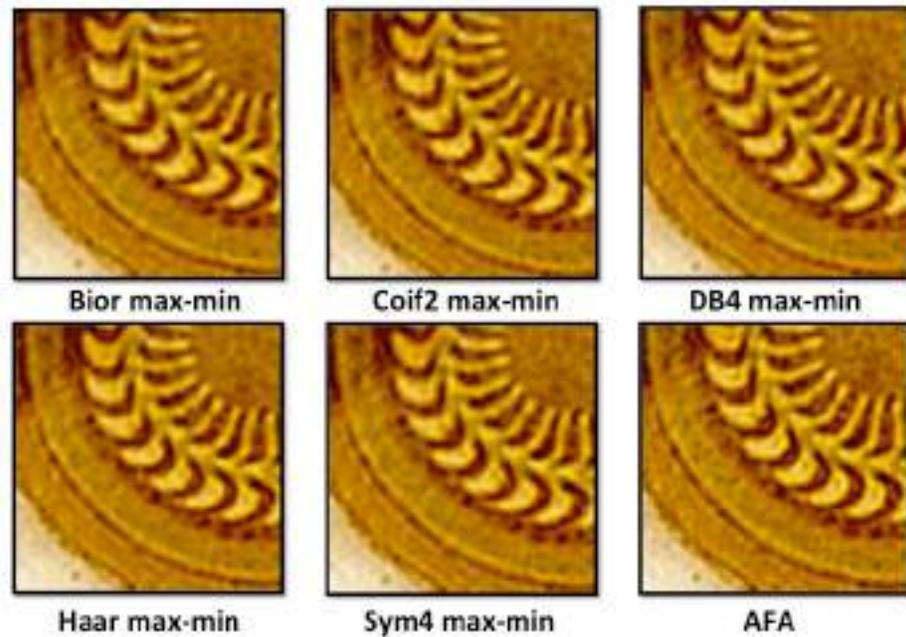


Fig. 27 Detalle de una sección de la imagen para comparar la calidad de fusión de las Wavelets con el método AFS .

La comparación de las medidas de calidad de imagen se describen en las **Tablas 4 y 5**. En éstas se muestran los porcentajes de los métodos que se implementaron, los métodos de fusión con wavelets y el método aquí propuesto, AFS. Se puede distinguir que el mejor resultado en esta evaluación fue obtenido por AFS. La medida de calidad de la fusión se define por las ecuaciones (41), (42) y (43).

Además de obtener la mejor fusión, AFS también optimiza los tiempos de fusión ya que al fusionar dos o más imágenes requiere de menos tiempo de máquina que al fusionar solo dos con otros métodos. El gráfico de la **Fig. 28** compara los tiempos de ejecución de los diferentes métodos de fusión de dos imágenes.

La métrica de calidad de enfoque Q_E emplea como uno de sus argumentos una función de detección de bordes de la imagen, el nivel utilizado de contribución de los bordes es $\alpha = 0.2$, el borde de la imagen se obtuvo por Laplacian de Gaussiano (LoG) método utilizando la función MATLAB "*edge*" [Mathworks, 2015c].

Fue necesario ajustar los parámetros umbral y sigma de la función LoG, a cada conjunto dependiendo de las características particulares, por ejemplo el conjunto de *Tricodina* utiliza el umbral $\tau = 0.01$, mientras que el conjunto de imágenes de limbo solar utiliza $\tau = 0.001$.

Tabla 4. Comparación de las medidas de calidad Q , Q_{WT} y Q_E de AFS con los otros algoritmos de fusión aquí analizados. Evaluación correspondientes empleando el conjunto de imágenes de Limbo Solar.

Algoritmo		Medida de calidad de la imagen			
		Q	Q_{WT}	Q_E	
AFS		0.9568	0.963	0.9593	
SA		0.8398	0.8758	0.7383	
SM		0.8554	0.8783	0.7433	
PCA		0.8952	0.903	0.8374	
DCT		0.92	0.9199	0.8468	
ROLP		0.867	0.8981	0.8314	
AFTPC		0.9466	0.9465	0.9208	
DWT	Mean	BIOR2.2	0.9553	0.9537	0.9537
		COIF2	0.9559	0.954	0.9549
		DB4	0.9559	0.954	0.9547
		HAAR	0.9553	0.9537	0.9537
		SYM4	0.9542	0.9532	0.952
	Max-Min	BIOR2.2	0.9462	0.9451	0.9531
		COIF2	0.9466	0.9449	0.9538
		DB4	0.9466	0.9449	0.954
		HAAR	0.9459	0.944	0.9531
		SYM4	0.9468	0.9449	0.9536
	Min	BIOR2.2	0.9442	0.9402	0.9492
		COIF2	0.9464	0.9409	0.9503
		DB4	0.9464	0.9409	0.9508
		HAAR	0.9455	0.9406	0.9503
		SYM4	0.9466	0.9409	0.9504
	Max	BIOR2.2	0.9353	0.9395	0.94
		COIF2	0.9362	0.9398	0.9404
		DB4	0.9363	0.9398	0.9408
		HAAR	0.9366	0.9396	0.9429
		SYM4	0.9359	0.9397	0.9402
Img1-img2	BIOR2.2	0.9372	0.9374	0.9339	
	COIF2	0.9373	0.937	0.9333	
	DB4	0.937	0.937	0.9335	
	HAAR	0.9381	0.9386	0.9373	
	SYM4	0.9374	0.9371	0.9336	

Tabla 5. Comparación de las medidas de calidad Q , Q_{WT} y Q_E de AFS con los otros algoritmos de fusión aquí analizados. Evaluación correspondientes empleando el conjunto de imágenes de Tricodina

Algoritmo		Medida de calidad de la imagen			
		Q	Q_{WT}	Q_E	
AFS		0.9845	0.9887	0.9443	
SA		0.864	0.901	0.745	
SM		0.88	0.9036	0.7501	
PCA		0.921	0.929	0.845	
DCT		0.9465	0.9464	0.8545	
ROLP		0.79	0.83	0.668	
AFTPC		0.9739	0.9738	0.9292	
DWT	Mean	BIOR2.2	0.9854	0.9886	0.9388
		COIF2	0.9854	0.9886	0.938
		DB4	0.9854	0.9886	0.9382
		HAAR	0.9854	0.9886	0.9388
		SYM4	0.9853	0.9886	0.939
	Max-Min	BIOR2.2	0.9815	0.9856	0.8801
		COIF2	0.9821	0.9871	0.9043
		DB4	0.9833	0.9871	0.9097
		HAAR	0.9812	0.9852	0.906
		SYM4	0.983	0.9869	0.909
	Min	BIOR2.2	0.9809	0.9872	0.9054
		COIF2	0.9811	0.9874	0.9059
		DB4	0.9832	0.9874	0.9063
		HAAR	0.985	0.9881	0.9046
		SYM4	0.9834	0.9875	0.9056
	Max	BIOR2.2	0.9739	0.9792	0.9154
		COIF2	0.9741	0.9794	0.9159
		DB4	0.9742	0.9794	0.9163
		HAAR	0.976	0.9811	0.9146
		SYM4	0.9744	0.9795	0.9145
Img1-img2	BIOR2.2	0.9783	0.9836	0.9265	
	COIF2	0.9786	0.9839	0.9266	
	DB4	0.9787	0.9839	0.927	
	HAAR	0.9788	0.9839	0.9263	
	SYM4	0.9787	0.9839	0.9268	

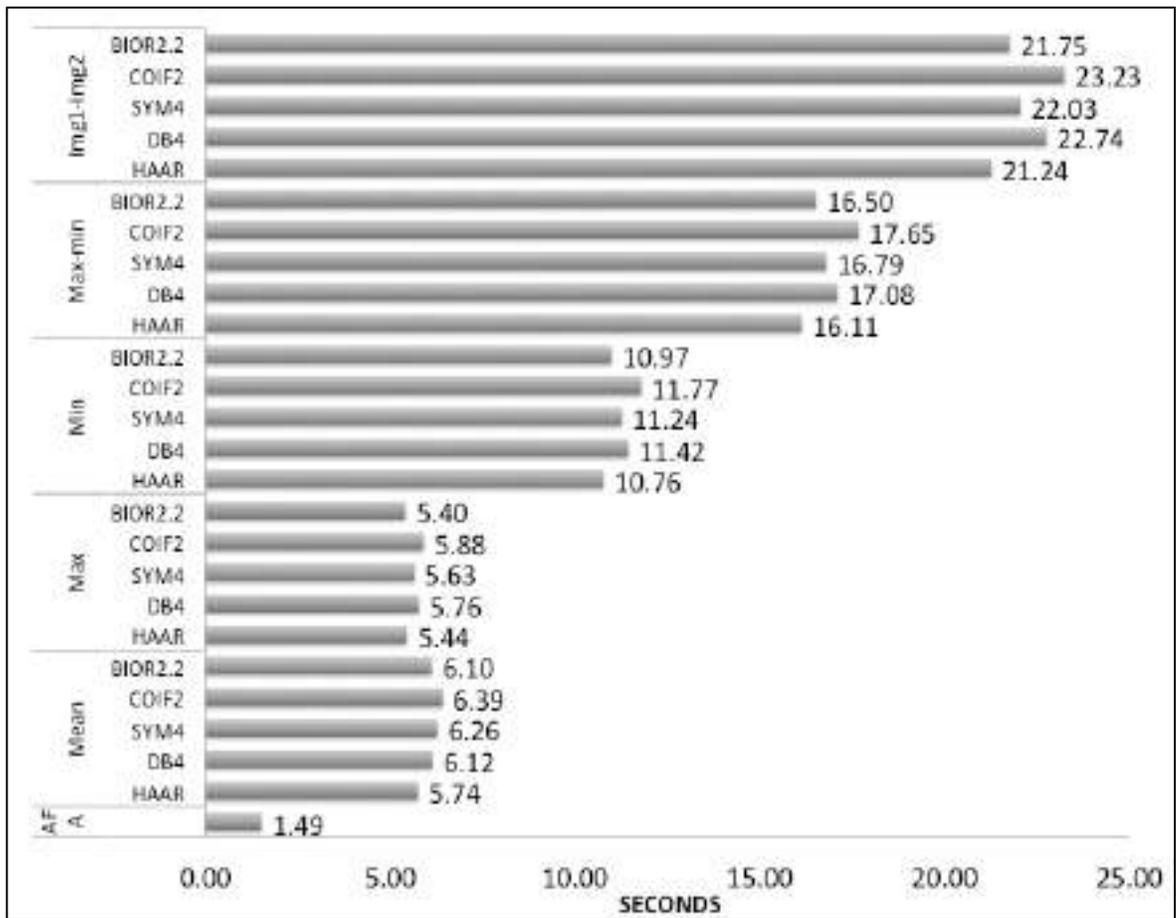


Fig. 28 Tiempo de ejecución en segundos, comparación de calidad de fusión DWT y AFS.

3.3 Discusión

La evaluación de los resultados muestra que AFS es un algoritmo robusto, eficaz y eficiente. Es robusto debido a que soporta el procesamiento de diferentes cantidades de imágenes, además emplea imágenes de diferentes formatos, tamaños, resolución y mapas de color. Es eficiente, porque cumple su objetivo al encontrar la imagen mejor enfocada y obtener una imagen mejorada como resultado de fusión. Es eficaz, porque es óptimo en complejidad en espacio y tiempo, es decir, optimiza el espacio de memoria requerido para procesar grandes cantidades de imágenes y es óptimo en tiempo, ya que obtiene los resultados del procesamiento en menor tiempo que otros algoritmos.

En comparación con los otros métodos de fusión que se estudiaron, AFS realiza la fusión con más de dos imágenes, eligiendo las imágenes a fusionar por un método automático de

selección que analiza los niveles de enfoque para fusionar solo las de mejores características, lo cual proporciona los elementos primordiales para obtener resultados de fusión de alta calidad.

El desarrollo de AFS cumple con los objetivos propuestos en esta tesis ya que de manera acertada selecciona la imagen con mejor enfoque en un tiempo óptimo, y como resultado de la fusión obtiene una imagen de alta calidad que supera a la de mejor enfoque, lo cual se comprueba en la evaluación de los resultados.

Se encontraron diferencias en los índices máximos de calidad de las imágenes pertenecientes a cada conjunto y la razón es la diferencia en las características en la captura de las imágenes, por ejemplo, las imágenes de limbo solar capturan un sujeto en movimiento y que cambia su apariencia en el tiempo, mientras que las imágenes de *Tricodina* capturan un sujeto estático que no cambia en el tiempo por ello las diferencias en los resultados de la evaluación.

Se implementaron variantes en los argumentos de las métricas para minimizar el tiempo de ejecución el cual era de orden exponencial. Se determinaron los valores óptimos para la detección de bordes, haciendo así una contribución a la medida de eficiencia.

Cómo trabajo futuro se sugiere la implementación del algoritmo en un sistema de adquisición de imágenes e integrarlo para el control del procesamiento de captura en tiempo real.

Capítulo 4

Conclusiones

- Se desarrolló un nuevo algoritmo AFS, para el autoenfoco y la fusión de imágenes el cual es robusto, eficaz y eficiente.
- La metodología propuesta optimiza los métodos de autoenfoco y fusión mediante la aplicación de la transformada de Fourier, la correlación no lineal y el filtro parabólico. También se reducen los tiempos de lectura y análisis de las imágenes empleando un método de barrido en espiral.
- El método de autoenfoco desarrollado en esta tesis selecciona la imagen mejor enfocada, IME, entre un conjunto de imágenes pre-capturadas y el método de fusión obtiene una imagen de alta calidad, conservando íntegra la cantidad y características de la información de la imagen.
- Soporta el procesamiento de imágenes de diferente formato, tamaño, resolución y mapa de color. Así también, pueden procesar conjuntos con diferentes cantidades de imágenes pre-capturadas.
- El algoritmo AFS es óptimo en su complejidad en tiempo y en espacio, ya que obtiene los resultados del procesamiento en menor tiempo que otros algoritmos y optimiza el espacio de memoria requerido para analizar grandes cantidades de imágenes.
- La metodología propuesta contribuye al procesamiento digital de imágenes, dado que aporta una opción de PDI en el dominio de frecuencias mejorada, aplicable a la investigación siendo posible su implementación en diversos campos de estudio donde interviene el PDI.

Referencias

1. Ahmed N, Natarajan T and Rao KR; 1974; “Discrete Cosine Transform”. IEEE Trans. On Computers, Vol.32, pp. 90- 93.
2. Akiyama A, Kobayashi N, Mutoh E, Kumagai H, Yamada H, Ishii H; 2008; “Space Imaging Infrared Optical Guidance for Autonomous Ground Vehicle”. Novel Optical Systems Design and Optimization XI, doi: 10.1117/12.794345; SPIE Vol. 7061, pp. 70610K-1-70610K-8.
<http://spiedigitallibrary.org>
3. Álvarez-Borrego J, Bueno-Ibarra MA, Chávez-Sánchez MC; 2008; “Autoenfoco y fusión”. En: Josué Álvarez-Borrego y María Cristina Chávez-Sánchez, Editores, “Introducción a la identificación automática de organismos y estructuras microscópicas y macroscópicas”. ISBN: 978-97027-1396-8. México. Capítulo 10, pp. 264
4. Baraiya S, Gagnani LP; 2014; “An Introduction of Image Fusion Techniques”. IJRST –International Journal for Innovative Research in Science & Technology, Volume 1, Issue 7, December 2014 ISSN (online): 2349-6010.
5. Brown G; 2000; "How Autofocus Cameras Work". HowStuffWorks.com.
<http://electronics.howstuffworks.com/autofocus.htm>
6. Bueno-Ibarra MA, Álvarez-Borrego J, Acho L, Chávez-Sánchez MC; 2005a; “Fast autofocus algorithm for automated microscopes”. Optical Engineering; 44(6), 063601
7. Bueno-Ibarra MA, Álvarez-Borrego J, Acho L, Chávez-Sánchez MC; 2005b; "Polychromatic image fusion algorithm and fusion metric for automatized microscopes". Optical Engineering, Vol. 44, No. 9, 093201, September 16. (PA: 33699). (SCI: 0.658, 2010).
8. Bueno-Ibarra MA; 2005c; “Desarrollo de una tecnología sistematizada para la adquisición y análisis de partículas biogénicas”. Tesis; Instituto Politécnico Nacional.
9. Deepak KS and Parsai MP; 2012; “Different Image Fusion Techniques –A Critical Review”. Department of Electronics & Communication Engineering, Jabalpur Engineering College, Jabalpur MP, India. International Journal of Modern Engineering Research (IJMER) www.ijmer.com Vol. 2, Issue. 5, Sep.- Oct. 2012 pp-4298-4301 ISSN: 2249-6645
10. Deepali AG and Dattatraya SB; 2011; “Wavelet based image fusion using pixel based maximum selection rule”. International Journal of Engineering Science and Technology (IJEST), Vol. 3 No. 7 July 2011, ISSN : 0975-5462
11. Desale R and Verma S; 2013; “Study and Analysis of PCA, DCT & DWT based Image Fusion Techniques”. International Conference on Signal processing, Image Processing and Pattern Recognition. [ICSIPRI], 978-1-4673-4862-1/13, ©2013 IEEE
12. Forero M, Sroubek F, Cristobal G; 2004; “Identification of tuberculosis bacteria wibased on shape and color”. Real-Time Imagen;10:251–262.

13. Geusebroek JM, Cornelissen F, Smeulders AW, Geerts H; 2000; "Robust autofocusing in microscopy". *Cytometry*, 39:1–9.
14. González RC and Woods, RE; 2008; "Digital image processing". By Pearson Education, Inc. Pearson Prentice Hall Pearson Education, Inc. Upper Saddle River, New Jersey 07458; ISBN 0-13-168728-x; 978-0-13-168728-8"
15. Groen FC, Young IT, Lighthart G; 1985; "A comparison of different focus functions for use in autofocus algorithms". *Cytometry*; Vol. 6, pp. 81–91.
16. Hilsenstein V; 2005; "Robust autofocusing for automated microscopy imaging of fluorescently labelled bacteria. DICTA ' 05". *Proceedings Digital Image Computing: Techniques and Applications*,. pp 95–101.
17. Horn B.K.P; 1968; "Focusing". *DSpace MIT*; <http://hdl.handle.net/1721.1/5863>
18. Jarvis RA; 1976; "Focus optimization criteria for computer image processing". *Microscope* 1976; 24(2), pp. 163–180.
19. Krotkov E; 1988; "Focusing. *Int J Comput Vision*". Vol 1, pp. 223–237.
20. Leica-Microsystems; 2013; "Leica Geschichte". <http://www.leica-microsystems.com/science-lab/history/videos-on-the-leica-history-german-only/>
21. Mansurov N; 2012; "How Phase Detection Autofocus Works". <https://photographylife.com/how-phase-detection-autofocus-works>
22. Mateos-Pérez JM, Redondo R, Nava R, Valdiviezo JC, Cristóbal G, Escalante-Ramírez B, Ruiz-Serrano MJ, Pascau J, Desco M; 2012; "Comparative evaluation of autofocus algorithms for a real-time system for automatic detection of Mycobacterium tuberculosis". *CIBERSAM*, Madrid, Spain.
23. MathWorks®; 2014a; "Introduction to Wavelet Families". Documentation. <http://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html?nocookie=true#f3-1009153>
24. MathWorks®; 2015b; "Image Fusion". http://www.mathworks.com/help/wavelet/examples/image-fusion.html?s_tid=srchtitle
25. Mathworks®; 2014c; "Edge". <https://www.mathworks.com/help/images/ref/edge.html?searchHighlight=log%20method>
26. Naidu VPS, Raol JR; 2008; "Pixel-level Image Fusion using wavelets and Principal Component Analysis". *Defence Science Journal*, Vol. 58, No. 3, May 2008, pp. 338-352 02008, DESIDOC.
27. Naidu VPS, 2012; "Discrete Cosine Transform based Image Fusion techniques", *Journal of Communication, Navigation and Signal Processing* (January 2012), Vol. I, No. I, pp. 35-45.
28. Pajares G and De la Cruz JM; 2004; "A wavelet-based image fusion tutorial". *Pattern Recognition Society*.
29. Piella G and Heijmans Henk; 2003; "A new quality metric for image fusion". *International Conference on Image Processing, CWI, Amsterdam, Netherlands*, Vol. 3, pp. III 173-6 vol. 2, Sep.
30. Rani K, Sharma R; 2013; "Study of Different Fusion Algorithm". *International journal of Emerging Technology and advanced Engineering*, Vol.3, May 2013.
31. Russell MJ, Douglas TS; 2007; "Evaluation of autofocus algorithms for tuberculosis microscopy". *Conference proceedings: Annual International*

- Conference of the IEEE Engineering in Medicine and Biology Society
2007:3489–3492.
32. Santos A, Ortiz De Solórzano C, Vaquero JJ, Peña JM, Malpica N, del Pozo F; 1997; “Evaluation of autofocus functions in molecular cytogenetic analysis”. *J Microscopy*; 188: 264–272.
 33. Saucedo-Morales J; 2015; “Imágenes del día”, “infraestructura”. Observatorio Solar Carl Sagan (OSCS). Área de Astronomía de la Universidad de Sonora, UNISON; Hermosillo, Sonora, México. <http://www.astro.uson.mx>; <http://www.astro.uson.mx/archimdia>; <http://www.astro.uson.mx/infraestructura>
 34. Schlag JF, Sanderson AC, Neuman CP, Wimberly FC; 1983; “Implementation of automatic focusing algorithms for a computer vision system with camera control”. Tech report CMU-RI-TR-83-14, Robotics Institute, Carnegie Mellon University, August, 1983.
http://www.ri.cmu.edu/publication_view.html?pub_id=61
 35. Subbarao M, Choi T, Nikzad A; 1992; “Focusing techniques”. *J Opt Eng* ; 32:2824–2836.
 36. Tenenbaum JM; 1970; "Accommodation in computer vision,". Ph.D. Thesis, Stanford Univ.
 37. Toet A; 1989; “Image fusion by a ratio of low pass pyramid”. *Pattern Recognition Letters* 9 (1989); Volume 9, Number 4, 245 -253; North-Holland.
 38. Vassil Dimitrov and Khan Wahid; 2004; “Multiplier less DCT Algorithm for Image Compression Applications”. *International Journal on Information Theories and Applications*, Vol.11, pp.162-169.
 39. Vollath D; 1988; “The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms”. *J Microsc* 151:133–146.
 40. Wang Z, Bovik AC; 2002; “A universal image quality index”. *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84.
 41. Zeder M, Pernthaler J; 2009; “Multispot live-image autofocusing for high-throughput microscopy of fluorescently stained bacteria”. *Cytometry Part A* 75A:781–788.

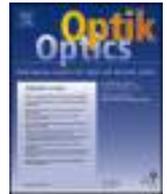
Anexos

1. Publicación en revista arbitrada

Cabazos-Marín, Alma Rocío & Álvarez-Borrego, Josué. (2018). “Automatic focus and fusion image algorithm using nonlinear correlation: image quality evaluation.”; *Optik* 164 10.1016/j.ijleo.2018.02.101; OPTIK ©2018 Elsevier, p. 224-242; <https://authors.elsevier.com/c/1Wj3Q6wQfGw6q>

2. Congreso Internacional

Alma Rocío Cabazos-Marín, Josué Álvarez-Borrego, and Ángel Coronel-Beltrán. (2014). “*Autofocus and Fusion using nonlinear Correlation.*”; *AIP Conference Proceedings* 1618, 147 (2014); doi: 10.1063/1.4897702; International Conference of Computational Methods in Sciences and Engineering (IMMSE 2014). Atenas, Grecia, 04-07 abril 2014. <https://aip.scitation.org/doi/abs/10.1063/1.4897702>



Original research article

Automatic focus and fusion image algorithm using nonlinear correlation: Image quality evaluation

Alma Rocío Cabazos-Marín^a, Josué Álvarez-Borrego^{b,*}

^a *Departamento de Investigación en Física, Universidad de Sonora (UNISON), Luis Encinas y Rosales S/N, Col. Centro, Hermosillo, Sonora, C.P. 83000, Mexico*

^b *Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), División de Física Aplicada, Departamento de Óptica, Carretera Ensenada-Tijuana No. 3918, Fraccionamiento Zona Playitas, Ensenada, Baja California, C.P. 22860, Mexico*

ARTICLE INFO

Article history:

Received 25 January 2018

Accepted 27 February 2018

Keywords:

Autofocus

Fusion

Digital image processing

Correlation

Image quality evaluation

Metrics

ABSTRACT

We propose a new autofocus and fusion algorithm to enhance the characteristics of the image in less run-time. The autofocus method selects the best-focused image (BFI) in a stack of images captured at different distances from the object. A vector is defined for each image in the stack, which contains elements selected by spiral scanning of the image. The spectrum of each vector is calculated using the Fourier transform then applying non-linear correlation to the reference vector spectrum and those of each of the corresponding to the images in the stack the BFI is determined. The fusion is carried out with a subset of images that have a focus measure value close to the BFI ones. The parabolic filter is applied to determine the relevant elements of the images that will be included in the fusion. The results evaluation is conducted using a multi-image metric, a quality measure that represents the percentage improvement of the fused image. The comparison was made with other fusion methods such as standard wavelets, resulting in that the autofocus and fusion algorithm (AFA) method obtained the highest quality indices. It is concluded that the AFA improves the quality of the images in less time than the conventional fusion methods.

© 2018 Elsevier GmbH. All rights reserved.

1. Introduction

Today, most imaging systems have automatic focus and fusion functions that minimize manual processes for obtaining high-quality images. Autofocus allows obtaining the best focal plane of an object or scene automatically. Fusion consists of taking the most characteristic visual features in different focal planes and merging them into a single image of higher quality. Both techniques play an essential role in digital image processing. The aim is to obtain a high-quality image that enhances the main features of the image object for further analysis, identification, or classification.

Most autofocus and fusion functions provide quality images, but this quality can be improved as well as the runtime. Most of the algorithms include only two pictures for fusion; in this work, a quality image obtained was enhanced by comprising more than two images for the fusion, achieving to fuse at less eight images in less time than are fused two by other methods. In this work, we propose an autofocus and fusion algorithm (AFA). The algorithm implements Fourier transform and non-linear correlation to find a best-focused image, as well as applies a parabolic filter to select the significant features of each picture.

* Corresponding author.

E-mail address: josue@cicese.mx (J. Álvarez-Borrego).

Autofocus and Fusion using Nonlinear Correlation

Alma Rocío Cabazos-Marín^a, Josué Álvarez-Borrego^{b,1}
and Ángel Coronel-Beltrán^a

^a Departamento de Investigación en Física, Universidad de Sonora (UNISON), Luis Encinas y Rosales S/N, Col. Centro, Hermosillo, Sonora, México C.P. 83000

^b Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), División de Física Aplicada, Departamento de Óptica, Carretera Ensenada-Tijuana No. 3918, Fraccionamiento Zona Playitas, Ensenada, Baja California, México C.P. 22860

Abstract. In this work a new algorithm is proposed for auto focusing and images fusion captured by microscope's CCD. The proposed algorithm for auto focusing implements the spiral scanning of each image in the stack $f(x, y)_w$ to define the V_w vector. The spectrum of the vector FV_w is calculated by fast Fourier transform. The best in-focus image is determined by a focus measure that is obtained by the FV_w nonlinear correlation vector, of the reference image, with each other FV_w images in the stack. In addition, fusion is performed with a subset of selected images $f(x, y)_{SBF}$ like the images with best focus measurement. Fusion creates a new improved image $f(x, y)_F$ with the selection of pixels of higher intensity.

Keywords: autofocus, fusion, nonlinear correlation, spiral scanning, parabolic filter.

PACS: 42.30.Va

INTRODUCTION

At present, most of the imaging systems have an autofocus function. The aim is to get a high quality image with a minimum of manual processes with the most similar characteristics possible to the real scene and also gather sufficient information for analysis. There are different autofocus algorithms for imaging quality [1-6]. One of the best autofocus systems can be found in [6], where the system is applied to an optic microscope. In this paper a modified algorithm is proposed following the main ideas given by Bueno, et al. [6].

In this work the nonlinear correlation is applied to obtain information on the similarity between the reference image and other images of the stack. The nonlinear correlation has the advantage that the correlation peak is well located and the output plane is less noisy [7], which provides an ability to discriminate between the best in-focus and less focused image in the stack of images.

METHODOLOGY

Getting Images Stack

The images used for testing the algorithms were captured by a CCD camera Leica model DC- 300 with a working resolution of 2088x1550 pixels. This camera is integrated on a Leica DM- RXA2 with motorized stage for X/Y and Z displacements.

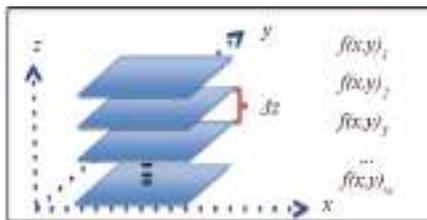


FIGURE 1. Stack of images captured by microscope, where $f(x, y)_w$ is the digital image with pixels (x, y) and Δz is the interval at different focal lengths on the axis Z.

¹ Corresponding author: josue@cicese.mx