

Imagen Log-Polar: Modelo de visión espacio-variante. Log -Polar Image: Vision model space -variant.

Omar Vidal-Pino  
Lic. Lisandra Guerrero-Pérez

## RESUMEN

La transformación *log-polar* es un modelo de visión espacio-variante ampliamente aceptado en la comunidad científica. La geometría *log-polar* juega un papel fundamental en el campo de la visión activa, pues supone una solución de compromiso entre dos objetivos contrapuestos (un amplio campo de visión y una alta resolución). Dicha solución consiste en una reducción selectiva de información, ofreciendo una resolución que es máxima en el centro de la imagen y que va disminuyendo con la excentricidad. En el artículo se muestran los conceptos fundamentales en los que se basan este tipo de imágenes, así como su utilidad, propiedades, ventajas y desventajas. Además, se plantean algunos modelos matemáticos para realizar su transformación a partir de imágenes cartesianas, profundizando en el modelo base utilizado en la realización de este trabajo.

Palabras clave: Imagen Log-polar, Visión Activa, Modelos Matemáticos, Visión Espacio-variante, Imagen Cartesiana.

## ABSTRACT

The log-polar transformation is a model of space-variant vision widely accepted in the scientific community. The log-polar geometry plays a fundamental role in the field of active vision, because it represents a compromise between two conflicting objectives (a wide field of view and high resolution). This solution is a selective reduction of information, offering a resolution that is greatest at the center of the image and decreases with eccentricity. The article shows the fundamental concepts underlying these images and their use, properties, advantages and disadvantages. Also deals with some mathematical models for processing Cartesian images in order to deepen the base model used in this work.

Key words: Log-polar Image, Active Vision, Mathematical Modeling, Space-variant Vision, Cartesian Image.

## INTRODUCCIÓN

La terminología *visión foveal* (o de espacio variante) basa su nombre en una característica puramente biológica, la fovea, conocida también como centro foveal, que es una parte del ojo, localizado en el centro de la mácula (región de la retina).

Esta importante zona es la responsable de la visión aguda central y es utilizada por los humanos para actividades como la lectura, mirar TV, conducir autos y muchas otras actividades donde el detalle visual es de vital importancia. Las imágenes log-polares constituyen, sin lugar a dudas, el modelo más popular de los sistemas de imágenes foveales. El verdadero origen de este modelo yace en lo profundo de la neuropsicología humana, donde se ha conseguido demostrar que la información capturada por el ojo es distinta de la que procesa el cerebro en el córtex, y que esta transformación retino-cortical puede ser modelada matemáticamente de forma muy exacta mediante una transformación logarítmica-polar (log-polar); la cual es definida en la literatura como mapeo retino-cortical. Este criterio constituye la base para plantear la existencia de una analogía entre la estructura de la retina humana y la de una imagen log-polar.

En el presente artículo se tratan algunos aspectos relacionados con el modelo log-polar, y se analizan varias de sus características, con el propósito de determinar hasta qué punto es factible el proceso de atención visual en este tipo de imágenes obtenidas, ya sea por sensores o por conversión de software. El objetivo de este estudio es aprovechar las facilidades que implica este modelo, como lo es la reducción de información para un mismo campo visual sin perder la conformalidad, propiedad indispensable para la aplicación posterior de las diferentes técnicas y algoritmos.

## MATERIALES Y MÉTODOS

Para el desarrollo de los objetivos propuestos en esta investigación se emplearon diversos métodos tales como:

El *análisis documental* para revisar la bibliografía disponible relacionada con los aspectos teóricos–conceptuales previstos, como son el modelo *log-polar*, la visión foveal, los modelos matemáticos, la visión espacio-variante, así como las imágenes cartesianas y la interrelación de conceptos.

El *Histórico-lógico* para analizar investigaciones precedentes sobre las imágenes log-polares y su ámbito de estudio en las Ciencias de la Computación, con el propósito de precisar sus regularidades y tendencias.

El *Lógico-deductivo* para la selección del modelo para la transformación de imagen cartesiana a *log-polar* pues el modelo teórico –matemático (Modelo Ideal) para representar a estas imágenes tiene sus cimientos en las ecuaciones (1) y (2).

Es este modelo el que plantea que todo punto del plano visual  $z$  puede representarse en el dominio complejo ( $z \in \mathbb{C}$ ) a través de (1) y que el punto correspondiente  $w \in \mathbb{C}$  del plano cortical es calculado mediante una función que calcula el logaritmo complejo de  $z$  (2). Debido a que esta función está definida para la  $z > 0$ , este punto se convierte en un problema para la implementación de este modelo y es la principal razón por la que el mismo es *ideal* debido a que es

puramente logarítmico, pero no puede ser usado directamente para una implementación. Su importancia es solamente teórica.

Por esta razón, existen diferentes modelos para el mapeo de coordenadas cartesianas a log-polares, los cuales difieren principalmente en la manera que ellos tratan la singularidad de la función logarítmica en el origen. El modelo utilizado para la realización de este trabajo es el del centro ciego o *Central Blind Spot* (por su nombre en inglés), por conservar las propiedades invariantes a cambios en escala y rotación. La ecuación (3) es la definición matemática de este modelo para calcular las coordenadas log-polares.

## RESULTADOS Y DISCUSIÓN

En efecto, la distribución de fotorreceptores en la retina sigue un modelo geométrico, el cual es similar a una celda *log-polar*. Pero esa celda, que es usualmente dibujada en el espacio cartesiano, no es realmente la de una imagen *log-polar*, sino que es tratada como un “esquema”. Dicho esquema es utilizado generalmente para comprender de qué manera la geometría *log-polar* actual puede ser calculada desde un dominio cartesiano, y visualizar el contenido de una imagen *log-polar* en un espacio más familiar e intuitivo.

Aunque estudios matemáticos y psicológicos sobre el mapeo de la retina comenzaron desde la década del 40 del siglo XX, fue probablemente Schwartz (1977) el primero que sugirió la existencia de una proyección matemática de la retina sobre la corteza, y propuso una expresión para ésta. Por lo que en este caso, se emplearon variables complejas para describir tal mapeo y se presentaron las principales propiedades geométricas ofrecidas por el mismo.



Figura 1: a) Imagen cartesiana tradicional (467x350), b) Imagen foveal (233x175).

Las ramas de la Ciencia de la Computación que estudian la visión foveal tienen una evidente inspiración biológica y estudian imágenes con una resolución variable. La resolución es máxima en el centro de la imagen y decrece con la excentricidad. Con ello se consigue disminuir drásticamente la cantidad de información a procesar. Este tipo de geometría resulta muy efectiva cuando el objetivo está centrado en la imagen y por lo tanto, se puede observar con gran nivel de detalle.

La menor resolución en la periferia del sensor permite un ahorro importante de recursos, pero evitando a su vez, que ciertos movimientos en la periferia pasen desapercibidos. Suele decirse que la visión central tiene una función reconocedora,

mientras que la visión periférica cumple un cometido alertador. Según la forma en la que el sensor procesa el campo receptivo, es lo que genera la existencia de diferentes diseños de sensores polares. (Ver Figura 2).

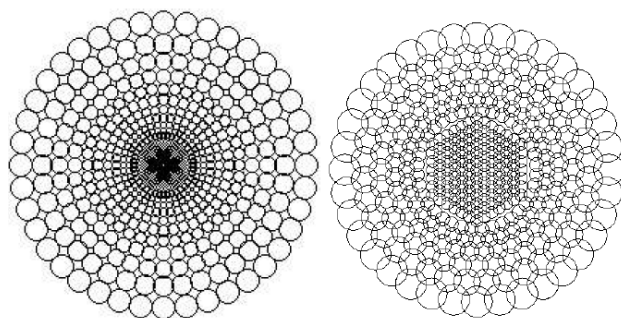


Figura 2: Muestra del diseño de dos tipos de sensores polares. De izquierda a derecha, sin solapamiento y con solapamiento del campo receptivo.

De este modo, la geometría *log-polar* se ha utilizado como modelo fundamental en la visión artificial para construir cámaras foveales, para simular la transformación por software o para la implementación de soluciones mixtas. En los últimos años, dado el aumento de la capacidad de procesamiento de información, han surgido innumerables nuevas aplicaciones para este tipo de geometría, entre ellos están los sistemas en tiempo real (de seguridad generalmente), por lo que su importancia ha crecido con el curso del tiempo.

### Propiedades y Ventajas del modelo log-polar

La principal característica de este modelo es, sin lugar a dudas, la reducción sustancial y selectiva de información, dada por la naturaleza polar del espacio y el incremento logarítmico en el tamaño de los campos receptivos (píxeles en el dominio polar). Propiedad que hace de este modelo un candidato excelente para la implementación de sistemas de detección y reconocimiento en tiempo real por la disminución de datos a procesar con respecto al modelo cartesiano tradicional.

El dominio logarítmico-polar posee además otras bondades, entre las más notables se encuentra la invarianza de rotación y escalado (Pla, 2005), donde rotaciones o escalados situados en el centro de la imagen pasan a ser simples traslaciones en el plano log-polar, simplificando el tiempo de cálculo, propiedad que es explotada en muchos trabajos de seguimiento y, lo que nos concierne, atención visual. Además, estas características aportan mayor robustez a la detección de puntos de interés salientes y, derivado de esto, a la detección de objetos.

Otra de las características del logarítmico-polar es el predominio de píxeles en el centro (fovea) de la imagen, dada la naturaleza del muestreo logarítmico, un número

mayor de píxeles son destinados al centro en contraposición con las regiones periféricas, lo cual se convierte en una poderosa característica que determina un foco de atención implícito, crucial en la atención de un objetivo.

Pero sin lugar a dudas, la propiedad más impresionante y a la vez útil de este tipo de imágenes es la conformalidad, referente al hecho de que los ángulos y vecindades locales son preservados en la imagen *log-polar*, por consecuencia se pueden aplicar la mayoría de los operadores locales que se utilizan en imágenes cartesianas sin realizar un cambio sustancial, lo cual no significa que cualquier técnica de procesamiento de imágenes cartesianas puede ser trasladada al dominio log-polar directamente sin variaciones (Ver Figura 5 c)).

Seguramente, esta última propiedad es la más explotada en el desarrollo e implementación de modelos para el procesamiento de este tipo de imágenes. A continuación, nos detendremos en el análisis de un modelo comúnmente usado para la representación de las imágenes log-polares, analizando sus principales características y su aprovechamiento.

### **Representación espacial y notación utilizada.**

Las imágenes cartesianas son, por lo general, un conjunto de regiones (píxeles) de forma cuadrada o rectangular dispuestas en un plano cartesiano en forma de matriz (filas y columnas), donde cada píxel representa una intensidad de color y con un tamaño único para todos ellos (Figura 3 a)).

Este tipo de imágenes es la que comúnmente vemos cuando nos muestran una fotografía, un cuadro o vemos la TV; y son sin lugar a dudas, la forma principal en que el hombre ha almacenado, expresado y procesado la información visual que le rodea desde todos los tiempos.

Por su parte las imágenes log-polares, no parten de un plano cartesiano, sino como su nombre lo indica, de un plano polar representado por anillos y sectores conformados por rejillas, dispuestas por este tipo de geometría, a las cuales llamaremos campos receptivos, los cuales, al igual que las imágenes cartesianas, representan una intensidad de color pero no con un único tamaño para todas ellas (Figura 3 b)).

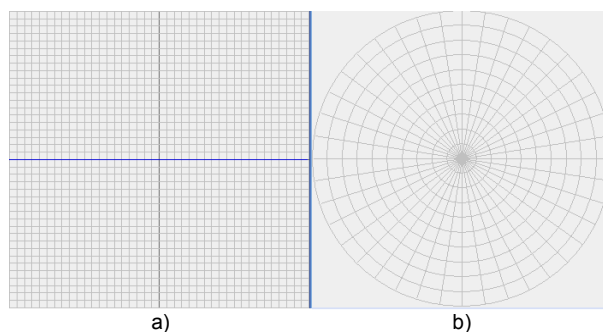


Figura 3: Ejemplos de ambos dominios: a) Dominio Cartesiano, b) Dominio Polar.

En lo sucesivo, a las imágenes cartesianas (también conocidas como *imágenes retinales*), las denotaremos por  $I$  y las log-polares o *corticales* por  $I'$ , en ambos casos algunos subíndices pueden ser adicionados. Los píxeles de una imagen cartesiana se denominan píxeles cartesianos, y equivalentemente píxeles log-polares los de la imagen log-polar (Ver Figuras 3 y 4).

El contenido de los píxeles log-polares es obtenido a partir de un conjunto de píxeles cartesianos. Este conjunto de píxeles, es conocido como área de soporte y es a lo que llamaremos *campo receptivo* (Figura 5 c)).

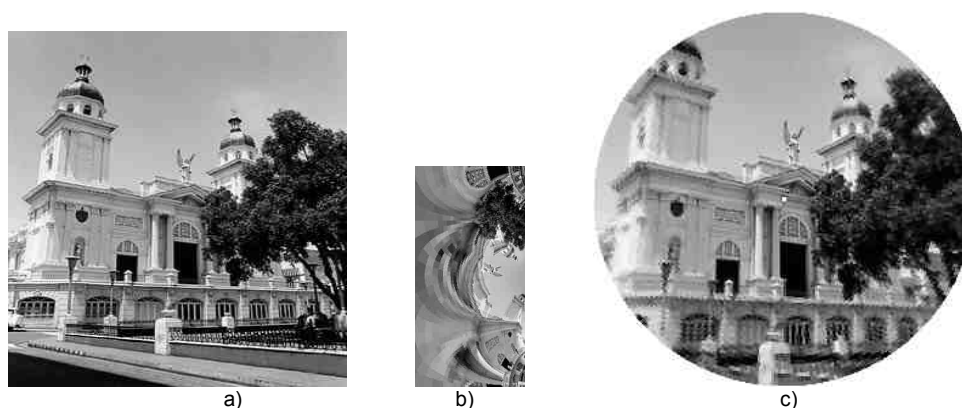


Figura 4: Transformación log-polar: a) Imagen cartesiana tradicional (256 x 240), b) Imagen cortical (50 x 90), c) Imagen retínica o foveal.

Las imágenes cartesianas serán descritas usualmente por su anchura  $M$  y su altura  $N$ , estos parámetros por lo general denotan el número de filas y columnas de la imagen respectivamente, por lo cual se puede denotar el tamaño de una imagen cartesiana discreta como  $M \times N$ . Equivalentemente, las imágenes log-polares discretas tienen  $R$  anillos y  $S$  sectores. Debido a que las imágenes log-polares pueden ser representadas en la memoria de la computadora como si fueran imágenes cartesianas convencionales,  $R \times S$  será la manera de representar su tamaño (número de columnas y filas respectivamente). Finalmente, otro parámetro importante, el campo de visión, será representado por  $\rho_{\max}$  (anillo máximo de visión polar).

Para referirnos a una localización en una imagen cartesiana usaremos las coordenadas continuas  $(x, y)$  o las discretas  $(i, j)$ , y en el plano log-polar,  $(\xi, \eta)$  denota un punto continuo mientras que  $(u, v)$  denota un punto discreto.

Las coordenadas polares serán tratadas con su significado continuo  $(\rho, \theta)$ , y son relacionadas con las coordenadas cartesianas por las expresiones conocidas:

$$\begin{aligned}\rho &= \sqrt{x^2 + y^2} \\ \theta &= \arctan \frac{y}{x}\end{aligned}\tag{1}$$

Donde  $\rho$  denota la distancia del punto  $(x, y)$  al origen y  $\theta$  al ángulo entre la línea que une al punto asociado al origen y el eje horizontal x.

### Modelos para la transformación log-polar.

- Modelo Ideal.

El mapeo retino-cortical puede ser descrito como un mapeo complejo del campo visual sobre la superficie cortical. Si  $z \in \mathbb{C}$  representa un punto en el plano visual, la función que mapea  $z$  con el punto correspondiente  $w \in \mathbb{C}$  en el plano cortical es dado por el logaritmo complejo:

$$w = \ln(z).\tag{2}$$

Este modelo es *ideal* debido a que es puramente logarítmico, dado que no puede ser usado directamente para una implementación, su importancia es solamente teórica.

La Figura 5 muestra la correspondencia entre los campos receptivos de la imagen cartesiana y los píxeles log-polares de la imagen cortical. A menudo se hace referencia al centro de una imagen, cuando se trate de una imagen cartesiana de  $M \times N$  este centro será el punto  $(x_c, y_c) = (M/2, N/2)$ . Este punto será estimado como centro de fijación, alrededor del cual la imagen log-polar será construida. Cuando la palabra *centro* sea usada en el contexto de imágenes log-polares, nos estaremos refiriendo a este punto y no al centro de la imagen log-polar  $(R/2, S/2)$ .

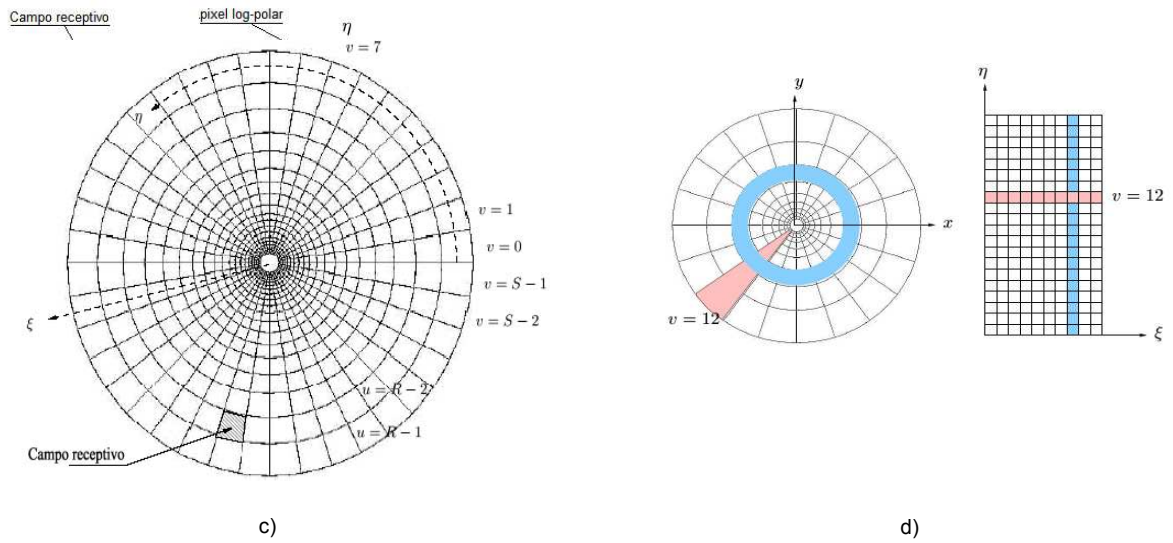


Figura 5: Correspondencia entre campos receptivos y píxeles log-polares. a) Imagen log-polar. b) Imagen cortical, c) elementos en el mapeo log-polar, d) Correspondencia entre anillos y sectores con columnas y filas respectivamente.

- Modelo del Centro Ciego (*Central Blind Spot*).

Existen diferentes modelos para el mapeo de coordenadas cartesianas a log-polares, los cuales difieren principalmente en la manera que ellos tratan la singularidad de la función logarítmica en el origen.

El modelo utilizado para la realización de este trabajo es el del centro ciego o *central blind* (por su nombre en inglés), por conservar las propiedades invariantes a cambios en escala y rotación. Es un modelo ampliamente difundido y aceptado por la comunidad científica y se ha convertido en uno de los más populares. Fue propuesto por Krüger (1995), Tistarelli (1992) y Sandini (1993). La ecuación (3) es la definición propuesta para calcular las coordenadas log-polares:

$$\xi = \log_a \left( \frac{\rho}{\rho_0} \right) \quad (3)$$

$$\eta = q * \theta$$

Donde  $\rho_0$ ,  $a$  y  $q$  son parámetros de la transformación. Podemos asumir que el radio más interior de la transformada (o sea, el radio del *Blind Spot*)  $\rho_0$  es conocido.

Aunque es válido aclarar que este parámetro puede ser elegido arbitrariamente. Como para el valor de  $a$ , la ecuación anterior puede ser instanciada para el caso del radio más exterior  $\rho_{\max}$ , el cual corresponde a la coordenada radial  $\eta = R$ . Entonces,



$$R = \log_a \left( \frac{\rho_{\max}}{\rho_0} \right) \quad (4)$$

Considerando que el logaritmo con base arbitraria  $a$  puede ser escrito como una función de logaritmos naturales  $\log_a b = \frac{\ln b}{\ln a}$  la expresión anterior puede ser escrita de nuevo como sigue:

$$R = \frac{\ln \left( \frac{\rho_{\max}}{\rho_0} \right)}{\ln a} \quad (5)$$

De esta manera resulta casi evidente el cálculo de  $a$ :

$$a = e^{\frac{\ln \left( \frac{\rho_{\max}}{\rho_0} \right)}{R}} \quad (6)$$

Respecto al parámetro  $q$  en (3) se puede decir que denota la resolución angular, por ejemplo, el número de sectores por radian,  $q = \frac{S}{2\pi}$ . Esto permite el mapeo de un ángulo  $\theta \in [0, \dots, 2\pi[$  a una coordenada log-polar angular  $\eta \in [0, \dots, S-1]$ .

## CONCLUSIONES

El modelo *log-polar* es un modelo de visión espacio-variante ampliamente aceptado en la comunidad científica, que ofrece facilidades en tanto su particularidad de reducción de información para un mismo campo visual, no implica la pérdida de conformalidad.

En el presente artículo se ha hecho referencia a los aspectos más relevantes y actualizados en la literatura, relacionados con los diferentes modelos existentes para el mapeo de coordenadas cartesianas a log-polares. En el caso del modelo del centro ciego enfatizamos en él por conservar las propiedades invariantes a cambios en escala y rotación; además de que es un modelo muy popular ampliamente difundido y aceptado por la comunidad científica.

Además, este modelo puede ser implementado sin muchas limitaciones y en los últimos tiempos se ha convertido en punto de mira para la utilización, en conjunto con las más renovadoras técnicas para la detección de puntos salientes en la escena, en el proceso de atención visual, por lo que resulta de gran interés el uso de

este modelo en el proceso de atención visual, directamente en el dominio *log-polar*, tema que lo convierte en uno de los modelos de transformación *log-polar* más usados en esta rama de la Visión por Computador.

## BIBLIOGRAFÍA

- Ahnelt, P.; Kolb, H. "Cells and cone photoreceptors in human retina: a golgi-electron microscopic study of spectral connectivity". Journal of comparative Neurology. 1994.
- Bernardino, Alexandre J. Malheiro. "Binocular Head Control with Foveal Vision : Methods and Applications". Dissertacao para a obtencao do Grau de Doutor em Engenharia Electrotecnica e de Computadores, Abril 2004.
- Gomes, H. M; Fisher, R.B. Primal sketch feature extraction from a log- polar image. Pattern Recognition Letters. 2003.
- Itti, L.; Koch, C.; Niebur, E. "A model of saliency-based visual attention for rapid scene analysis". IEEE PAMI, 1998.
- Itti, L.; Koch, C. "A saliency-based search mechanism for overt and covert shifts of visual attention". Vision Research, 2000.
- Itti, L.; Koch, C. "Computational modelling of visual attention". Nature Reviews Neuroscience, 2001.
- Jordan, T. "Anticipation of moving stimuli by the retina". Nature (Long), 1999.
- Koch, C.; Ullman, S. Shifts in selective visual attention: Towards the underlying neural circuitry. Human Neurobiology, 1985.
- Koenderink, J. "The structure of images", Biological Cybernetics, 1984, no. 50, pag. 363-396.
- Leventhal, A.G. The Neural Basis of Visual Function: Vision and Visual Dysfunction, vol. 4. Boca Raton, Fla.: CRC Press, 1991.
- Mahowald, Misha. "An Analog VLSI system for Stereoscopic Vision". Kluwer Academic Publishers, 1994.
- Martínez Álvarez, A. Generación automática de sistemas bioinspirados de visión en hardware reconfigurable". Tesis Doctoral, Universidad de Granada, Mayo 2006.
- Pajares Martinsanz, Gonzalo." Visión por Computador: Imágenes Digitales y aplicaciones". Universidad Complutense de Madrid. 2001
- Schnitzer, M.J. "Multineuronal firing patterns in the signal from eye to brain". Neuron, 37:499-511, 2003.

Recibido: 6 de abril del 2010

Aceptado: 30 de marzo del 2012

Omar Vidal Pino

Licenciado en Ciencias de la Computación

Especialista en Administración y Gestión de Redes Informáticas

Profesor de la Sede Universitaria Ernesto Guevara de la Serna, Universidad Vladimir I. Lenin, Filiación Institucional, Las Tunas, Cuba.

e-mail: [mrvidalpino@gmail.com](mailto:mrvidalpino@gmail.com)

Lic. Lisandra Guerrero Pérez  
Licenciada en Bibliotecología y Ciencias de la Información  
Especialista de la Biblioteca de la Universidad Vladimir Illich Lenin, Filiación  
Institucional, Las Tunas, Cuba.  
e-mail: [lisandragp@ult.edu.cu](mailto:lisandragp@ult.edu.cu)