

# Estudio comparativo de modelos de conversión de imágenes log-polar\*

Joaquin López Sánchez, V. Javier Traver, Filiberto Pla  
jlopez@linx.act.uji.es, {vtraver|pla}@uji.es  
*Grup de Visió per Computador · Universitat Jaume I  
Campus Riu Sec · 12080 – Castelló, Spain*

**Resumen:** *La información capturada por el sistema visual humano sufre una transformación hasta llegar al córtex visual, a través del nervio óptico. Estudios neurofisiológicos han demostrado que dicha transformación puede modelarse mediante una expresión logarítmico-compleja (más conocida como log-polar). Este modelo biológico presenta propiedades que resultan muy interesantes en ámbitos de la visión artificial, tales como la visión activa y el reconocimiento de formas. En este trabajo presentamos diversos modelos matemáticos de transformación log-polar que se han propuesto en la literatura. Además de analizar las propiedades que presenta cada uno de estos modelos, realizamos un estudio comparativo atendiendo a diferentes criterios. Esta comparación puede ser útil para ayudar a seleccionar un determinado modelo en función de la tarea para la que se va a utilizar.*

**Palabras clave:** visión artificial, transformación log-polar, fóvea, retina, visión activa, reconocimiento de formas.

## 1 Introducción

Estudios neurofisiológicos han demostrado que la información visual en determinados animales superiores sufre una transformación retino-cortical (de la retina al córtex visual), y que dicha transformación viene dada por una sencilla ley matemática, conocida como transformación log-polar [11]. Este modelo biológico presenta interesantes propiedades que lo hacen idóneo en aplicaciones de visión activa y reconocimiento de formas.

Por ejemplo, para poder cumplir con los requerimientos de tiempo real, típicos en aplicaciones de visión activa, resulta interesante explotar la reducción selectiva de información que presentan las imágenes log-polares. Esta drástica reducción de información

---

\*Trabajo parcialmente financiado por los proyectos GV97-TI-05-27 de la *Conselleria d'Educació, Cultura i Ciència, Generalitat Valenciana*, y CICYT TIC98-0677-C02-01 del *Ministerio de Educación y Cultura*.

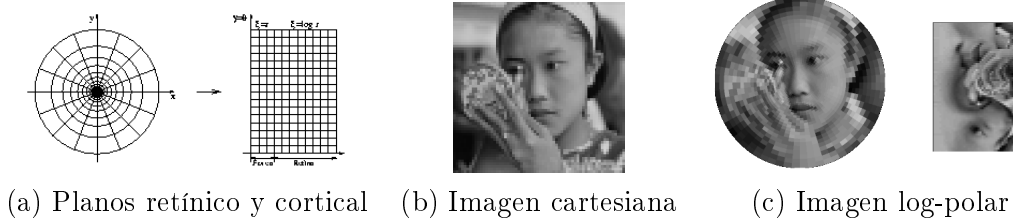


Figura 1. Planos retínicos y cortical. Ejemplo de transformación log-polar.

es posible por presentar una resolución variable que es máxima en el centro de la imagen (la fovea) y decrece con la distancia al centro. Además, en el contexto de la visión activa, la geometría log-polar presenta otras ventajas: simplifica el cálculo del tiempo al impacto [14], facilita el control de vergencia [5], presta ayuda en la navegación autónoma de robots móviles [2], etc. Por otro lado, sus conocidas propiedades de invarianza a rotaciones y escalados resultan útiles en reconocimiento de formas [13, 7].

Matemáticamente, la representación log-polar es la transformación de puntos del plano polar o plano retínico  $(\rho, \theta)$ , al plano plano cortical  $(\xi, \eta)$ . Un punto del plano retínico puede escribirse como un número complejo  $z = \rho \exp(i\theta)$ , donde  $\rho$  es la distancia al origen de coordenadas,  $\theta$  es el ángulo con el eje  $x$ , e  $i = \sqrt{-1}$ . De esta forma, la transformación al plano cortical se puede escribir como  $w = \log(z) = \log \rho + i\theta$ . Es decir, las coordenadas corticales  $(\xi, \eta)$  son:

$$\begin{cases} \xi = \log \rho \\ \eta = \theta \end{cases} \quad (1)$$

En la figura 1(a) se puede observar el plano retínico y el cortical de una transformación log-polar. También se muestra un ejemplo de transformación: la imagen cartesiana (figura 1(b)) y su conversión a log-polar (figura 1(c)).

Aparte de la reducción selectiva de información, hay dos ventajas matemáticas directas que facilitan el tratamiento de rotaciones y escalados respecto al origen. Una rotación respecto al centro de la imagen en el plano retínico se transforma en una traslación respecto al eje  $\eta$  del plano cortical, mientras que una escalado, también centrado en el origen, en el plano retínico se transforma en una traslación respecto al eje  $\xi$  del plano cortical, como se ilustra en la figura 2.

## 2 Transformación log-polar por *software*

### 2.1 Modelos log-polar

Presentamos aquí cuatro modelos de conversión log-polar. Aunque se han desarrollado cámaras foveales [9, 10] o transformaciones por *hardware* [6], nos centramos aquí en modelos que se utilizan normalmente en implementaciones de conversores por *software*. El trabajo de Bolduc y Levine [4], que presenta una revisión de algunos modelos, no

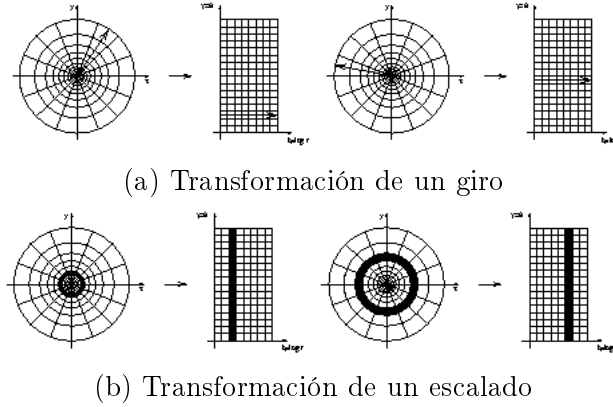


Figura 2. Rotaciones y escalados en el plano retínico se convierten en simples traslaciones en el plano cortical.

incluye el más recientemente propuesto [8]. La principal diferencia entre las distintas propuestas reside, básicamente, en la forma en que tratan la singularidad que presenta la función logaritmo en el origen.

En la conversión por software, una imagen cartesiana de dimensión  $M \times N$  se convierte en una imagen log-polar de dimensión  $R \times S$ , donde  $R$  denota el número de anillos y  $S$  el número de sectores. La imagen log-polar también se conoce como imagen cortical. La transformación inversa permite “reconstruir” la información contenida en el plano cortical utilizando el espacio retínico. Esta transformación log-polar inversa resulta útil a efectos de visualización. La figura 3 muestra un ejemplo de transformación de cada uno de los modelos considerados en este trabajo. Todos estos modelos tienen en común los parámetros  $R$ ,  $S$ , y  $\rho_{\max}$ , siendo este último el radio del anillo más externo (o el radio del campo visual). Además cada modelo tiene sus parámetros particulares, como veremos a continuación.

## Modelo de Sandini

Este modelo, utilizado, entre otros, por el grupo de Sandini [5], la transformación log-polar viene dada por

$$\begin{cases} \xi = \log_a \left( \frac{\rho}{\rho_0} \right) \\ \eta = \frac{S}{2\pi} \left( \theta + \frac{\pi}{2} \right) \end{cases} \quad (2)$$

Los parámetros específicos de esta transformación son  $a$  y  $\rho_0$ . La base del logaritmo,  $a$ , gobierna la tasa con que va decreciendo la resolución del sensor al alejarnos del centro. Con esta definición, al anillo más interno tiene asociado un radio  $\rho_0 > 0$ , de modo que la singularidad en el origen se soluciona mediante una pequeña zona “ciega” (ver figura 3(a)).

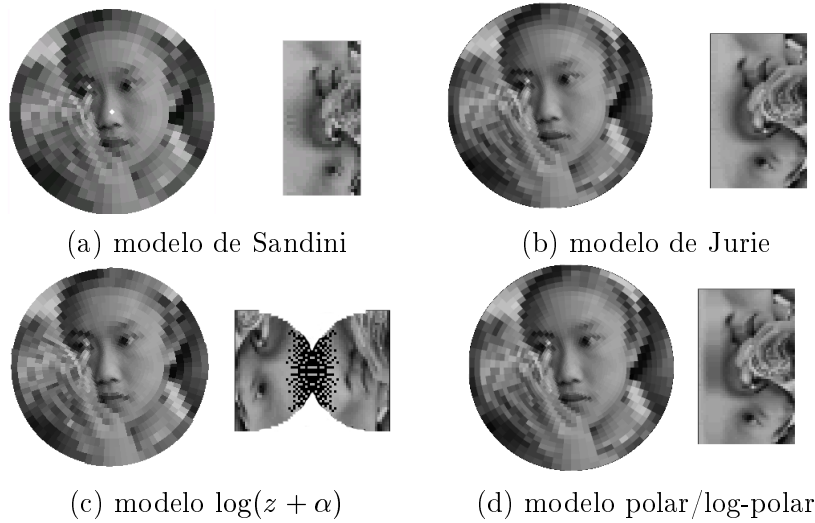


Figura 3. Ejemplo de transformación log-polar para los cuatro modelos considerados. En cada caso, a la derecha se muestra la imagen cortical y a la izquierda la imagen retinica que resulta de la transformación inversa.

## Modelo de Jurie

En este caso, las coordenadas en el espacio cortical vienen dadas por [8]:

$$\begin{cases} \xi = \log_a \left( \frac{\rho + \rho_0}{\rho_0} \right) \\ \eta = \theta \end{cases} \quad (3)$$

Los parámetros específicos de este modelo son  $a$  y  $\rho_1$ , el radio del primer anillo. Aunque pueden plantearse otras posibilidades, Jurie propone calcular  $a$  y  $\rho_0$  a partir de  $\rho_1$  y  $\rho_{\max}$ . También en este caso se soluciona la singularidad en el origen, pero sin que haya una área “ciega” (ver figura 3(b)).

## Modelo $\log(z + \alpha)$

En el modelo de Wallace *et al.* [16], un punto en el espacio cartesiano se considera como un número complejo  $z = x + iy$ , y la transformación consiste en aplicar el logaritmo de  $z$ :

$$w = \log(z + \alpha) \quad (4)$$

Como vemos, la singularidad en el origen se evita al sumar una cantidad  $\alpha > 0$  a la parte real de  $z$ , es decir, a la coordenada  $x$ . Esto equivale a dejar fuera del dominio de la función logarítmica una franja vertical centrada en el origen. Este modelo es el más fiel al modelo biológico. De hecho, incluso su representación (ver figura 3(c)) nos hace recordar los dos hemisferios cerebrales.

Para su implementación, Wallace *et al.* utilizan dos tablas de búsqueda,  $LUT_R(i, j)$  y  $LUT_S(i, j)$ , que contienen, respectivamente, el anillo y el sector al que se mapea cada

pixel  $(i, j)$ . Además, aprovechan la simetría que existe en los cuatro cuadrantes para calcular la transformación solamente en uno de ellos, y obtener el resto a partir de éste.

## Modelo polar/log-polar

Una forma obvia de evitar la singularidad de la función logarítmica en el origen es no usar ninguna función logarítmica. Por otro lado, el carácter logarítmico de la coordenada log-polar radial dificulta cierto tipo de procesamientos en la imagen si se requiere linealidad. Además, cerca del origen los píxeles de la imagen log-polar pueden ser incluso más pequeños que los píxeles de la imagen cartesiana, dando lugar a un submuestreo innecesario de la información. Por estos motivos, puede plantearse definir una función simplemente polar en la fóvea y log-polar en la retina:

$$\xi(\rho) = \begin{cases} k\rho & \text{si } \rho \leq \rho_T \\ \log_a\left(\frac{\rho+\rho_0}{\rho_0}\right) & \text{si } \rho > \rho_T \end{cases} \quad (5)$$

Imponiendo una condición de continuidad en la distancia de cambio  $\rho_T$ , tenemos que  $k\rho_T = \log_a\left(\frac{\rho_T+\rho_0}{\rho_0}\right)$ , por lo que  $k = \frac{\log_a\left(\frac{\rho_T+\rho_0}{\rho_0}\right)}{\rho_T}$ . Aunque aquí fijamos  $\rho_T$ , otra posibilidad sería encontrar un valor  $\rho_T$  apropiado, de forma automática [1].

## 2.2 Obtención de los niveles de gris

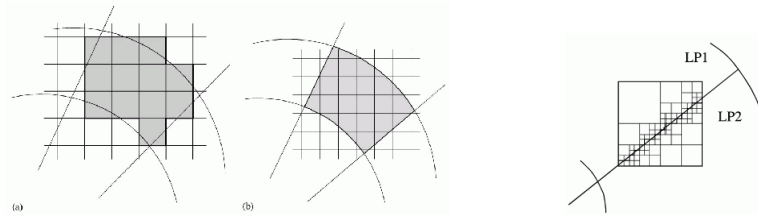
Además de la definición matemática de la función log-polar continua, existe la cuestión de la obtención de los niveles de gris de la imagen log-polar discreta a partir de la imagen cartesiana. Evidentemente este problema no se da en las cámaras foveales, puesto que los fotorreceptores son sensores reales. Al respecto existen diferentes soluciones, como utilizar campos receptivos (los píxeles de la imagen log-polar, o LP-píxeles en adelante) que se solapan entre sí. De esta manera, un mismo píxel de una imagen cartesiana (C-píxel en adelante) puede participar en varios LP-píxeles. Otras propuestas [4] utilizan funciones Gaussianas centradas en el LP-píxel para ponderar la contribución de los diferentes C-píxeles.

En [8] se propone usar LP-píxeles que son fracciones de C-píxeles, como se ilustra en la figura 4. De esta manera, un C-píxel se subdivide recursivamente hasta un cierto nivel de precisión prefijado.

## 3 Comparación de los modelos

Comparamos a continuación los diferentes modelos según diferentes criterios.

**Representación de la imagen.** En general la imagen log-polar es representable como cualquier imagen cartesiana convencional. Simplemente cambia el significado de los ejes (pasan a ser  $\xi$  y  $\eta$ ) y hay que tener en cuenta aspectos como que cada píxel representa un área física diferente. Por eso, además de la imagen, hay que guardar, dependiendo del



(a) LP-píxels son fracciones de C-píxels (b) Subdivisión recursiva

Figura 4. Formación de LP-píxels a partir de C-píxels

tipo de procesamiento que se vaya a realizar, el área que representa ese píxel. En todos los modelos presentados, la representación es básicamente la misma. Los autores del modelo  $\log(z + a)$ , sin embargo, proponen una representación, el *grafo de conectividad* [16], que permita representar cualquier tipo de geometría de sensor, independientemente de la complejidad de las relaciones de adyacencia. Esta representación facilitaría el procesamiento de imágenes espacio variantes de complejidad arbitraria.

**Invarianza a escalados.** Todos los modelos cumplen la relación de invarianza a rotaciones, por su carácter polar. La situación es distinta respecto a la invarianza a escalados. Básicamente, esta propiedad se basa en que el logaritmo de un producto es una suma de logaritmos. El modelo de Sandini cumple perfectamente la propiedad. De forma estricta, el modelo de Wallace *et al.* ( $\log(z + \alpha)$ ), no cumple esta propiedad. En la práctica, en cambio, se puede considerar que sí la cumple para  $|z| \gg \alpha$ . Sin embargo, el modelo de Jurie, al utilizar una suma dentro de la función logarítmica, impide que se cumpla esta propiedad. En cuanto al modelo polar/log-polar, sólo se daría la invarianza al escalado en la zona retínica, cuando es la fóvea la zona de interés en muchas aplicaciones.

**Coste espacial y temporal.** En general, el coste espacial es el coste de almacenar el mapa de superficies y el mapeado log-polar propiamente dicho. El de superficies depende del número de sectores y anillos con que se cree el mapeado y el del mapeado log-polar depende del número de columnas y filas de la imagen original. El coste tiende en todos ellos al orden  $O(M \cdot N)$ .

En cuanto al coste temporal, todos los algoritmos son  $O(M \cdot N)$ , por los cálculos necesarios para calcular el mapeado, aunque en realidad, el método  $\log(z + \alpha)$  tiene un coste menor, debido a dos razones. Por un lado, explota la simetría entre los cuadrantes, como se ha comentado anteriormente. Por otro lado, este método no realiza la subdivisión de píxeles, lo que supone una importante ganancia temporal. Sin embargo, en este sentido, la eficiencia puede conseguirse también en el esquema más general, con precisión subpixel, sin más que disminuir la precisión deseada. Estos costes están asociados sólo a la creación del mapeado. Es decir, para una misma geometría de las imágenes cartesiana y log-polar, estos cálculos se pueden calcular una sola vez, *off-line*, y después se puede reutilizar para realizar las transformaciones, suponiendo un ahorro de recursos computacionales.



Figura 5. Comparación de la relación calidad/tamaño en los modelos (a,e) Sandini, (b,f) Jurie, (c,g)  $\log(z + \alpha)$ , y (d,h) polar/log-polar, a resoluciones (a,b,c,d)  $32 \times 32$  y (e,f,g,h)  $64 \times 64$ .

**Relación calidad/tamaño.** Se pretende estudiar qué método ofrece mayor calidad visual con un menor tamaño de anillos y sectores. Aunque podría cuantificarse de algún modo, de forma subjetiva se puede observar en la figura 5 que el mapeado que obtiene una mayor calidad con una menor resolución es el modelo de Jurie. El modelo  $\log(z + \alpha)$  es el que presenta una menor resolución periférica. El tamaño de la zona “ciega” central afecta a la visualización del modelo de Sandini, tanto más cuanto mayor sea esta zona.

**Evolución de  $\xi(\rho)$ .** Durante una tarea de seguimiento activo, el objeto de interés estará centrado en la imagen. El tamaño del objeto puede representar sólo un reducido porcentaje del tamaño físico del plano imagen. Sin embargo, al trabajar en visión log-polar, el objeto representará, en términos de pixels de la imagen log-polar, un mayor porcentaje. Este hecho es importante al utilizar métodos de correlación de imágenes, puesto que el fondo de la imagen afecta poco las medidas de correlación. De forma implícita a la geometría log-polar, el objeto de interés, al estar centrado, cobra mayor importancia que el fondo. Una ventaja que se deriva de ello es que no se necesita realizar, en la mayoría de los procesos, ninguna segmentación del objeto. Algunos autores han mostrado el interés de este aspecto para visión binocular [3] o monocular [15]. De este modo, sería interesante estudiar la relación entre el tamaño de un objeto y lo que ocupa en la imagen log-polar, y comparar esta relación en distintos métodos de conversión log-polar. El tamaño del objeto podemos representarlo por la distancia al centro  $\rho$ , y el tamaño en la imagen por el número de anillos que ocuparía este objeto.

En la figura 6(a) se muestra pues las curvas  $\xi(\rho)$  para los distintos modelos. Se puede apreciar que el comportamiento es muy similar en todos los casos, si bien el modelo  $\log(z + \alpha)$  presenta un crecimiento más pronunciado en la parte central. La figura 5(c) confirma, visualmente, esta característica.

Puede ser también interesante estudiar cómo cambia el comportamiento de estas gráfi-

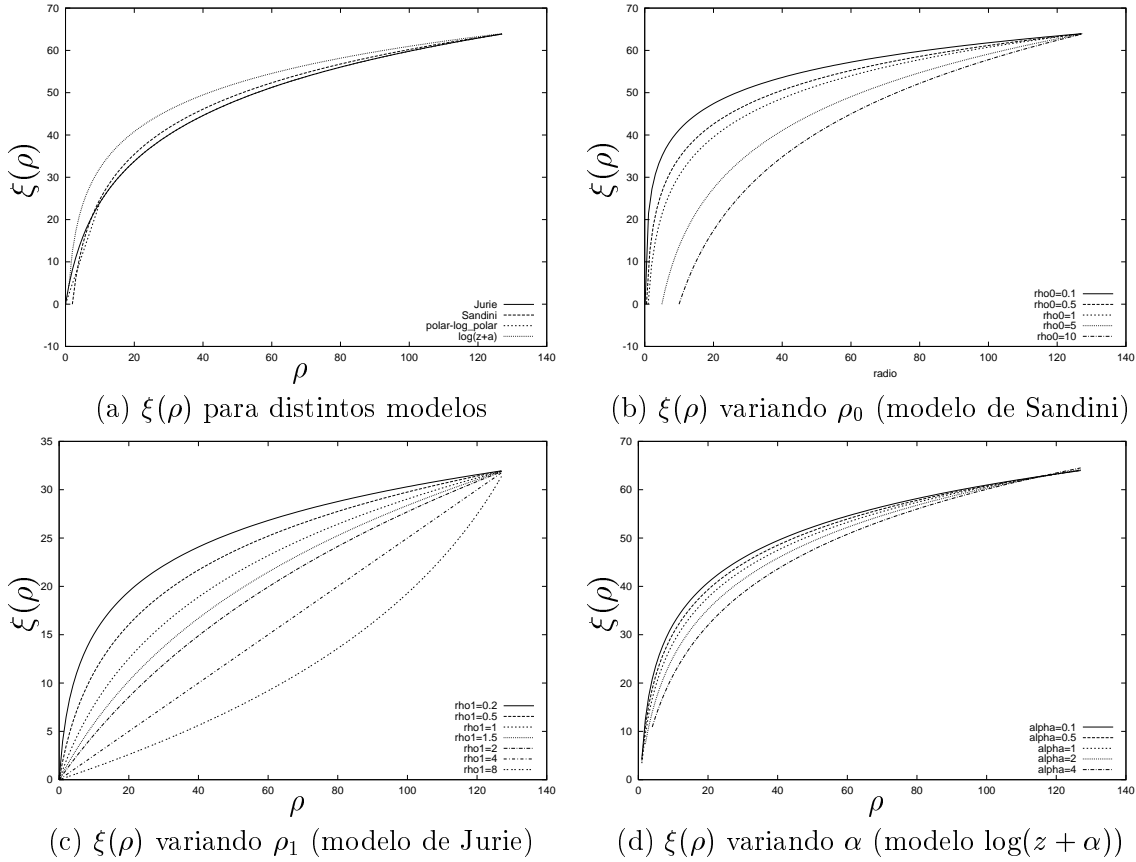


Figura 6. Comparación de  $\xi(\rho)$  (a) para distintos modelos y (b,c,d) variando un parámetro en cada modelo.

cas variando algún parámetro de cada modelo. En figura 6(b) se realiza una comparación de la evolución de  $\xi(\rho)$  para diferentes valores de  $\rho_0$  del modelo de Sandini. A medida que vamos aumentando el valor de  $\rho_0$  tenemos, en el centro de la fovea, una mayor área “ciega”, y el crecimiento de  $\rho$  es más suave. Es decir, a medida que aumenta  $\rho_0$  la resolución a lo largo de todo el rango de  $\rho$  se va haciendo cada vez más lineal.

En la figura 6(c) se realiza una comparación de la evolución de  $\xi(\rho)$  para diferentes valores de  $\rho_1$  del modelo de Jurie. A medida que va aumentando el valor de  $\rho_1$  se observa que la resolución de la imagen log-polar va pasando de ser mayor en el centro de la imagen a ser mayor en el exterior. En principio esto tiene poco sentido, porque nos interesa centrarnos en la zona de la fovea. Pero recientemente algunos autores [12] encontraban interesante aumentar la resolución en la retina (hacia el exterior) cuando el objetivo ya se tiene centrado y ocupa (casi) toda la imagen. Para ello, definen un mapeado log-polar diferente al convencional. En este sentido, podría ser útil el modelo de Jurie, de forma que, variando el valor de  $\rho_1$ , se obtendrían diferentes mapeados, siempre bajo el mismo modelo. Por último, en la figura 6(d) se muestra cómo afecta  $\alpha$  a  $\xi(\rho)$  en el modelo  $\log(z + \alpha)$ .



## 4 Conclusiones

En este trabajo se ha estudiado, implementado por *software* y comparado diferentes modelos de transformación log-polar. Se han comentado algunas propiedades de las imágenes log-polares, como son: (a) reducción selectiva de información, importante para aplicaciones en tiempo real; (b) tratamiento de rotaciones y escalados como simples traslaciones, relevante en reconocimiento de formas y en estimación del movimiento; (c) mecanismo implícito de ponderación de pixels de la fovea, que resulta útil en aplicaciones de seguimiento activo.

En cuanto al coste temporal de la transformación, resulta interesante utilizar tablas de búsqueda, y explotar la simetría entre los cuadrantes, como se hace en [16]. Si se utiliza subdivisión recursiva de pixels para promediar los niveles de gris, la precisión utilizada afecta de forma importante al coste de la transformación, tanto en la obtención *off-line* del mapa como en cada una de las conversiones.

Respecto a la relación calidad visual/tamaño, los métodos son similares, aunque posiblemente el de Jurie ofrezca una mejor relación. Por otro lado, sólo el modelo de Sandini presenta la propiedad de invarianza a escalados. La falta de esta propiedad es incluso más evidente en el modelo de Jurie, lo cual limita su interés.

La función logarítmica, que indica la variación de la resolución radial a lo largo del sensor, es también apropiada para medir cuántos pixels se dedican al centro respecto al exterior. Todos los métodos tienen un comportamiento similar, aunque el modelo  $\log(z + \alpha)$  presenta una tendencia mayor que los demás a dedicar más píxels al centro que a la periferia. También se ha estudiado cómo varía esta función al variar un parámetro de cada método. En este sentido, el de Jurie ofrece cierta flexibilidad que podría ser interesante en ciertas aplicaciones, como la indicada en [12].

## Referencias

- [1] Richard Alan Peters II, Magued Bishay, and Tamara Rogers. On the computation of the log-polar transform. Technical report, Vanderbilt University, School of Engineering, March 1996.
- [2] Nick Barnes and Giulio Sandini. Active docking based on the rotational component of log-polar optic flow. In W.-H. Tsai and H.-J. Lee, editors, *Asian Conf. on Computer Vision*, pages 955–960, Taiwan, January 2000.
- [3] Alexandre Bernardino and José Santos-Victor. Visual behaviors for binocular tracking. *Robotics and Autonomous Systems*, 25:137–146, 1998.
- [4] Marc Bolduc and Martin D. Levine. A review of biologically motivated space-variant data reduction models for robotic vision. *Computer Vision and Image Understanding (CVIU)*, 69(2):170–184, February 1998.

- [5] C. Capurro, F. Panerai, and G. Sandini. Dynamic vergence using log-polar images. *Intl. Journal of Computer Vision*, 24(1):79–94, 1997.
- [6] J. Ruiz del Solar, C. Nowack, and B. Schneider. VIPOL: A virtual polar-logarithmic sensor. In *Scandinavian Conf. on Image Analysis (SCIA)*, pages 739–744, Finland, 1997.
- [7] Kazuhiro Hotta, Takio Kurita, and Takethoshi Mishima. Scale invariant face recognition method using spectral features in log-polar image. In *Applications of Digital Image Processing XXII (SPIE Proc.)*, pages 33–43, 1999.
- [8] F. Jurie. A new log-polar mapping for space variant imaging. Application to face detection and tracking. *Pattern Recognition*, 32:865–875, 1999.
- [9] Fernando Pardo-Carpio. *Sensor Retínico Espacio Variante Basado en Tecnología CMOS*. PhD thesis, Dept. Informàtica i Electrònica, Universitat de València, September 1997.
- [10] Giulio Sandini, Paolo Questa, Danny Scheffer, Bart Dierickx, and Andrea Mannucci. A retina-like CMOS sensor and its applications. In *Proc. 1st. IEEE SAM Workshop*, Cambridge, USA, March 2000.
- [11] Samir Shah and Martin D. Levine. Information processing in primate retinal cone pathways: A model. Technical Report TR-CIM-93-18, Center for Intelligent Machine, McGill University, Montreal, Québec, Canada, December 1993.
- [12] Won-Kyung Song, Jong-Sung Kim, and Zeungnam Bien. Visual servoing for human-robot interaction in the wheelchair-based rehabilitation robot. In *IEEE Intl. Conf. on Systems, Man and Cybernetics*, pages 1811–1816, Nashville, Tennessee, USA, October 2000.
- [13] A. L. Thornton and S. J. Sangwine. Log-polar sampling incorporating a novel spatially variant filter to improve object recognition. In *Intl. Conf. on Image Processing and its Applications (IPA)*, pages 776–779, Dublin, July 1997.
- [14] Massimo Tistarelli and Giulio Sandini. On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 15:401–410, 1993.
- [15] V. J. Traver and F. Pla. An optimization approach for translational motion estimation in log-polar domain. In *Intl. Conf. on Computer Analysis of Images and Patterns (CAIP)*, Warsaw, Poland, September 2001. Accepted.
- [16] Richard S. Wallace, Ping-Wen Ong, Benjamin B. Bederson, and Eric L. Schwartz. Space variant image processing. Technical Report 633, Courant Institute of Mathematical Sciences, New York University, NY, April 1993.