

ADQUISICIÓN Y VISUALIZACIÓN DE ESCENAS CON FOCO VARIABLE UTILIZANDO PROCESAMIENTO DIGITAL DE IMÁGENES



PEDRO SANDINO ATENCIO ORTIZ¹

AUGUSTO ARIAS GALLEGO²

ALBERTO MAURICIO ARIAS CORREA³

RESUMEN

La profundidad de campo (DOF, por sus siglas en inglés) o profundidad de foco es una característica de un sistema óptico con una configuración en particular, que hace referencia al rango espacial o de distancia que dicha configuración permite mantener de forma nítida o enfocada. Una DOF alta en un sistema óptico es deseada en aplicaciones de microscopía y fotografía macro entre otras aplicaciones. Una solución al problema de una DOF limitada en un sistema óptico consiste en tomar secuencias de imágenes con diferentes distancias focales y posteriormente fusionar dichas imágenes en una sola imagen completamente enfocada. Este proceso es conocido como fusión de imágenes multi-foco (FIMF). La literatura sobre técnicas de FIMF es extensa y de actual interés en la comunidad científica de fusión de imágenes y fotografía computacional. Por el contrario, la literatura que ilustra el proceso completo de FIMF es escasa. En este trabajo es propuesto un método general de FIMF, el cual comienza con la configuración y calibración del sistema óptico, seguido de la implementación de la técnica de fusión y termina con la visualización de dichas imágenes. En este trabajo se utilizó un sistema óptico de bajo costo con una profundidad de campo variable en el rango $[0,18 m, \infty]$. El método propuesto es aplicable de forma directa para sistemas de adquisición con configuraciones ópticas diferentes.

PALABRAS CLAVE: procesamiento digital de imágenes; fusión de imágenes multi-foco; fotografía computacional.

ACQUISITION AND VISUALIZATION OF VARIABLE-FOCUS SCENES USING DIGITAL IMAGE PROCESSING

ABSTRACT

Depth of field (DOF) or focus range is a feature of an optical system with a specific configuration, that refers to the spatial or distance range that such configuration allows to maintain sharp or focused. A high DOF in an optic system is desirable in microscopy and macro photography among other applications. A solution to a limited DOF in an optic system is

¹ Ingeniero de sistemas, Universidad del Magdalena (Colombia). Magíster en Ingeniería en Sistemas, Universidad Nacional de Colombia, sede Medellín. Profesor, Instituto Técnico Metropolitano, Facultad de Ingenierías, Medellín (Colombia).

² Ingeniero físico, Universidad Nacional de Colombia, sede Medellín, (Colombia). Profesor, Instituto Técnico Metropolitano, Facultad de Ingenierías, Medellín (Colombia).

³ Ingeniero en instrumentación y control, Politécnico Jaime Isaza Cadavid, (Colombia). Magíster en Ingeniería de Sistemas, Universidad Nacional de Colombia, sede Medellín (Colombia). Profesor, Instituto Técnico Metropolitano, Facultad de Ingenierías, Medellín (Colombia).



Autor de correspondencia: Atencio-Ortiz, P.A. (Pedro-Asencio). Instituto Técnico Metropolitano, Facultad de Ingenierías, Medellín, Colombia. Correo electrónico: pedroatencio@itm.edu.co. Tel: (57+4) 446 10 43

Historia del artículo:

Artículo recibido: 29-XI-2013 / Aprobado: 30-IX-2015

Disponible online: 30 de octubre de 2015

Discusión abierta hasta noviembre de 2016



to take a sequence of images with different focal distances and subsequently fuse those images in one that is completely focused. This process is known as multi-focus image fusion (MFIF). Literature about techniques of MFIF is extensive and of current interest in scientific community of image-fusion. In contrast, literature about complete MFIF process is scarce. In this work it is proposed a general method of MFIF, which begins with configuration and calibration of optical system, followed by implementation of fusion technique and ends with visualization of fused image. In this work was used a low-cost optical system with a variable DOF in range $[0,18 m, \infty]$. Proposed method is directly applicable for acquisition systems with different optical configurations.

KEYWORDS: Digital Image Processing; Multi-Focus Image Fusion; Computational Photography.

AQUISIÇÃO E VISUALIZAÇÃO DE CENAS COM FOCO VARIÁVEL USANDO PROCESSAMENTO DE IMAGENS DIGITAIS

RESUMO:

A profundidade de campo (DOF, por sua sigla em Inglês) ou a profundidade de foco é uma característica de um sistema óptico com uma configuração particular, que se refere ao intervalo espacial ou a distância que esta configuração permite manter de modo afiada ou focada. Uma alta DOF num sistema óptico é desejado em aplicações de microscopia e macro fotografia e outras aplicações. Uma solução para o problema da limitada DOF num sistema óptico é levar seqüências de imagens com diferentes distâncias focais e, em seguida, mesclar as imagens em uma imagem completamente focada. Este processo é conhecido como fusão de imagens multi-foco (FIMF). A literatura sobre técnicas FIME é extensa e de atual interesse na comunidade científica de fusão de imagens e fotografia computacional. Por outro lado, a literatura que ilustra o processo completo de FIMF é escassa. Neste trabalho, propõe-se um método FIMF geral, que se inicia com a configuração e calibração do sistema óptico, seguido pela aplicação da técnica de fusão e termina com a visualização de tais imagens. Neste trabalho foi usado um sistema óptico de baixo custo com uma profundidade variável de campo no intervalo $[0,18 m, \infty]$. O método proposto é aplicável de maneira direta para sistemas de aquisição com diferentes configurações ópticas.

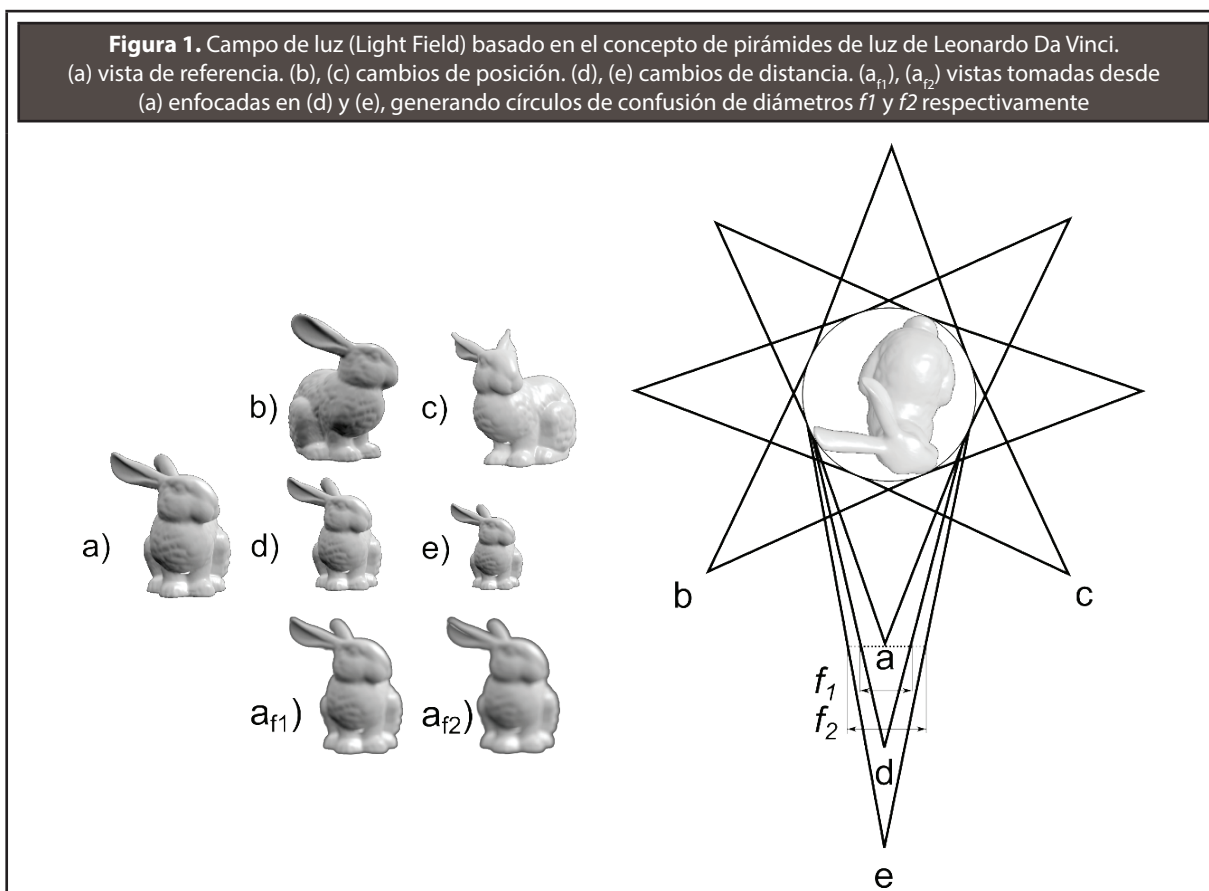
PALAVRAS-CHAVE: processamento de imagem digital; fusão de imagens multi-focais; fotografia computacional.

1. INTRODUCCIÓN

Cualquier sistema actual de cámara fotográfica de uso común solo puede adquirir un estado del campo de luz (Levoy, 2006) o según Da Vinci (Richter, 1970) como describe Adelson *et al.*, en (Adelson y Wang 1992), una «pirámide o cono de luz», esto es, la luz reflejada por un objeto vista desde una posición, distancia y enfoque respecto al mismo. En la **Figura 1** se ilustra el concepto de campo de luz utilizando el concepto de pirámides de luz. Las cámaras convencionales solo permiten adquirir una de estas vistas para un instante de tiempo t , lo cual se convierte en una limitación debido a que solo se

puede tomar una porción ínfima de la información visual que contiene la escena real.

Obtener toda o gran parte de dicha información, constituye un ideal en múltiples disciplinas como la óptica y la fotografía computacional. Con esta información una escena podría ser vista desde infinitas orientaciones y distancias focales, posterior a su adquisición. Dicha información es conocida en fotografía computacional como campo de luz (*Light Field*) y se refiere a una función que describe la cantidad de luz viajando en cada dirección y punto en el espacio. Adquirir este de forma completa este campo sin embargo, converge en una imposibilidad práctica.



La naturaleza parcial o limitada de la captura del campo de luz de una escena, conlleva a problemas inevitables tales como oclusiones entre objetos de la escena, el problema de la óptica inversa (Zygmunt, 2001) y profundidad de campo limitada.

Una forma aproximada de obtener una porción del campo de luz, consiste en obtener vistas parciales del campo de luz (multi-temporales, multi-modales, multi-vista o multi-foco) y posteriormente fusionar dichas imágenes o relacionarlas para reconstruir parte del campo de luz. El proceso anterior recibe el nombre de fusión de imágenes.

El proceso de fusión de imágenes depende de la naturaleza de la fuente de las imágenes o de la porción del campo de luz que se esté reconstruyendo. Dicha fuente puede ser temporal (capturas en distintos instantes de tiempo), modal (capturas a través de múltiples sensores), espacial (capturas de una escena desde distintas posiciones) o focal

(capturas con distinto plano focal). El proceso de fusionar imágenes del último tipo (focal), se conoce como fusión de imágenes multi-foco, el cual tiene como objetivo mejorar la *DOF* de un sistema de óptico, mediante la captura de múltiples imágenes con focos en objetos a distintas distancias y posteriormente generar una imagen completamente enfocada, en la que todos los objetos de la imagen aparezcan enfocados.

La fusión de imágenes multi-foco es de especial interés en aplicaciones en microscopía (Song *et al.*, 2006), reconstrucción tridimensional (Saeed y Choi 2008; Favaro y Soatto 2007), fotografía computacional y visión por computador (Wan *et al.*, 2013).

Los métodos de FIMF pueden agruparse en dos categorías: dominio de transformación y dominio espacial.

La primera categoría se basa en transformar la imagen del dominio espacial a otro dominio, fusionar los coeficientes de transformación y posteriormente obtener la imagen fusionada en el dominio espacial mediante una transformación inversa de los coeficientes fusionados (Liu *et al.* 2015; Nejati *et al.* 2015). Los métodos de dominio de transformación están basados principalmente en transformaciones multi-escala, por ejemplo descomposición piramidal (Liu *et al.*, 2001), transformada discreta *wavelets* (DWT, por sus siglas en inglés) (Santhosh *et al.*, 2014; Wang *et al.*, 2013). Algunos algoritmos novedosos como análisis de componentes principales robusto (*robust principal component analysis*) (Wan *et al.*, 2013) y representaciones escasas (*sparse representations*) también se han utilizado para la fusión de imágenes (Yang y Li, 2012; Chen *et al.*, 2013).

Por el contrario, en los métodos de FIMF basados en dominio espacial, los mecanismos de fusión son aplicados directamente para cada píxel en las imágenes de entrada o regiones de píxeles en las mismas. Según la forma en que estos métodos operan sobre las imágenes, se pueden clasificar en: basados en píxel/bloques y basados en regiones. Los métodos de FIMF basados en píxel/bloques utilizan operaciones píxel a píxel de las imágenes de entrada o sobre bloques rectangulares de tamaño definido. En algunos casos, la FIMF basada en píxel/bloques puede entenderse como un problema de clasificación (Saeedi y Faez, 2009; Li *et al.*, 2002) en el cual por cada píxel o bloque de las imágenes de entrada, debe decidirse cuál de todos los píxeles o bloques presenta un mayor grado de nitidez, y posteriormente se reemplaza el píxel o bloque completo en la imagen fusionada con el píxel o bloque seleccionado. Utilizar píxeles o bloques para la FIMF puede generar artefactos en la imagen fusionada en imágenes de escenas naturales, debido a que la disposición y forma de los objetos en la misma no es regular. Por esta razón trabajos recientes utilizan estructuras de particionamiento irregular tipo *QuadTree* (Bai *et al.*, 2015; De y Chanda, 2015). Los métodos basados en regiones van acompañados de

métodos de segmentación adaptativos, mediante los cuales se determina la región exacta que está enfocada dentro de una imagen. Estos métodos normalmente utilizan algoritmos de optimización para generar las regiones o para refinar el resultado parcial del algoritmo de segmentación. Li *et al.* (2013) proponen un método que inicialmente genera una segmentación gruesa, en la cual se seleccionan píxeles con altos niveles de enfoque y posteriormente en una segunda se optimizan las regiones segmentadas mediante *image matting* para generar una separación precisa entre el fondo (región desenfocada) y el primer plano (región enfocada). Nejati *et al.* (2015) presentan un método basado en clasificación mediante representaciones escasas (*sparse representations*) basadas en diccionarios para detectar píxeles con valor alto de enfoque y posteriormente utiliza campos aleatorios de Markov (*Markov Random Fields*) para optimizar las regiones clasificadas y generar un mapa de decisión suave en cada imagen de entrada. Los métodos de FIMF que utilizan optimización producen los mejores resultados reportados en la literatura, sin embargo todos presentan problemas de alto costo computacional.

Todo proceso de FIMF, independiente del método aplicado, requiere utilizar una métrica de estimación de foco, la cual entrega un valor directamente proporcional al grado de nitidez de un píxel o regiones de píxeles en la imagen. Generalmente cualquier operador que describa cambios en los niveles de intensidad de la imagen puede ser utilizado como métrica de enfoque. Los descriptores de textura pueden brindar información sobre el nivel de nitidez o enfoque de una imagen. Lorenzo *et al.* (2008) realizan una exploración sobre la utilización del descriptor de texturas *Local Binary Pattern* como descriptor de enfoque. Algunos de los descriptores comúnmente utilizados en tareas de FIMF y tecnologías de autoenfoco en cámaras son: *varianza de la imagen*, *Laplacianos*, *energía de los gradientes* (*EOG*, *Energy Of Gradients*), *energía de los Laplacianos* (*EOL*) (Subbarao *et al.* 1993). Modificaciones

recientes y mejoras de los descriptores anteriores son: *Laplaciano Modificado*, *Suma del Laplaciano Modificado*, *Frecuencia Espacial* y *Tenengrad*. Un estudio y evaluación de distintos descriptores de foco puede ser consultado en (Huang y Jing, 2007).

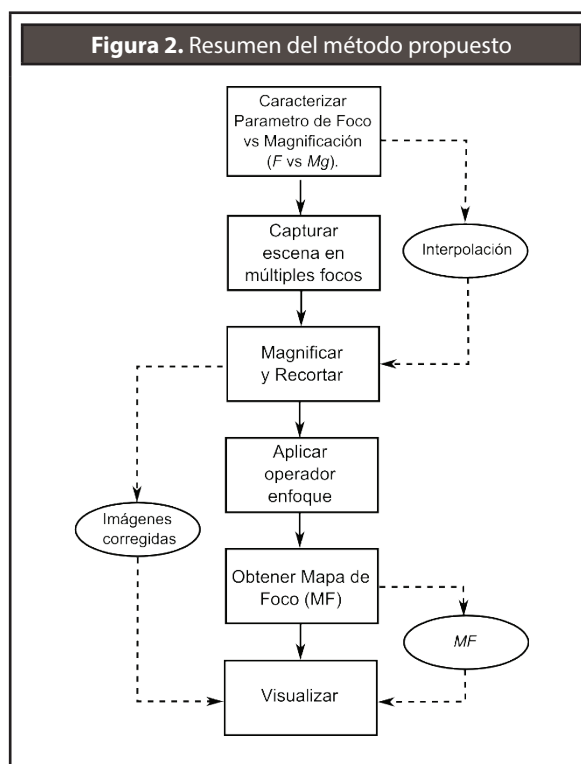
Aunque la literatura es extensa en métodos y técnicas de FIMF, usualmente solo abarcan la etapa de procesamiento de imágenes, es decir, el análisis de las imágenes de entrada y la generación de la imagen fusionada. Sin embargo, aplicar un proceso de FIMF implica resolver distintos retos técnicos que van desde la selección de la óptica, calibración del sistema óptico y la corrección de las imágenes multi-foco (también conocido como proceso de registro), hasta el proceso de visualización de imágenes de este tipo. Por otra parte, los trabajos encontrados en la literatura utilizan sistemas ópticos de laboratorio para la toma de las imágenes haciendo impráctico el proceso de FIMF fuera del mismo.

En este trabajo se propone un método para adquirir, procesar y visualizar escenas con múltiple enfoque, utilizando para ello, sistemas fotográficos de tipo consumidos, con foco variable y técnicas de procesamiento digital de imágenes. El trabajo se encuentra organizado de la siguiente manera: En el capítulo 2 se ilustra el método propuesto que comienza con la calibración del sistema óptico y corrección de las imágenes, así como las técnicas de descripción de enfoque, fusión de las imágenes de entrada y visualización de la imagen fusionada. En el capítulo 3 se detalla la experimentación y los resultados obtenidos y por último en el capítulo 4 se exponen las conclusiones y trabajos futuros.

2. MATERIALES Y MÉTODOS

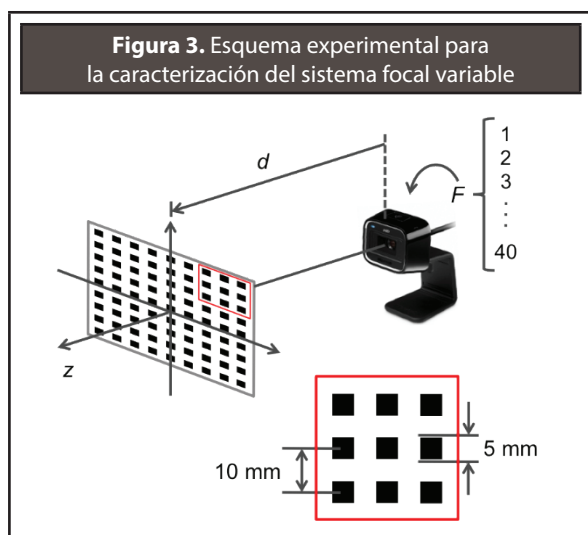
El método propuesto inicia con la caracterización del enfoque digital F respecto a la magnificación producida en la imagen. Dicha caracterización consiste en realizar una interpolación que relacione ambos parámetros. A continuación, es realizada la adquisición de las imágenes de la escena en distintos valores de F y se aplica la interpolación, para escalar

y recortar cada una de estas con el fin de eliminar el efecto de la magnificación generada en cada toma. Posteriormente es aplicado un *operador de enfoque/desenfoque* a cada píxel de cada una de las imágenes adquiridas, con lo cual es obtenida una matriz de $n \times m \times k$ donde $n \times m$ hace referencia al tamaño de la imagen y k al número de tomas con distintos F . Cada valor de esta matriz indica que tan enfocado está un píxel para un valor específico de F . Utilizando esta información se procede a obtener un *mapa de foco MF* de dimensión $n \times m$ el cual contiene el valor de F en el cual cada píxel de la imagen alcanza el mayor valor enfoque. Finalmente, el visualizador funciona desplegando la imagen en el foco F_1 , y permitiendo que el usuario seleccione un píxel en una posición (i, j) arbitraria en la misma. Luego se promedian los valores de una ventana de tamaño $l \times l$ centrada en la posición (i, j) en la matriz MF y se procede a visualizar la imagen correspondiente al valor promedio obtenido. En la **Figura 2** se puede observar el resumen del método propuesto, por medio de un diagrama de flujo.



A. Caracterización de F vs magnificación

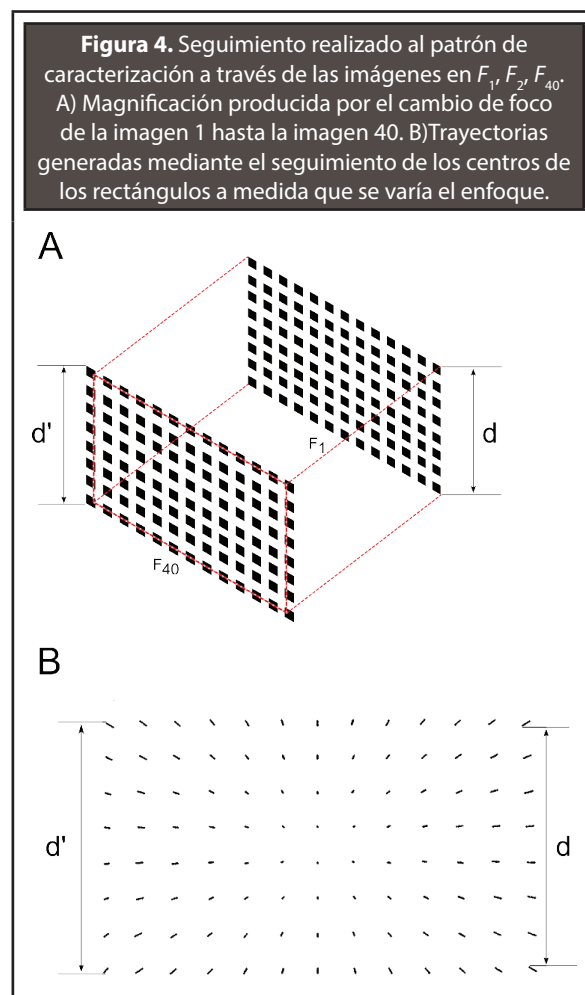
Debido a que el cambio del valor F genera una magnificación de la imagen, esta debe ser caracterizada con el fin de poder realizar una corrección posterior y mantener la escena original en cada una de las imágenes. Este proceso de caracterización consiste en realizar un ajuste de curvas o interpolación entre alguna medida de magnificación y el respectivo valor F . Para ello es utilizada una cuadrícula que permita realizar un seguimiento de los cuadrados en las imágenes de la secuencia F_1, F_2, \dots, F_{40} manteniendo fija la distancia d entre la cuadrícula y la cámara (ver **Figura 3**).



Utilizando la imagen F_1 como referencia se puede calcular el cambio de posición en altura y ancho de cada punto de interés en la imagen y con ello obtener una diferencia de altura y ancho en el patrón utilizado, lo cual equivale a una medida de magnificación en la imagen.

En la **Figura 4** (A) se muestra la magnificación producida de la imagen con foco F_1 a la imagen con foco F_{40} . Esta magnificación produce pérdida de parte de la información presente en la primera imagen. En la **Figura 4** (B) se pueden observar las trayectorias obtenidas mediante el seguimiento de los centros de los cuadrados del patrón en las 40 imágenes adquiridas con diferentes valores de F . Nótese que

la distribución de la magnificación en la imagen es homogénea del centro hacia los bordes de la misma. Esto permite eliminar el efecto de la magnificación generada por el enfoque digital mediante un algoritmo de escalado digital.

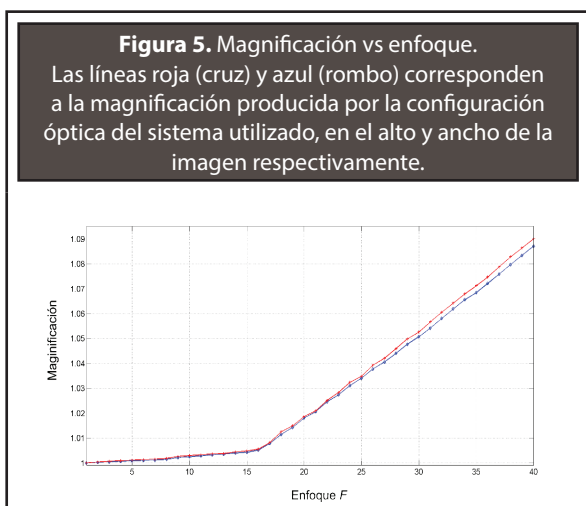


En la **Ecuación 1** se utiliza el promedio de los cocientes de las posiciones en filas o columnas, de las trayectorias mencionadas anteriormente para obtener los valores de magnificación en cada valor de foco F_i respecto al primer valor de foco F_1 .

$$Mg_F = \frac{\sum_{i=1}^n \frac{P_i}{P_1}}{n} \quad (1)$$

Donde Mg_F es la magnificación obtenida (en uno de los dos ejes de la imagen, *filas* o *columnas*)

para el valor de foco F , n es el número de trayectorias obtenidas equivalente al número de cuadrados del patrón utilizado y P_{iF} es la posición (*filas o columnas*) del centro del rectángulo i en el foco F . Cabe aclarar que la magnificación para las filas (alto de la imagen) y para las columnas (ancho de la imagen) debe ser calculada por separado.



En la **Figura 5** se puede observar la curva de magnificación obtenida para el sistema óptico utilizado en este trabajo. En esta se ve claramente que es una curva compuesta por dos tramos lineales. El primer tramo va desde F_1 a F_{17} el cual tiene un crecimiento lento. El segundo tramo va desde F_{18} a F_{40} y tiene un crecimiento en magnificación mucho mayor.

Para este caso, el primer tramo consiste en los valores de enfoque correspondientes a distancias de 27 cm (aprox.) para F_{17} y al infinito para F_1 , pero siendo la variación en enfoque muy pequeña y una profundidad de campo muy grande. Por lo anterior, solo resulta de interés para esta aplicación el segundo grupo de tramo (F_{18} a F_{40}).

Debido a la alta linealidad de la magnificación Mg_F respecto al valor de F en el segundo tramo de la caracterización realizada al sistema óptico utilizado (ver **Figura 1**), se pudo realizar una interpolación lineal simple (**Ecuación 2**), utilizando los primeros y últimos valores de F y Mg_F obtenidos mediante la **Ecuación 1**.

$$M_F = Mg_{min} + (Mg_{max} - Mg_{min}) \frac{F - F_{min}}{F_{max} - F_{min}} \quad (2)$$

Donde M_F es el valor de magnificación estimado para el valor de foco digital F . Mg_{min} y Mg_{max} son los valores de magnificación obtenidos mediante la **Ecuación 1** para valores de enfoque F_{min} y F_{max} respectivamente. En este trabajo, $F_{min} = 18$ y $F_{max} = 40$.

Generalmente los fabricantes de dispositivos ópticos con valor de enfoque paramétrico, se aseguran de que la relación del parámetro de control respecto a la magnificación generada en la imagen sea de tipo lineal. Sin embargo, el tipo de interpolación aplicada en la caracterización del sistema óptico, puede variar según las características propias del sistema utilizado.

B. Corrección

El proceso de adquisición de imágenes consiste en tomar n imágenes con valores diferentes de F , manteniendo la escena estática y la configuración espacial del montaje de la cámara. Para cada imagen es conocido su valor de F .

Como fue mencionado anteriormente, este proceso de adquisición con foco variable genera una magnificación (**Figura 5**), la cual debe ser corregida con el fin de no se aprecie en el posterior proceso de cálculo del mapa de enfoque. Partiendo del hecho de que el tamaño de la imagen se conserva, es claro que existirán zonas de la imagen sin magnificación (F_{18}) que no aparecen en las imágenes en (F_{19}, \dots, F_{40}). Por lo tanto, la mayor información recuperable de la escena será la que aparece en la imagen con F_{40} . Debido a lo anterior, el proceso de corrección consta de dos etapas; la primera consiste en cambiar el tamaño de la imagen dependiendo de su factor de magnificación, y la segunda en recortar la misma para que solo aparezca la información presente en la imagen F_{40} . De lo anterior, es claro que el cambio de tamaño consiste en llevar cualquier imagen entre F_{18}, \dots, F_{39} a la magnificación generada en F_{40} , lo cual es posible calculando el valor inverso de la magnificación de la imagen que se quiere corregir, debido a que el comportamiento

de la magnificación es lineal. Por lo tanto el factor de corrección para la magnificación de la imagen quedaría descrito mediante la **Ecuación 3**.

$$CM_F = 0,034(58 - F) + 0,9386; F = \{18, \dots, 39\} \quad (3)$$

Sea $\delta(I, k)$ una función que magnifica una imagen I en un factor k entonces la primera etapa de corrección de la imagen quedaría descrita por la **Ecuación 4**.

$$IR_F = \delta(I_F, CM_F); F = \{18, \dots, 39\} \quad (4)$$

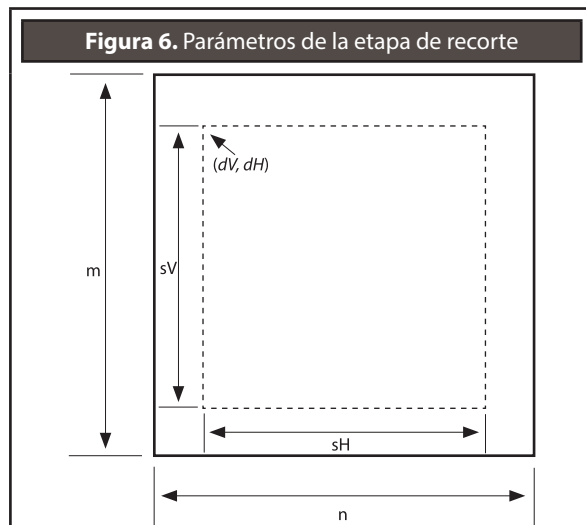
Donde IR_F es la imagen corregida en la primera etapa, I_F la imagen correspondiente al foco digital F .

La segunda etapa de corrección consiste en recortar la porción de las imágenes F_{18}, \dots, F_{39} que no es visible en la imagen de mayor magnificación F_{40} . Este proceso consiste en recortar los bordes de dichas imágenes de tal forma que la dimensión final de estas sea igual a la resolución de la cámara.

Sean sV y sH el número de píxeles verticales y horizontales de la cámara, y $\varepsilon(I, dH, dV, sH, sV)$ una función que retorna una porción de la imagen I de resolución (m, n) , equivalente a un rectángulo situado en el punto (dV, dH) y con incremento horizontal y vertical de sV, sH respectivamente (ver **Figura 6**); dV, dH estarán definidos mediante las **Ecuaciones 5 y 6**.

$$dV = \lfloor 0,5 * |m - sV| \rfloor \quad (5)$$

$$dH = \lfloor 0,5 * |n - sH| \rfloor \quad (6)$$



C. Cálculo del mapa de foco

Determinar el valor de F que alcanza el mayor valor, para cada píxel de la imagen es el objetivo del cálculo del mapa de foco. Para ello se requiere una métrica de enfoque que permita medir numéricamente que tan enfocado está un píxel en particular. En este trabajo se utiliza la «Suma de los Laplacianos Modificados» (*Sum of Modified Laplacian*) (Nayar y Nakgawa, 1994) como operador de enfoque/desenfoque debido a la robustez del mismo para responder variaciones de alta frecuencia en la imagen (bordes o regiones de alto contraste) y a la simplicidad de su implementación, lo cual no implica que otros operadores puedan ser utilizados en implementaciones futuras. Este operador es definido mediante las **Ecuaciones 7 y 8**.

$$ML(x, y) = |2 * IR(x, y) - IR(x - \text{paso}, y) - IR(x + \text{paso}, y)| + |2 * IR(x, y) - IR(x, y - \text{paso}) - IR(x, y + \text{paso})| \quad (7)$$

$$G_F(i, j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} ML(x, y) \quad (8)$$

Donde $ML(x, y)$ es el operador *Laplaciano Modificado* para el píxel en la posición x, y , y paso es un valor que permite «saltar» el vecino cercano más próximo en dicha proporción. $G_F(i, j)$ es una función que retorna el valor de la suma del Laplaciano modificado para un píxel en la posición i, j utilizando una ventana de $N \times N$ vecinos cercanos para la imagen correspondiente al foco digital F .

Utilizando las **Ecuaciones 7 y 8**, el proceso de calcular el mapa de foco consiste en encontrar por cada píxel de las imágenes adquiridas, el F con mayor valor, equivalente a mayor enfoque (**Ecuación 9**).

$$MF(i, j) = \text{argmax}(G_F(i, j)); F = \{18, \dots, 40\} \quad (9)$$

Finalmente, se aplica el filtro bilateral propuesto en (Tomasi y Manduchi, 1998) para evitar efectos producidos por el ruido del sistema de adquisición.

$$MF(i, j) = BF(MF, k, \sigma) \quad (10)$$

Donde $BF(MF, k, \sigma)$ es una función que retorna la imagen producto de la aplicación del filtro bilateral sobre la imagen de entrada MF utilizando un tamaño de vecindario local k y una desviación estándar σ . La imagen MF debe ser de tipo flotante en el rango $[0,1]$.

D. Selección de la imagen

Esta etapa consiste en visualizar la imagen que presenta mayor foco en una región seleccionada por el usuario mediante una interfaz gráfica.

Sea $SF(i,j,k)$ una función que retorna el promedio de F en una región definida por un kernel cuadrado de lado k y centrada en el píxel en la posición i, j , entonces:

$$SF(i,j,k) = \left\| \frac{\sum_{x=i-t}^{x=i+t} \sum_{y=j-t}^{y=j+t} ML(x,y)}{k^2} \right\| ; t=k \setminus 2 \quad (11)$$

El objetivo de utilizar un kernel en el cálculo anterior radica en utilizar información de foco alejada al píxel por el cual se está preguntado, partiendo de la suposición de que píxeles cercanos deben presentar un nivel de foco similar. Esto minimiza el efecto de artefactos generados en los procesos anteriores en el cálculo del valor de F .

Finalmente, la visualización de la imagen enfocada consiste en desplegar la imagen corregida, correspondiente al valor SF .

E. Obtención de una imagen con enfoque completo

Una imagen reenfocada IFa puede ser calculada mediante el promedio de las imágenes calculadas en diferentes focos (Levoy, 2006) (Ecuación 12).

$$IFa = \frac{\sum_{F=18}^{40} IR_f}{n} ; n = 18 - 40 \quad (12)$$

Donde IR_f a la imagen corregida en el foco F . Sin embargo, la imagen resultante presenta un efecto de suavizado o borrosidad (blur) debido a que

para cada zona de la imagen solo se tiene una imagen en la cual presenta enfoque máximo.

Utilizando el mapa de foco MF previamente calculado, es posible obtener una imagen IFb en la cual todos sus píxeles aparezcan en el enfoque máximo (Ecuación 13).

$$IFb(i) = IR_f(i); i = (MF == F) \quad (13)$$

Donde i es un vector R^2 de posiciones con valor F en el mapa de foco MF , es decir, un vector de posiciones de píxeles que presentan un máximo valor para el enfoque F . Sin embargo esta imagen presenta ruido producido por el carácter local del operador de enfoque.

En este trabajo se propone la combinación de las ecuaciones anteriores para obtener una imagen con enfoque completo IFc tal que presente un alto contraste y bajo nivel de ruido (Ecuación 14).

$$IFc = \frac{IF_a + IF_b}{2} \quad (14)$$

3. RESULTADOS Y DISCUSIÓN

Se realizó un experimento para validar la viabilidad de la utilización del método propuesto en escenas simples y escenas complejas, así como para determinar las condiciones necesarias para la correcta implementación del mismo.

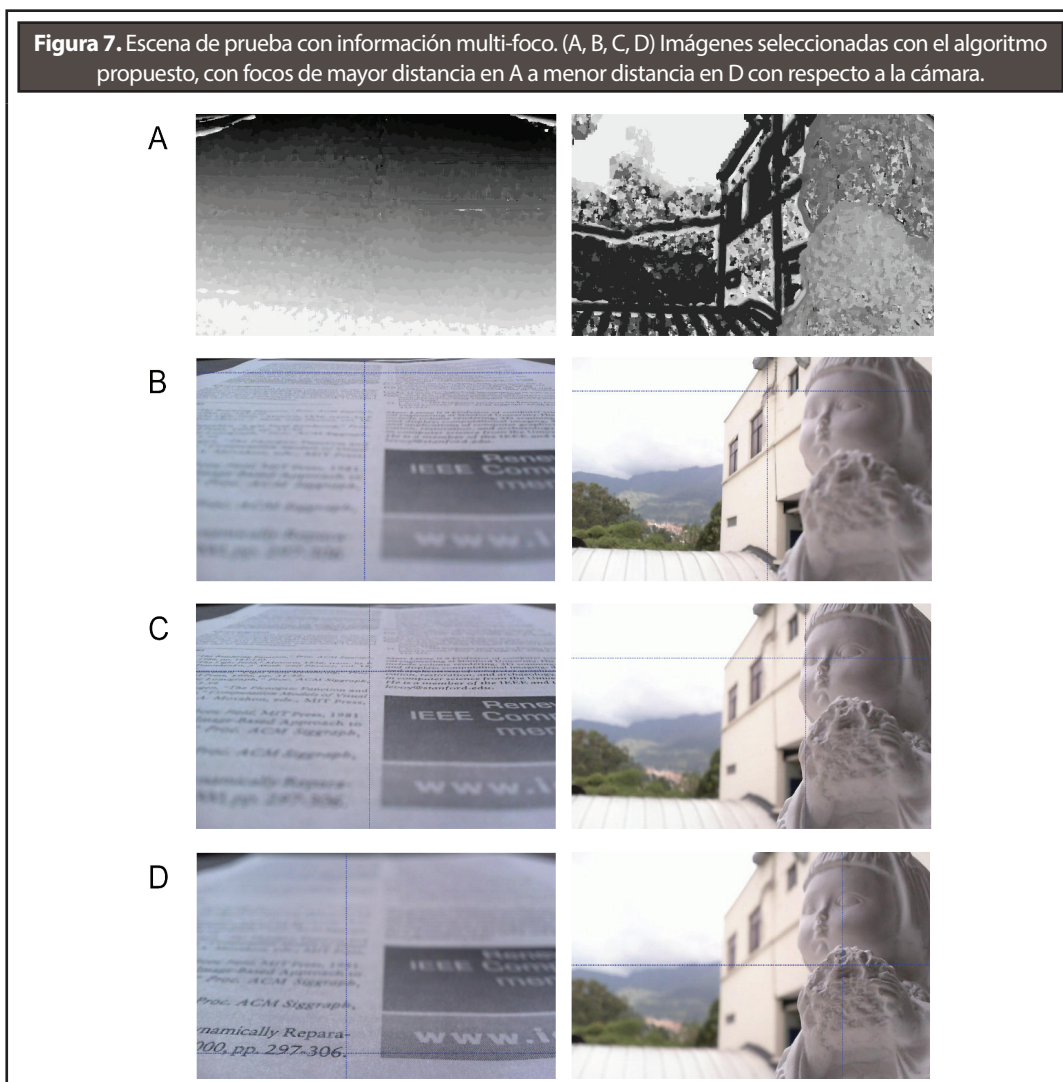
Para la experimentación fueron realizadas adquisiciones de escenas con condiciones que representan posibles casos de aplicación. Los algoritmos fueron implementados sin ningún tipo de optimización o paralelización de código, debido a que el objetivo del experimento no fue el de determinar la complejidad computacional del método.

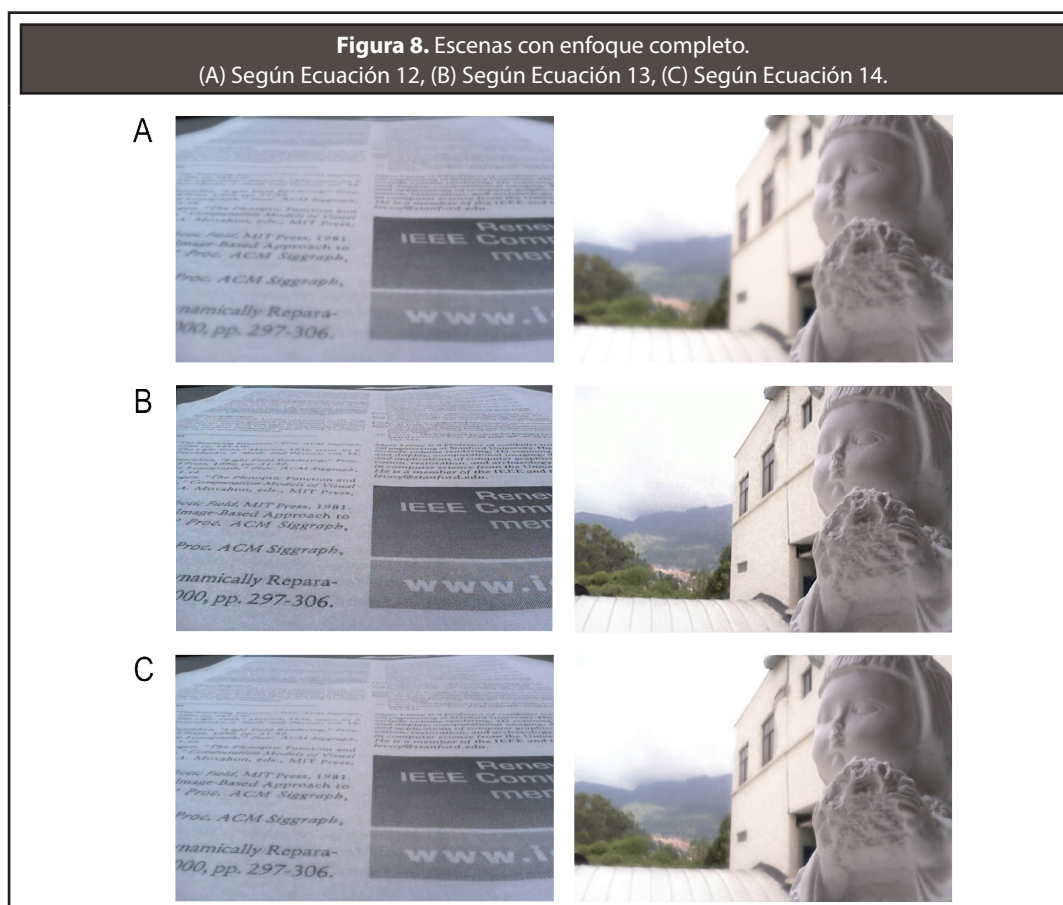
En la **Figura 7** se observa el resultado del método propuesto, aplicado a dos escenas; la primera una escena con alto contraste en condiciones de laboratorio y la segunda una escena natural que presenta zonas con alto y bajo contraste. En el grupo A se puede observar el mapa de enfoque obtenido mediante la **Ecuación 10**. Nótese que las regiones de bajo contraste en la escena 2 (nubes, pared de

edificación), generan zonas de confusión en el mapa de foco, debido a que el operador de enfoque retorna un valor parecido para cada una de las imágenes en distintos valores de F . Por el contrario, en la escena 1, el mapa de foco corresponde de forma precisa a la información de enfoque real de la escena, debido al alto contraste que genera el texto impreso. Nayar *et al.* (1996) proponen la utilización de un patrón de luz blanca sobre la escena con el fin mejorar el contraste y por ende la precisión del operador de enfoque. Sin embargo, esto no es una opción viable en escenas naturales, ya sea porque el patrón de luz debe proyectarse con un valor alto de lumen en escenas con mucha iluminación

o porque este patrón altera la imagen original. Una posible opción para afrontar esta limitación en un sistema comercial radicaría en estimar de manera automática un determinado número de puntos de enfoque que el usuario puede seleccionar, esto es, puntos con alto contraste local.

Los grupos B, C y D corresponden a la selección de imagen determinada mediante la **Ecuación 11** para zonas con distinto enfoque, del más lejano al más cercano respectivamente, utilizando una interfaz gráfica de usuario para seleccionar los puntos en la escena en los cuales se deseaba enfocar (línea azul).





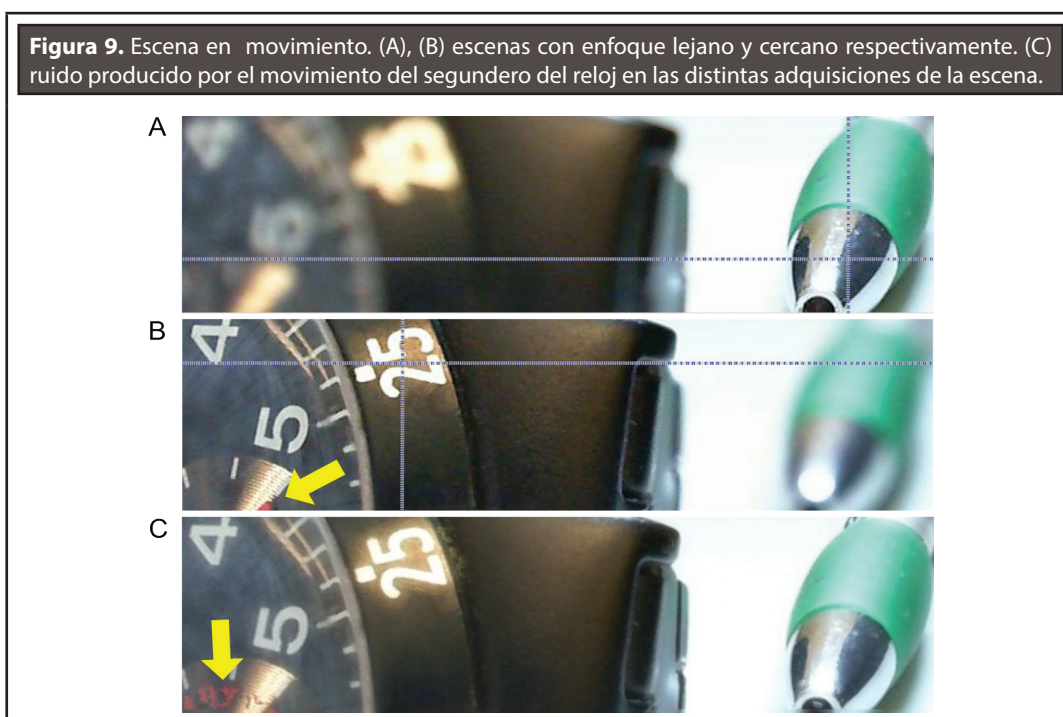
En la **Figura 8** se observa el resultado de la obtención de imagen con foco completo para las dos escenas de la **Figura 7**, resultado de utilizar la técnica del promedio de imágenes (A), el mapa de foco (B) y la combinación de ambas (C). Nótese que el grupo (A) presenta un nivel de suavizado o blur alto; El grupo (B) presenta un mayor contraste, pero también ruido generado por la naturaleza local de la obtención mapa de foco; el grupo (C) muestra el promedio de ambas imágenes (A y B) logrando una apariencia visual agradable, con un nivel de contraste menor que el obtenido en (B) pero con un nivel de menor de ruido.

Una limitación importante de este trabajo radica en la imposibilidad de adquirir escenas en movimiento, debido a que las imágenes para cada enfoque se toman en tiempos distintos, lo cual implica

que la escena se modifique en el tiempo y por ende el mapa de foco calculado no sea absoluto.

En la **Figura 9** se puede apreciar lo anterior. Nótese que debido al movimiento del segundero del reloj en la escena se producen artefactos en la imagen con enfoque completo (C).

Tecnologías tales como las lentes líquidas (Oku y Ishikawa, 2010) y las cámaras MEMS (Wei *et al.*, 2012) que permiten velocidades de enfoque aproximadas de 10 ms permitirían afrontar esta limitación. Por otra parte, debido a que el mapa de enfoque tiene relación con el valor de enfoque de la escena, realizando una calibración previa del valor de enfoque vs distancia de enfoque, sería posible utilizar el método propuesto para calcular el mapa de rango de la escena para aplicaciones de reconstrucción tridimensional en tiempo real tales como microscopía tridimensional y navegación robótica.



4. CONCLUSIONES

En este trabajo se presentó un método para la adquisición y visualización de imágenes de escenas con múltiples enfoques. La implementación del método es simple y permite ser utilizado en escenas con condiciones no controladas, lo cual resulta de interés para su implementación en teléfonos móviles o cámaras personales.

Este método contempla la caracterización inicial del sistema de adquisición, lo cual permite la implementación del mismo, en sistemas con configuraciones ópticas distintas a las utilizadas en este trabajo.

Debido a que el método requiere de un operador de enfoque y este a su vez trabaja con medidas locales de contraste, se pueden obtener resultados indeseados en el mapa de enfoque de las regiones de la escena con poco contraste, sin embargo esto depende en gran parte de la resolución del sensor de la cámara, ya que a mayor resolución, mayor es el nivel de detalle y por ende de contraste de los objetos en la escena.

Como trabajo futuro se proponen la extensión de este método para obtener el mapa de profundidad de

la escena adquirida y la implementación de sistemas de enfoque rápido como lentes líquidas para su aplicación en escenas con movimiento.

REFERENCIAS

- Adelson, E. H. and Wang, J. (1992). Single Lens Stereo with a Plenoptic Camera. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(2), February, pp. 96-106.
- Xiangzhi, Bai, Yu Zhang; Fugen Zhou and Bindang Xue (2015). Quadtree-Based Multi-Focus Image Fusion Using a Weighted Focus-Measure. *Information Fusion*, 22, March, pp. 105-118.
- Chen, L., Li, J. and Philip-Chen, C.L. (2013). Regional Multifocus Image Fusion Using Sparse Representation. *Optics Express*, 21(4), pp. 5182-5197.
- De, I. and Chanda, B. (2015). Multi-Focus Image Fusion Using a Morphology-Based Focus Measure in a Quad-Tree Structure. *Information Fusion*, 14(2), April, pp.136-146.
- Favaro, P. and Soatto, S. (2007). 3-D Shape Estimation and Image Restoration: Exploiting Defocus and Motion-Blur. Springer-Verlag London.

- Huang, W. and Jing, Z. (2007). Evaluation of Focus Measures in Multi-Focus Image Fusion. *Pattern Recognition Letters*, 28(4), March, pp. 493-500.
- Levoy, M. (2006). Light Fields and Computational Imaging. *Computer, IEEE Computer Society*, 39(8), August, pp. 46-55.
- Li, S. *et al.* (2013). Image Matting for Fusion of Multi-Focus Images in Dynamic Scenes. *Information Fusion*, 14(2), April, pp. 147-162.
- Li, S., Kwok, J. T. and Wang, Y. (2002). Multifocus Image Fusion Using Artificial Neural Networks. *Pattern Recognition Letters*, 23(8), June, pp. 985-997.
- Liu, C. *et al.* (2015). Multi-Focus Image Fusion Based on Spatial Frequency in Discrete Cosine Transform Domain. *Signal Processing Letters*, 22(2), February, pp. 220-224.
- Liu, Z. *et al.* (2001). Image Fusion by Using Steerable Pyramid. *Pattern Recognition Letters*, 22(9), July, pp. 929-939.
- Lorenzo, J. *et al.* (2008). Exploring the Use of Local Binary Patterns as Focus Measure. In *Computational Intelligence for Modelling Control & Automation*, 2008 International Conference on. Vienna, 10-12 December.
- Nayar, S.K. and Nakgawa, Y. (1994). Shape From Focus. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(8), August, pp. 824-831.
- Nayar, S.K.; Watanabe, M. and Noguchi, M. (1996). Real-Time Focus Range Sensor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(12), December, pp. 1186-1198.
- Nejati, M.; Samavi, S. and Shirani, S. (2015). Multi-Focus Image Fusion Using Dictionary-Based Sparse Representation. *Information Fusion*, 25, pp. 72-84.
- Oku, H. and Ishikawa, M. (2010). High-Speed Liquid Lens for Computer Vision. In *Robotics and Automation (ICRA), IEEE International Conference on. Anchorage, AK*, 3-7 May.
- Richter, J.P. (1970). *The Notebooks of Leonardo Da Vinci*. 2nd ed. Dover Publications. 514 p.
- Saeed, A. and Choi, T.-S. (2008). A Novel Algorithm for Estimation of Depth ap Using Image Focus for 3D Shape Recovery in the Presence of Noise. *Pattern Recognition*, 41(7), pp. 2200-2225.
- Saeedi, J. and Faez, K. (2009). Fisher Classifier and Fuzzy Logic Based Multi-Focus Image Fusion. In *Intelligent Computing and Intelligent Systems*, 2009. ICIS 2009. IEEE International Conference on, 4, pp. 420-425.
- Santhosh, J., *et al.* (2014). Application of SiDWT with extended PCA for multi-focus images. In *Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, 2014 International Conference on, November, pp. 55-59.
- Song, Y. *et al.* (2006). A New Wavelet Based Multi-focus Image Fusion Scheme and Its Application on Optical Microscopy. In *Robotics and Biomimetics*, 2006. ROBIO '06. IEEE International Conference on, December, pp. 401-405.
- Subbarao, M., Chor, T.-S. and Nikzad, A. (1993). Focusing Techniques. *Optical Engineering*, 32(11), November, pp. 2824-2836.
- Tomasi, C. and Manduchi, R. (1998). Bilateral Filtering for Gray and Color Images. In *Computer Vision*, 1998. Sixth International Conference on. January, pp. 839-846.
- Wan, T., Zhu, C. and Qin, Z. (2013). Multifocus Image Fusion Based on Robust Principal Component Analysis. *Pattern Recognition Letters*, 34(9), July, pp. 1001-1008.
- Wang, X., Wei, Y.-L. and Liu, F. (2013). A New Multi-Source Image Sequence Fusion Algorithm Based on SIDWT. In *ICIG'13 Proceedings of the 2013 Seventh International Conference on Image and Graphics*, pp. 568-571.
- Wei, H. *et al.* (2012). Controlling a MEMS Deformable Mirror in a Miniature Auto-Focusing Imaging System. *Control Systems Technology, IEEE Transactions on*, 20(6), pp. 1592-1596.
- Yang, B. y Li, S. (2012). Píxel-Level Image Fusion with Simultaneous Orthogonal Matching Pursuit. *Information Fusion*, 13(1), January, pp. 10-19.
- Zygmunt, P. (2001). Perception Viewed as an Inverse Problem. *Vision Research*, 41(24), November, pp. 3145 - 3161.

**PARA CITAR ESTE ARTÍCULO /
TO REFERENCE THIS ARTICLE /
PARA CITAR ESTE ARTIGO /**

Atencio-Ortiz, P.S.; Arias-Gallego, A.; Arias-Correa, M.A. (2015). Adquisición y visualización de escenas con foco variable utilizando procesamiento digital de imágenes. *Revista ELA*, 12(24), Julio-diciembre, pp. 13-25. [Online]. Disponible en: DOI: <http://dx.doi.org/10.14508/reia.2015.12.24.13-25>