

Algoritmos para el tratamiento de imágenes de campos de luz

Gustavo Bizai, Facundo Peiretti, Adrián Salvatelli, Bartolomé Drozdowicz,
Alejandro Hadad, Diego Evin

Laboratorio de Sistemas de Información, Facultad de Ingeniería, Universidad
Nacional de Entre Ríos
Ruta 11 Km.10 Oro Verde, Entre Ríos, Argentina
E-mail: {gbizai, asalvatelli, bdrozdo}@bioingenieria.edu.ar

Resumen

El Proyecto de Investigación y Desarrollo “Sistema de Información Plenóptica como medio diagnóstico para Lámparas de Hendidura” propone la utilización de campos de luz como estrategia para agregar información tridimensional a la imagen del fondo ocular. Una de las líneas de trabajo consiste en el desarrollo de algoritmos para la obtención de enfoque dinámico, multiperspectiva y mapa de profundidades a partir de imágenes plenópticas que se generan en una única toma. La percepción de profundidades constituye un valioso apoyo a la toma de decisiones en múltiples aplicaciones. Se presentan algunas metodologías para el tratamiento computacional de capturas con una cámara plenóptica, utilizando imágenes de dominio público mientras se está trabajando en la implementación de la óptica asociada a una lámpara de hendidura. Se muestra resultados de los algoritmos y se discute su eficiencia en términos de su complejidad y tiempo de procesamiento. El trabajo futuro es optimizar estas propuestas, implementar nuevas y obtener mapas de profundidad.

Palabras clave: Campos de Luz, Fotografía Plenóptica, Pila Focal.

Contexto

La línea de I+D+i presentada es parte del Proyecto de Investigación y Desarrollo PIDUNER N°6146 “Sistema de Información Plenóptica como medio diagnóstico para Lámparas de Hendidura”, financiado por la Universidad Nacional de Entre Ríos y ejecutado en el ámbito de la Facultad de Ingeniería (FI-UNER).

Introducción

La fotografía de fondo de ojos es de especial interés para la detección temprana de patologías [1-3] así como también el seguimiento de la evolución de los tratamientos [4]. Una de las complicaciones de la diabetes es el edema macular, cuyo diagnóstico temprano es fundamental. Su evaluación clínica es siempre difícil, siendo la fotografía estereó y la biomicroscopía con lámpara de hendidura las técnicas más utilizadas para mostrar cambios en el volumen retiniano. Sin embargo, ambas técnicas son dependientes del observador y no permiten la medición objetiva del cambio [5]. Asimismo, la lámpara de hendidura, equipo presente en cualquier consultorio oftalmológico, no ha evolucionado significativamente desde el punto de vista tecnológico [6]. La posibilidad de obtener imágenes de la topografía del fondo

ocular es de gran interés para el médico oftalmólogo como soporte objetivo para la toma de decisiones, máxime si se utiliza un equipo económico y de difusión masiva como lo es dicha lámpara. El valor agregado de información 3D de estas imágenes, incrementa su valor diagnóstico.

El PIDUNER abarca tres líneas de trabajo: a) caracterización de la lámpara de hendidura y obtención de datos para la decisión de compra de equipamiento de fotografía plenóptica; b) desarrollo de la algorítmica para la obtención de imágenes multifoco, multiperspectiva y mapa de profundidades; c) aplicación de metodologías para extracción de características de imágenes 3D. Las primeras dos líneas mencionadas se trabajan en paralelo y mientras se está desarrollando la óptica de captura, se ha comenzado a estudiar los algoritmos con imágenes plenópticas públicas dado que las metodologías son las mismas.

Este trabajo presenta la generación y evaluación de diferentes algoritmos para el renderizado de imágenes a partir de los campos de luz [7]. El campo de luz de una escena es el conjunto de todos los rayos de luz que refleja, en cualquier punto del espacio, en cualquier dirección. En otras palabras, es la radiancia en el espacio como función de la posición y dirección. Al conjunto de todos los rayos de luz que ocupa el espacio durante todo el tiempo se los describe con la función plenóptica [8]. La insistencia de la “totalidad” de esta función es la razón de su nombre: óptica plena, completa, llena. La función plenóptica describe la luz en 7 dimensiones: espacio (3D), dirección (2D), frecuencia (1D) y tiempo (1D). Es posible reducir esta alta dimensionalidad a 4D si se utiliza una parametrización de dos planos paralelos [9] y se consideran imágenes estáticas. Con esta información 4D es posible obtener imágenes 2D

enfocadas en diferentes profundidades de la escena, así como también mostrar la imagen desde diferentes ángulos. Con este conjunto de imágenes denominado “pila focal”, puede reconstruirse un modelo de la topografía de la escena.

En este trabajo se presentan algunos algoritmos para renderizar imágenes en diferentes profundidades y perspectivas, a partir de los campos de luz.

Líneas de Investigación, Desarrollo e Innovación

La investigación pertenece al Área Computación Gráfica, Imágenes y Visualización. La Fotografía Plenóptica [10] es una forma de Fotografía Computacional, en la cual la cámara plenóptica [11] recolecta información más rica que una cámara convencional para luego someterla a procesos computacionales que dan como resultado una visualización de la escena en diferentes profundidades y perspectivas.

Resultados y Objetivos

La captura del campo de luz se realiza con una cámara a cuyo sensor de imagen se antepone una matriz de lenticulas.

Codificación del Campo de Luz

Las imágenes crudas consisten en un conjunto de $M \times N$ subimágenes de $P \times Q$ pixeles cada una (figura 1). $M \times N$ es la cantidad de lenticulas que constituyen la matriz de microlentes, mientras que $P \times Q$ son los pixeles del elemento sensor de la cámara, que constituyen las subimágenes de cada lenticula. El tamaño de la imagen cruda de campo de luz es de $(M \times P) \times (N \times Q)$ pixeles. A partir de estas capturas es posible una visualización multiperspectiva y refocalización

computacional. La información de cada pixel se codifica en una matriz 5D que permite direccionarlo según la subimagen a la que pertenece (posición de la lenticula M,N) y su posición dentro de ella (P,Q) , para cada canal de color (ch) . El script consiste en varios bucles anidados que distribuyen cada pixel a una posición (M,N) y una dirección (P,Q) para cada canal de color. Requiere como variables la imagen a ser decodificada, la cantidad de lenticulas y la cantidad de pixeles por lenticula y devuelve la matriz $LF(M,N,P,Q,ch)$.

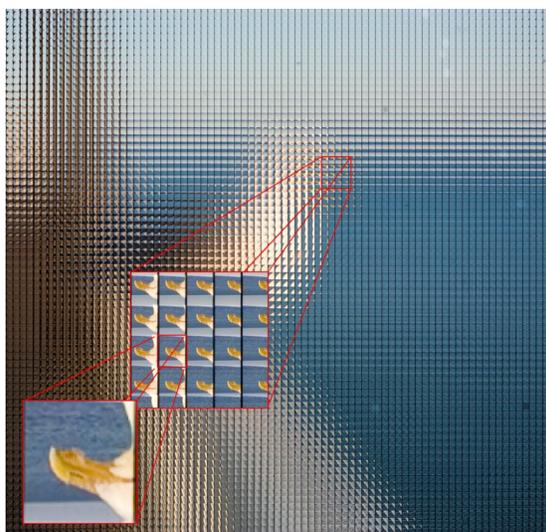


Figura 1. Fotografía Plenóptica obtenida por Georgiev et al. [11].

Para obtener una imagen en determinado ángulo, se selecciona una misma posición (P,Q) para todas las lenticulas y se genera una imagen de $M \times N$ pixeles. Recuérdese que la posición de los pixeles debajo de cada microlente guarda la información de la dirección de los rayos. En la figura 2 se observa una secuencia de imágenes en diferentes perspectivas, reconstruidas a partir de una única captura.

Filtro Desplazamiento-Suma

Uno de las estrategias más sencillas para obtener foco en diferentes profundidades es aplicando el Filtro Desplazamiento-Suma. El cómputo de refocalización consiste en desplazar cada imagen en perspectiva por un factor de $(P(1-1/\alpha), Q(1-1/\alpha))$ y sumarlas. Cuando $\alpha=1$ significa que no hay refocalización y el desplazamiento es nulo para el conjunto de las imágenes de perspectiva. En la figura 3 se observa una secuencia de imágenes enfocadas a diferentes profundidades, según el parámetro alfa.

Las imágenes refocalizadas presentan borrono, una manifestación del aliasing asociado con un resampleo inapropiado del campo de luz dado que en el cómputo de los desplazamientos los índices de posición calculados son reales y deben redondearse al entero más cercano.

Fast Discrete Focal Stack Transform

Otra aproximación a la refocalización digital es aplicar un caso particular de la Transformada de Radón y su extensión N -dimensional. En [12] se propone una Transformada para generar la pila focal, basada en la Transformada Discreta de Radon. Sus autores proponen el algoritmo FDFST (Fast Discrete Focal Stack Transform) para obtener un conjunto de imágenes enfocadas en diferentes profundidades. La figura 4 muestra el resultado de aplicar FDFST a una imagen. Se observa resultado similar al obtenido con el algoritmo de desplazamiento-suma, excepto por un nivel de brillo menor.

Filtro Planar 4D de dominio frecuencial

Otra alternativa de filtrado es trabajar en el dominio de Fourier. El aporte de Ng [13] fue la generalización del Teorema del Corte, que permite obtener imágenes refocalizadas al aplicar la transformada inversa de Fourier a un plano cortado del hiperespacio transformado del campo de

luz. Dicho plano se obtiene construyendo un filtro frecuencial H que toma como parámetro una pendiente (correspondiente a la inclinación del plano de corte). Asociada a dicha pendiente está la profundidad a la que se obtendrá el foco en la imagen resultado. Se multiplica punto a punto el espacio transformado por el filtro H y se aplica la transformada inversa de Fourier para generar la imagen mencionada. En la figura 5 se muestran algunas imágenes refocalizadas por corte del hiperespacio transformado.

Trabajo futuro

Actualmente se están implementando y evaluando otros algoritmos con el propósito de encontrar aquel que se ajuste a los requerimientos del Proyecto. Una vez seleccionado, deberá construirse la pila focal e implementarse el algoritmo para crear el mapa de profundidades del fondo ocular. La imagen del fondo visualizada como superficie 3D se espera que incremente el valor diagnóstico de los estudios retinográficos.

Formación de Recursos Humanos

El equipo de trabajo está conformado por especialistas del área de sistemas de información, inteligencia artificial y bioingeniería. Relacionadas con el Proyecto de Investigación se han generado dos Tesinas de Grado (una aprobada y la otra en curso) y una de Maestría (en curso).

Referencias

[1] Salvatelli A, Bizai G, Martina M, Hadad A and Drozdowicz B 2013 Detección temprana de patologías diabéticas oculares en retinografías utilizando un Sistema Multiagentes 42 *JAIHO - CAIS 2013* - ISSN: 1853-1881 p 263

[2] Salvatelli A, Bizai G, Barbosa G, Drozdowicz B and Delrieux C 2007 A comparative analysis of pre-processing techniques in color retinal images *Journal of Physics: Conference Series* **90** 012069. IOP Publishing. doi:10.1088/1742-6596/90/1/012069

[3] Insfran M, Yancovich J, Salvatelli A, Bizai G and Drozdowicz B 2015 Workflow based on pixel processing for automatic segmentation of vasculature in retinographies *IFMBE Proc* **49** ISBN/ ISSN 1680-0737

[4] Gelroth Balbi A, Rodriguez D, Salvatelli A, Drozdowicz B and Bizai G 2011 Analysis and implementation of methodologies for the monitoring of changes in eye fundus images *Journal of Physics: Conference Series* **332** 012036 IOP Publishing.

[5] Cunha-Vaz J 1998 Diabetic macular edema *European Journal of Ophthalm.* **8** n.3 pp. 127–130

[6] González G, Faisano A and Torres R 2015 The slit-lamp survey in Entre Ríos, Argentina *IFMBE Proc* **49**, pp 117–119.

[7] Gershun A 1939 The light field. Moscow, 1936. *Journal of Mathematics and Physics* **XVIII**, 51–151 Translated by P. Moon and G. Timoshenko.

[8] Adelson E and Bergen J 1991 The plenoptic function and the elements of early vision *Computational models of visual processing* **91** n.1 pp 3–20

[9] Levoy M and Hanrahan P 1996 Light field rendering *SIGGRAPH ACM* pp 31–42

[10] Adelson E and Wang J 1992 Single lens stereo with a Plenoptic Camera *IEEE transactions on pattern analysis and machine intelligence* **14** n.2

[11] Georgiev, T.: New results on the Plenoptic 2.0 camera. *Signals, Systems and Computers, Conference Record of the Forty-Third Asilomar Conference*. Pp 1243–1247 (2009)

[12] Pérez Nava F, Marichal-Hernández J and Rodríguez-Ramos J 2008 The discrete Focal Stack transform *Proceedings of Eusipco'08*, August 25–29 2008

[13] Ng R 2006 Digital light field photography *Ph.D. dissertation, Stanford University*



Figura 2: Dos perspectivas de la misma escena. En el círculo, detalle de una de las diferencias.



Figura 3: Resultados obtenidos del algoritmo Shift Sum para distintos valores de 'alfa'.

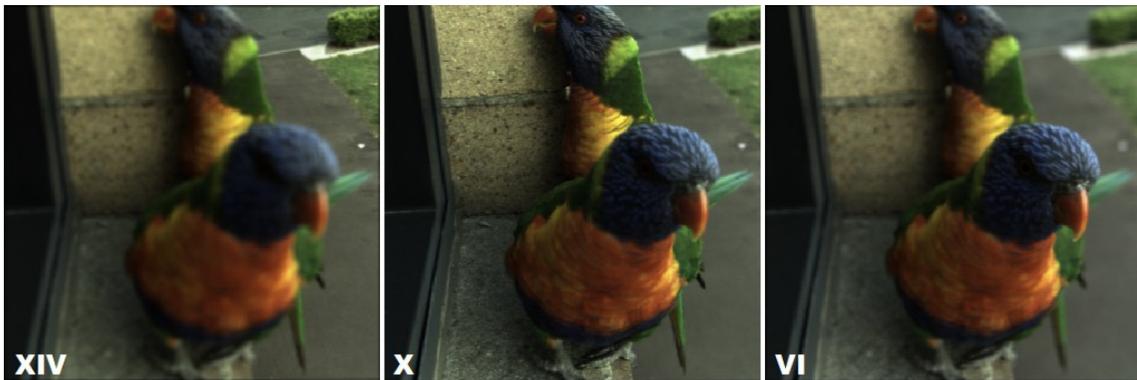


Figura 4: Tres de las 17 imágenes de una pila focal.

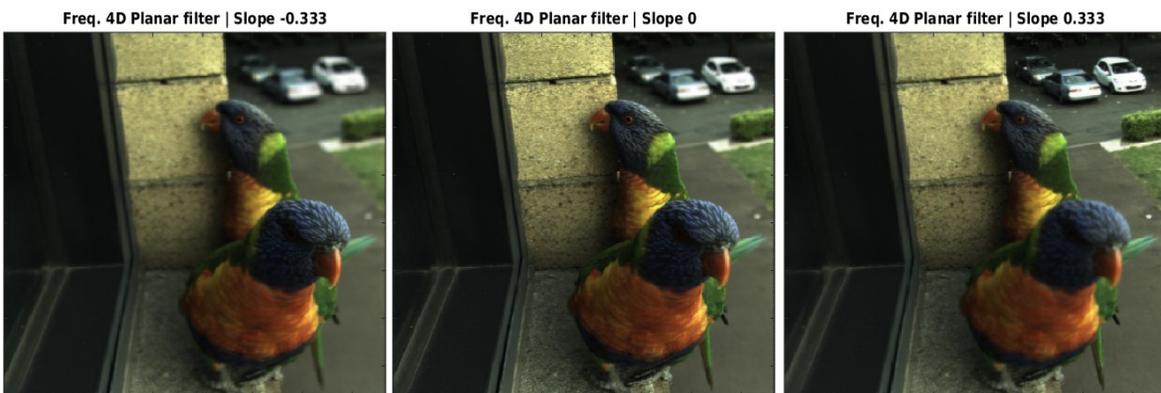


Figura 5: Filtro planar frecuencial 4D con diferentes pendientes.