

Modelo foveal rectangular a partir de un sensor de imagen comercial

José Antonio Loaiza Brito, Luis Niño de Rivera y Oyarzábal

Instituto Politécnico Nacional, Sección de Estudios de Posgrado e Investigación,
ESIME Culhuacan, México

jloaiza@ipn.mx, luisnino@rivera@gmail.com

Resumen. Este trabajo propone la implementación del modelo foveal tipo rectangular a la imagen generada en un sensor convencional. La principal problemática en los sensores foveales basados en el modelo log polar es que cada fotosensor está diseñado con geometrías y tamaños diferentes. Los sensores convencionales tienen una distribución de M por N píxeles en un arreglo rectangular, en tanto que en los modelos foveales, los elementos de imagen se encuentran distribuidos en círculos concéntricos. El diseño de cada sensor debe ser específico de acuerdo a la aplicación que se requiera. Esto ocasiona una limitación en el desarrollo de sistemas con distintas configuraciones de retinas, debido a que es necesario cambiar los diseños para diferentes aplicaciones con el respectivo análisis. La propuesta consiste en una nueva estrategia para generar una imagen foveal desde un sensor de imagen convencional, que utiliza un algoritmo de transformación log polar en un modelo rectangular, donde se promedian los valores de los elementos fotosensores vecinos en una regla de crecimiento geométrico. Los resultados muestran las imágenes obtenidas desde un sensor de imagen comercial, aplicando el algoritmo de transformación log polar a un esquema de rectángulos circunscritos con crecimiento geométrico.

Palabras clave: modelo foveal, modelo log polar, sensor de imagen, visión foveal.

Foveal Rectangular Model from a Commercial Image Sensor

Abstract. This work proposes the implementation of the rectangular model foveal to the image generated in a conventional sensor. The main problem in foveal sensors based on the log polar model is that each photosensor is designed with different geometries and sizes. Conventional sensors have a distribution of M by N pixels in a rectangular array, whereas in foveal models, the elements of image are distributed in concentric circles. The design of each sensor must be specific according to the application that is required. This causes a limitation in the development of systems with different configurations of retinas, because it is necessary to change the designs for different applications with the respective

analysis. The proposal consists of a new strategy to generate a foveal image from a conventional image sensor, which uses a log polar transformation algorithm in a rectangular model, where the values of the neighboring photo-sensor elements are averaged in a geometric growth rule. The results show the images obtained from a commercial image sensor, applying the log polar transformation algorithm to a scheme of circumscribed rectangles with geometric growth.

Keywords: foveal model, log polar model, image sensor, foveal vision.

1. Introducción

El modelo foveal está inspirado en los sistemas de visión de los animales debido a que se trata de emular el comportamiento de la retina biológica, basándose en la distribución geométrica de las células fotorreceptoras en humanos o animales [1,2,3]. En la retina biológica se tiene la mayoría de los fotorreceptores concentrados en la zona foveal y disminuyen en número en dirección a la periferia [4,5,6]. Esta organización de los fotorreceptores tiene propiedades significativas debido a la variación en la resolución en el campo visual. A esto se le conoce como *visión espacio-variante* [5,7]. Existen numerosos estudios en que han desarrollado sensores con estructuras foveales, destacando la tecnología CMOS debido a sus capacidades y calidad en resolución [6,7,8]. Sin embargo, el diseño y fabricación de este tipo de sensores aún tienen problemas que deben ser resueltos.

Aunque algunos autores han sugerido una aplicación ingeniosa en VLSI, éstas aún requieren un proceso de fabricación no estándar [7,8]. Pardo et al. presentaron el diseño de un sensor foveal CMOS utilizando ajustes en la geometría de cada fotosensor de acuerdo a la ubicación desde el centro hasta la periferia, utilizando una rotación de ejes, los cuales, van creciendo en tamaño en los diferentes niveles de la fovea para producir corrientes fotogeneradas proporcionales al área de incidencia [7,8,9]. Sin embargo, ésta se debe compensar debido al crecimiento de los sensores. Nuevas soluciones se encuentran en la literatura especializada que proponen la construcción de elementos de detección de tipo logarítmico para compensar dichos problemas [4,5,7,8].

F. Pardo, B. Dierickx, D. Schaffer, F Paillet y P. Zeferino analizaron en detalle el procedimiento de diseño de una celda básica para un fotodiodo con tecnología CMOS [7,8,9,10]. El Patrón de Ruido (FPN), debido al factor de escala en la fabricación y los efectos de canal angosto, es un gran reto que todavía no está resuelto. Pardo et al. proponen reducir el FNP generado por la escala de diseño, utilizando una estructura cartesiana de píxeles para la región de mayor resolución y posteriormente una transformación log polar en la región periférica. P. Pardo muestra una transformación de ejes cartesianos para rotar los polígonos [7]; sin embargo, al momento de realizar el diseño de todo el sensor, es necesario construir un modelo de anillos concéntricos con un factor de escalamiento geométrico. Es por ello que sería necesario rotar la célula básica y escalarla, (figura 1). El modelo foveal representa la base de los sistemas de visión foveal; sin embargo, el problema principal para poder llevar diversos análisis de estructuras foveales, representa el inconveniente de su fabricación.

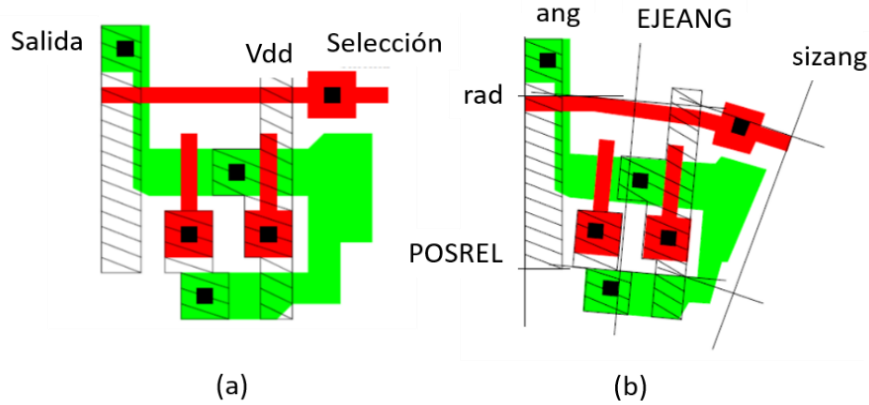


Fig. 1. Diseño de una celda básica de fotodetección. a) Celda original con sus líneas de acceso. b) Rotación establecida para el campo receptivo, en la que se muestran los ejes modificados [7].

2. El modelo foveal

El modelo foveal es una transformación matemática de una imagen desde un plano polar a un plano rectangular a través de un mapeo logarítmico y luego regresarla para su interpretación. La representación log-polar [7,8,11,12,13] es el mapeo de puntos del plano polar o también llamado plano retínico (ρ, η) a un plano complejo logarítmico, nombrado plano cortical (ξ, γ) .

De acuerdo con [11,14], la imagen foveal es, en resumen, una representación en un plano discreto de la transformación log-polar. Esto es, una imagen es representada en coordenadas rectangulares de acuerdo con:

$$z = x + jy, \quad (1)$$

$$z = \rho e^{j\eta}, \quad (2)$$

donde:

$$\rho = \sqrt{x^2 + y^2}, \quad (3)$$

$$\eta = \tan^{-1}\left(\frac{y}{x}\right). \quad (4)$$

Por lo que las coordenadas x , y se pueden expresar como:

$$x = \rho \cos(\eta), \quad (5)$$

$$y = \rho \sin(\eta). \quad (6)$$

Lo anterior representa la ubicación de un elemento de imagen en el plano rectangular. Cualquier pixel queda perfectamente ubicado bajo estas coordenadas. De la investigación de Volker Krüger, una imagen cortical quedaría representada a través de un mapeo al plano logarítmico complejo por la siguiente regla de correspondencia:

$$w = \ln(\rho e^{j\eta}). \quad (7)$$

El origen del plano x-y no se puede mapear por la singularidad presentada en $\rho=0$. Por lo que es necesario definir un intervalo muy cercano al origen de la imagen. La representación de la imagen tiene un crecimiento geométrico o exponencial, por lo que se utiliza un logaritmo arbitrario de base “a” [14]. Así, para el mismo punto de imagen z:

$$z = \rho a^{j\eta}. \quad (8)$$

La base “a” es la contraparte del *factor de magnificación cortical*. Entonces:

$$w = \log_a(\rho) + j\eta. \quad (9)$$

Para realizar el proceso de visión por computadora o utilizar un procesador digital, los componentes radiales y angulares deberán ser digitalizados. Por ello se consideran las siguientes coordenadas: ξ = Componente radial, γ = Componente angular.

Las variables N_r y N_a definen la resolución máxima radial y angular del plano log polar, respectivamente. El mapeo de las coordenadas polares (ρ, η) a las coordenadas (γ, ξ) en el plano log polar, está dada por [14]:

$$\xi = \log_a\left(\frac{\rho}{\rho_0}\right), \quad (10)$$

$$\gamma = \frac{N_a}{2\pi}\left(\eta + \frac{\pi}{2}\right). \quad (11)$$

La figura 2 muestra una representación de la transformación con las coordenadas de (10) y (11). El plano cortical aún sigue siendo un plano rectangular en donde queda representada la imagen en coordenadas logarítmicas. Para obtener la representación foveal de dicha imagen, es necesario obtener la transformación inversa, que es en realidad, la transformación log polar. Las coordenadas se definen como:

$$x = \rho_0 a^\xi \text{Cos}(\eta), \quad (12)$$

$$y = \rho_0 a^\xi \text{Sen}(\eta), \quad (13)$$

donde los valores de las coordenadas quedan en un intervalo de $0 \leq \xi \leq N_r, 0 \leq \gamma \leq N_a$. De igual manera, en [13,14], se establece lo siguiente:

$$\rho_0 a^{N_r} = \rho_{max}, \quad (14)$$

donde es posible observar que el tamaño de la imagen transformada queda relacionado por el factor “a”. La forma de la ecuación es una función de crecimiento geométrico.

3. Metodología

En este trabajo se discute la manera de utilizar la transformación log polar para generar un modelo rectangular e implementarlo desde un sensor de imagen comercial.

Así, cualquier elemento de fotodetección queda especificado por un pixel, llamado *celda básica*, (figura 2). La propuesta consiste en realizar agrupaciones de celdas básicas para hacer crecer zonas de incidencia hacia la periferia. La distribución en un plano polar (figura 2a.) queda determinada por medio de campos receptivos. Así, al colocar agrupaciones de celdas básicas, éstas pueden cubrir áreas completas si se realiza una distribución rectangular o cuadros concéntricos. Los campos receptivos se presentan como rectángulos con crecimiento geométrico.

Una primera evaluación consiste en superponer el conjunto de celdas rectangulares sobre la estructura polar, generando diferentes zonas blancas sobre el plano polar. Sin embargo, las zonas blancas decrecen cuando se trata de rectángulos concéntricos. Este nuevo enfoque reduce las zonas blancas y produce una mejor distribución sobre el diseño.

Las trayectorias rectas concéntricas en la figura 2-b se mueven gradualmente hacia la periferia siguiendo el modelo log polar conocido. Lo relevante es que la ubicación de cada celda básica sigue las ecuaciones (12) y (13) haciendo posible reorganizar los pixeles de una imagen en una nueva estructura foveal, pero rectangular. Esta agrupación respeta el modelo espacio-variante del sensor foveal estudiado. El ajuste se realiza utilizando asociaciones de elementos de imagen siguiendo una función polinomial:

$$f(k) = b^k, \quad (15)$$

donde “b” es la base en la que se trabaja la serie y “k” equivale al número de capas o niveles en los que dividirá la imagen foveal.

En esta propuesta, se establece la base 2, debido a que las asociaciones serán en grupos de dos, cuatro o más celdas básicas, como se puede observar en la figura 2-b.

Este nuevo modelo es equivalente al modelo log polar convencional, ya que las trayectorias se extienden por la misma ley de crecimiento geométrico de los círculos de la estructura espacio-variante. Es importante mencionar que la ley de crecimiento de las celdas que componen cada rectángulo circunscrito en la nueva distribución foveal está controlada por el Factor de Crecimiento S.

De acuerdo con (14) y (15), se propone que, el elemento de imagen más pequeño se designa como W_0 , el cual tiene un crecimiento geométrico rectangular con un factor de crecimiento S. El tamaño del i-ésimo elemento de imagen está dado por:

$$W_i = S^i W_0. \quad (16)$$

Por lo que el tamaño semi radial del sensor será:

$$rd = \sum_{i=0}^N W_i, \quad (17)$$

$$rd = W_0 \left(\frac{S^{N+1}-1}{S-1} \right), \quad (18)$$

donde N es el número de capas del modelo foveal. Sin embargo, si se toma como centro el elemento mínimo de imagen, se presentará una singularidad (cuando S=1). Es por ello, que se deberá tomar lo siguiente:

$$rd_{max} = rd_0 + W_0 \left(\frac{S^N+S-2}{S-1} \right), \quad (19)$$

rd_{max} queda determinado por la suma del radio del elemento mínimo de imagen ubicado en el centro más el radio (rd) de todo el sensor. Cada elemento rectangular de tamaño W_i queda representado en el plano rectangular, como se observa en la figura 3.

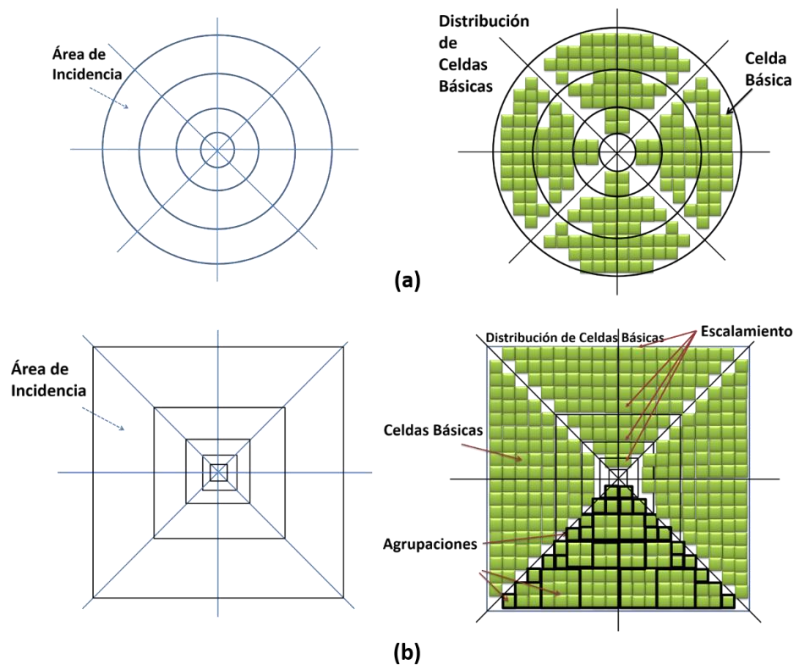


Fig. 2. Correspondencia entre un modelo de círculos concéntricos y uno rectangular. En (a) se muestra la distribución de celdas básicas en los diferentes anillos. En (b) se puede observar una propuesta de distribución de celdas en un formato rectangular basado en agrupaciones de acuerdo al área de incidencia.

La magnitud de W_0 tiene valores reales positivos que dependen de la tecnología de fabricación para el fotodetector o simplemente adquieren valores enteros positivos cuando se trata de una imagen en píxeles. El crecimiento geométrico de los elementos de imagen, sirve de base para la transformación de la imagen original en un modelo polar, visto desde un plano rectangular.

Una vez definida la celda básica, es posible generar toda la retina y la fóvea en una matriz concéntrica rectangular distribuyendo los píxeles. Los fotorreceptores están representados con cuadros distribuidos en polígonos concéntricos alrededor del centro del sensor cuya separación aumenta de forma exponencial. Lo anterior posibilita que cualquier circuito de fotodetección sea una celda básica, evitando una rotación de ejes en el diseño de la estructura.

La figura 4 representa la forma de agrupar los elementos de imagen para llevar a cabo la distribución foveal de la imagen. La selección de filas y columnas es necesaria al tratarse de un arreglo matricial; sin embargo, la línea de selección de “anillo” es fundamental

para seleccionar un contorno cuadrado similar a los anillos o circunferencias en una representación polar.

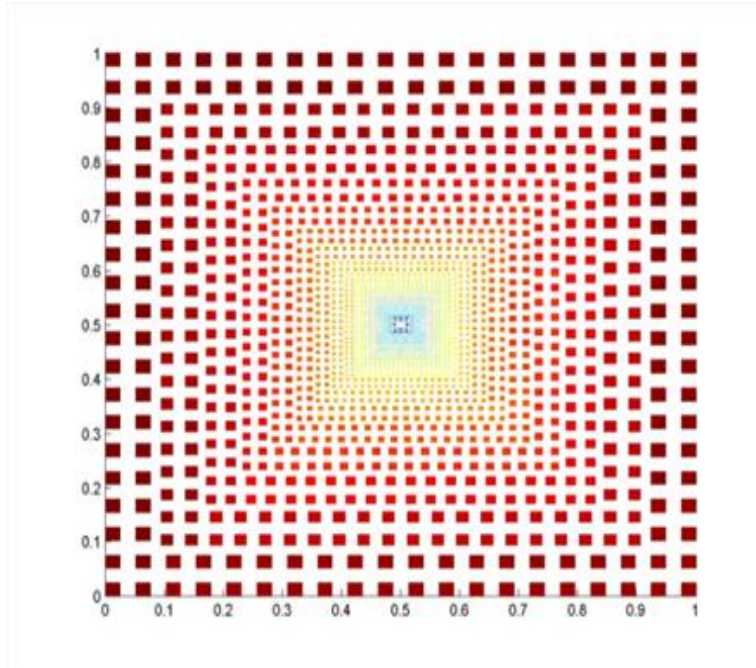


Fig. 3. Representación de la transformación por agrupación de celdas utilizando el modelo rectangular.

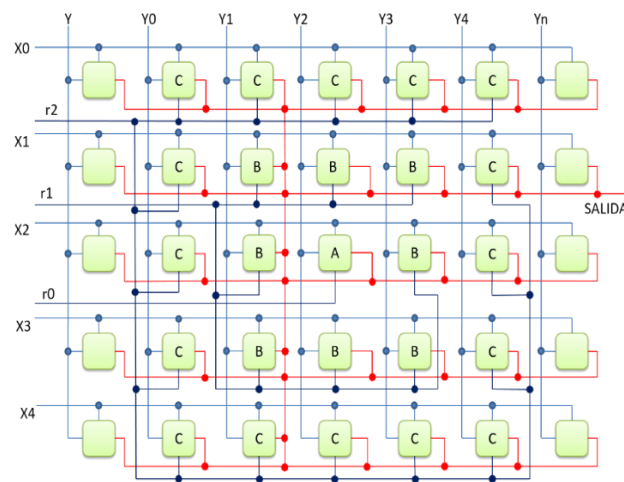


Fig. 4. Estructura del modelo foveal rectangular donde las coordenadas X, Y son para seleccionar cada pixel. La variable r representa la manera con la cual se agrupan los elementos de imagen.

El modelo conecta celdas básicas formando un contorno. Por ejemplo, la celda básica central es individual y se selecciona mediante las líneas (r_0, X_2, Y_2) . El contorno siguiente puede seleccionarse mediante $(r_1, X_1-X_2-X_3, Y_1-Y_2-Y_3)$, lo anterior significa que se pueden generar las siguientes coordenadas para seleccionar cada una de las celdas (r_1, X_1, Y_1) , (r_1, X_2, Y_1) , (r_1, X_3, Y_1) , etc. Para el siguiente contorno, se realiza de la misma manera $(r_2, X_0-X_1-X_2-X_3-X_4, Y_0-Y_1-Y_2-Y_3-Y_4)$.

4. Resultados

El modelo propuesto presenta que es posible realizar agrupaciones de celdas con el objeto de tener áreas de incidencia mayores sin modificar el tamaño del pixel. La salida será la suma de las corrientes fotogeneradas o del valor promedio de los pixeles en las celdas agrupadas. De acuerdo con (18) y (19) y considerando el elemento mínimo igual 1 y 480 como el máximo, se obtienen los valores correspondientes al Factor de Crecimiento S, (figura 5).

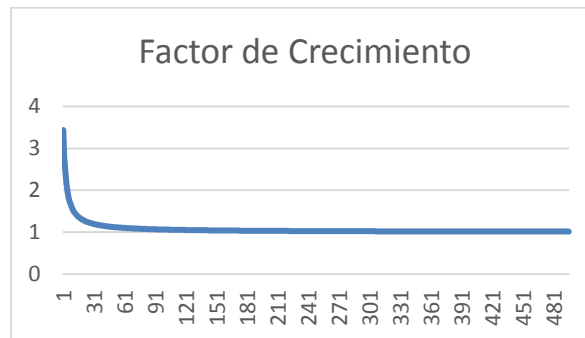


Fig. 5. Gráfica del número de capas (resolución en pixeles) contra el factor de crecimiento (escalamiento).

Al aproximarse a los 480 pixeles el factor de crecimiento varía muy poco. Aplicando el procedimiento a las imágenes de prueba es posible obtener una imagen foveal como se muestra en la figura 6.

El valor de $N=240$ se toma como un valor máximo de acuerdo a la resolución de la imagen original que es de 480 pixeles. Las figuras en 6(a) son imágenes tomadas por un sensor de imagen comercial OV7670 con una resolución de 640×480 pixeles. Desarrollando el procedimiento de transformación foveal rectangular (figura 3 y 4), es posible observar que se mantiene una alta resolución en el centro de la imagen y que decrece hacia la periferia.

El cuadro blanco en el centro de la imagen se debe a que se consideró como el elemento más pequeño a un grupo de pixeles para conformar la fovea; sin embargo, eso no es necesario ya que es completamente viable considerar a un pixel como el elemento

básico de imagen. Lo anterior se puede observar en la primera figura foveal, en la que el cubo tiene un intervalo de área central de píxeles.

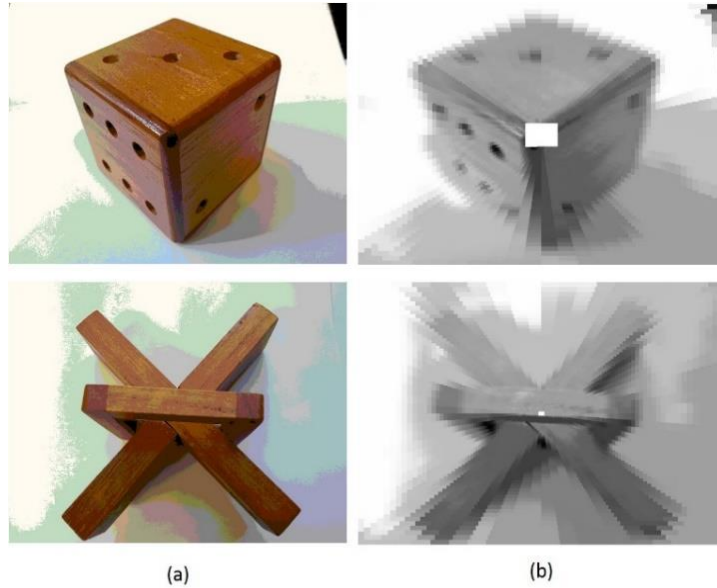


Fig. 6. Transformación Foveal Rectangular. a) Imágenes originales tomadas del sensor comercial. b) Imágenes que han sido transformadas siguiendo la estructura de la figura 4.

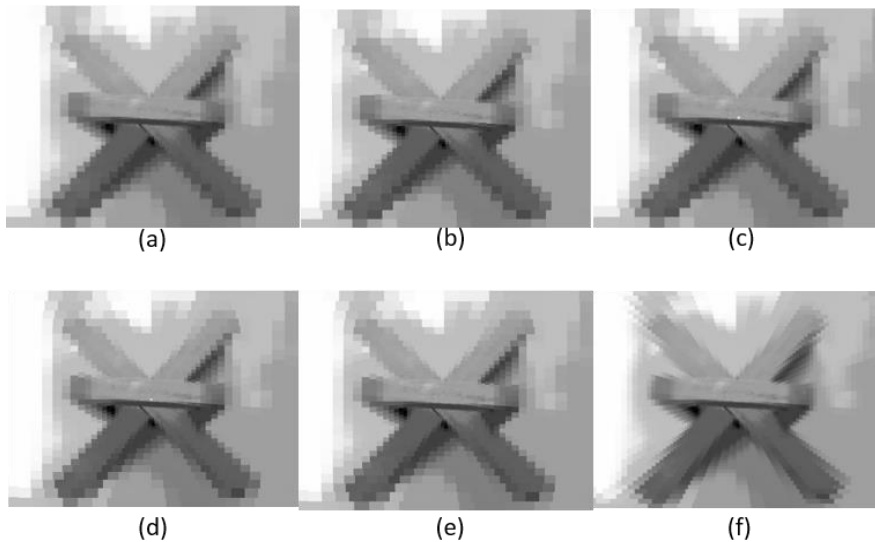


Fig. 7. Conversión foveal rectangular con una imagen tomada desde el sensor OV7670 con una resolución de 640x480 píxeles. (a) $S=10$, $N=120$. (b) $S=10$, $N=240$. (c) $S=50$, $N=120$. (d) $S=50$, $N=240$. (e) $S=100$, $N=120$. (f) $S=100$, $N=240$.

El valor se toma considerando un píxel a la derecha e izquierda de un punto central. De igual manera, un píxel superior e inferior. Debido al escalamiento, esto se ve reflejado en un cuadro blando notorio en el centro de la imagen. Para la imagen inferior, solo se tomó un píxel como intervalo central. Nótese que el área blanca se encuentra reducida.

Las imágenes de la figura 7, muestran el efecto con diferentes valores de S y un valor constante del número de capas en 120 y 240.

5. Conclusiones

Los resultados muestran que la estructura propuesta basada en fotosensores rectangulares funciona correctamente para implementar sensores de visión foveal a partir de estructuras de sensores convencionales. Este nuevo enfoque de los sensores foveales preserva las propiedades de la transformada log polar. Además, es posible trabajar con diseños VLSI convencionales y sin requisitos previos de fabricación de sensores estándar. Por el contrario, los chips foveales (que son de fabricación no estándar) son una de las razones que han dificultado la aplicación en la visión foveal. Este modelo puede aplicarse a nuevos chips foveales dedicados basados en fotodiodos, fototransductores CMOS convencionales o fototransistores.

La celda básica puede ser cualquiera que haya sido empleada en diferentes modelos retinales, estableciendo un modelo foveal ortogonal con cuadros concéntricos y agrupación de elementos de imagen del mismo tipo sin realizar ninguna transformación de ejes. El punto clave es asegurar que esas estructuras sean capaces de tener un acceso independiente a cada píxel a través de una lógica de control apropiada.

La transformación log polar a través de software dedicado ha visto una aplicación limitada debido a los tiempos de transformación y procesamiento impuestos por la lógica de programación. El modo de transformación propuesto en este artículo, que utiliza un espacio variable rectangular, permite realizar la transformación foveal directamente sin llevar a cabo una transformación previa cortical y así obtener de forma directa una reducción en la cantidad de información que podrá ser procesada posteriormente.

Referencias

1. Palanker, D., Vankov, A., Huie, P., Baccus, S.: Design of a high resolution optoelectronic retinal prosthesis. Department of Ophthalmology and Hansen Experimental Physics Laboratory, Stanford University, Institute of Physics Publishing 2 S105–S120 (2002)
2. Weber, C., Trisch, J.: Implementations and Implications of Foveated Vision. *Recent Patents on Computer Science* 2, 75–85 (2009)
3. Bolduc, L.: A review of biologically motivated space variant data reduction models for robotic vision. *Computer Vision and Image Understanding*, vol. 69, 170–184 (February 1998)
4. Traver, V. J., Pla, F.: Designing the lattice for log-polar images. *Discrete Geometry for computer Imagery, Proceedings*, vol. 2886, 164–173 (2003)
5. Melnyk, P.: Biologically Inspired Composite Image Sensor for Deep Field Target Tracking. Thesis. University of New Hampshire.

6. Niño-de-Rivera, L., Calzada Salas, R., Duchén Sánchez, G., Loaiza Brito, J.A.: Visual Simulation of Retinal Images Through Microstructures. *Microelectronic Engineering*. Vol. 90, pp. 159–162 (2012)
7. Pardo, F., Dierickx, B., Scheffer, D.: Space-Variant Nonorthogonal Structure CMOS Image Sensor Design. *IEEE Journal of Solid-State Circuits*, Vol. 33, No. 6 (June 1998)
8. Pardo, F., Dierickx, B., Scheffer, D.: CMOS foveated image sensor: Signal scaling and small geometry effects. *IEEE Transactions on Electron Devices*, Vol. 44, No. 10 (October 1997)
9. Paillet-Damien, F., Thierry, M., Bernard, M.: Second Generation Programmable Artificial Retina. CTA/GIP ENSTA/LEI. 32. 75015 Paris, France (2000)
10. Zeferino-Tomas, P. F.: Bio-Inspired Processing Module for the development of an artificial retina. Universidade Tecnica de Lisboa. Instituto Superior Tecnico. Licenciatura em Engenharia Electrotecnica e de Computadores. Relatório de Trabalho Final de Curso 157/2002/M (Setembro de 2003)
11. Pardo-Carpio, F.: Sensor Retínico Espacio Variante basado en Tecnología CMOS. Tesis. Departamento de Informática y Electrónica Universidad de Valencia (24 de septiembre de 1997)
12. Christos-Savvas, D. B.: Reconfigurable Foveated Active Vision System. In: 3rd International Conference on Sensing Technology, Taiwan (2008)
13. Javier-Traver, V., Bernardino, A.: A review of log-polar imaging for visual perceptions in robotics. *Robotics and Autonomous Systems* 58 (2010)
14. Volker, K.: Optical Flow Estimation in the Complex Logarithmic Plane. Instituts für Informatik und Praktische Mathematik der Christian Albrechts Universität zu Kiel (1995)