

SEGUIMIENTO DE OBJETOS MEDIANTE UNA CÁMARA LOG POLAR CON MOVIMIENTO PROPIO

Jose A. Boluda¹, Fernando Pardo¹, Juan J. Pérez¹, Juan Domingo², Joan Pelechano²

¹Departament d'Informàtica i Electrònica. Universitat de València
C/ Doctor Moliner, 50. 46100 Burjassot. Spain.
email: jboluda@glup.eleinf.uv.es

²Institut de Robòtica. Universitat de València
C/Hugo de Moncada, 4. 46010 Valencia. Spain.

1. Resumen

En el presente artículo se presenta una nueva aplicación para el seguimiento de objetos con una cámara log-polar. El algoritmo aplicado fue diseñado previamente en coordenadas cartesianas y detecta la presencia de objetos con movimiento propio, independientemente del movimiento de la cámara. Las restricciones impuestas en el método original traducidas a nuestra adaptación log-polar pasan a ser que la traslación de la cámara se realice a través del eje óptico y que la rotación, si la hay, sea también en ese eje. Esta aplicación del formalismo log-polar parece muy orientado a la detección de objetos en movimiento desde plataformas de navegación automática. La simplicidad del método hace que sea realizable mediante hardware específico para aplicaciones en tiempo real. Se incluyen imágenes reales que demuestran el funcionamiento de la aplicación.

2. Introducción

El problema del seguimiento de movimiento en tiempo real, desde plataformas que a su vez se muevan, es un aspecto importante para sistemas de navegación autónoma. Esto se realiza típicamente para el cálculo del tiempo al impacto y navegación evitando obstáculos. Las restricciones para la implementación de estos algoritmos en tiempo real son la gran cantidad de datos que deben ser procesados y el alto coste computacional de los algoritmos implementados.

La mayoría de estos métodos usan el cálculo previo del flujo óptico [1]. Esta aproximación es robusta y razonablemente precisa pero tiene una implementación con alto coste computacional. El cálculo del flujo óptico es en cualquier caso una metodología lenta y poco realizable en tiempo real [2]. Hay también algoritmos para la detección de objetos en movimiento sin un cálculo previo del flujo óptico, pero de todas maneras no son muy sencillos [3].

No parece que las coordenadas cartesianas sean las naturales cuando nos estamos refiriendo al cálculo del flujo óptico desde una plataforma con traslación en el sentido del eje óptico. El tipo de desplazamiento en la imagen de los objetos que se aproximan sugieren unas coordenadas polares. Existen varios trabajos que demuestran la simplificación de estos cálculos en coordenadas polares [4]. Finalmente se ha optado por adaptar un algoritmo ya existente en coordenadas cartesianas. Este algoritmo está muy orientado a una implementación hardware [5] y, traduciendo los supuestos a coordenadas log-polares, el método se simplifica extraordinariamente.

3. La cámara log-polar

La geometría espacio variante realiza una transformación topológica, llamada transformación log-polar, desde el plano retínico, representado con coordenadas (x,y) , al plano cortical representado con coordenadas (ξ,γ) . En la figura 1 podemos observar gráficamente la transformación realizada que sigue las ecuaciones (1).

$$\begin{cases} \xi = \log r = \log \sqrt{x^2 + y^2} \\ \gamma = \theta = \arctg \frac{y}{x} \end{cases} \quad (1)$$

Esta transformación hace una reducción selectiva de la información, dando una mayor densidad de información en el centro y una menor cantidad de pixels en la periferia. Esta representación de la información es muy similar a la distribución de células fotosensibles en el ojo humano con una mayor resolución en el centro (fovea) y menor resolución en la periferia. Con esta distribución se consigue un compromiso entre la cantidad de información a procesar y la amplitud del campo visual.

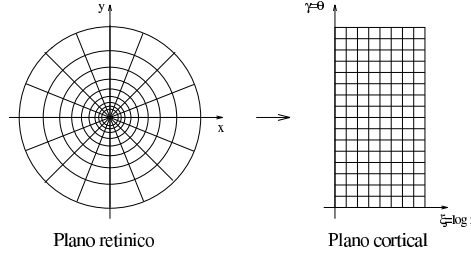


Figura 1: La transformación log-polar.

Esta disposición ha sido físicamente realizada en un sensor CMOS [6] con una resolución de 76 anillos y 128 pixels por anillo. Una cámara ha sido implementada con el sensor [7] consiguiendo 100 imágenes por segundo, una alta tasa que será muy útil en uno de los supuestos del algoritmo. El layout del sensor tiene el aspecto mostrado por el plano retinico, siguiendo la distribución de fotocélulas la distribución log-polar. El acceso a la información es muy sencillo ya que se accede directamente a un número de celda en un anillo, es decir, se accede a la matriz del plano cortical. Es como si la retina se hubiese desdoblado y extendido hasta obtener el plano cortical.

El sensor tiene además la interesante propiedad de que permite el acceso aleatorio a cada pixel de la misma manera que a una ROM, lo cual es muy ventajoso para así solo almacenar/procesar los pixels interesantes. Esta peculiaridad hace que la cámara CMOS previamente desarrollada sea muy diferente a una cámara CCD.

4. Metodología

4.1. Algoritmo original

El método para la detección de objetos en movimiento independientemente de que la cámara se mueva o no ha sido desarrollado en coordenadas cartesianas por Cheng y Nandhakumar [5]. La formulación en el plano cartesiano es como sigue.

Sea $E(x,y,t)$ la secuencia de imágenes en tonos grises que varía con el tiempo. Sea $T(\Sigma)$ la proyección de una superficie Σ en el plano de la imagen Π y $\partial T(\Sigma)$ los puntos de la frontera de $T(\Sigma)$. Los supuestos para el método son los siguientes:

- $E(x,y,t)$ es lineal a tramos con respecto a x e y en cualquier punto de la proyección de una superficie en el plano imagen Π . Esto significa que se cumple que $\forall (x,y) \in T(\Sigma) - \partial T(\Sigma)$

$$\frac{\partial^2 E}{\partial x^2} = \frac{\partial^2 E}{\partial y^2} = 0 \quad (2)$$

- El movimiento de la imagen es suave con respecto al tiempo es decir, $\forall (x,y) \in T(\Sigma)$

$$\frac{\partial^2 x}{\partial t^2} = \frac{\partial^2 y}{\partial t^2} = 0 \quad (3)$$

Con estos supuestos y a partir de la ecuación del flujo óptico en coordenadas cartesianas se demuestra que $\forall (x,y) \in \partial T(\Sigma)$

$$\frac{\partial^2 E}{\partial t^2} = 0 \quad (4)$$

Mediante el cálculo de la derivada temporal de segundo orden de la secuencia de imágenes es posible detectar los bordes de los objetos que se muevan. El método tiene en cuenta el cálculo del movimiento de la cámara en cualquier dirección, siempre que este sea lineal.

El supuesto de imagen lineal a tramos se consigue mediante el cálculo de un filtro óptimo para eliminar las transiciones bruscas en los bordes. El supuesto de linealidad en el movimiento se asume aproximadamente cierto para secuencias de 3 imágenes tomadas consecutivamente de forma muy rápida. Una vez se computa la derivada segunda temporal se binariza la imagen con un cierto valor umbral para detectar el borde del objeto que se mueve.

4.2. Adaptación a las coordenadas log-polares

En este caso el espacio de trabajo será el plano cortical que es la transformación que el sensor log-polar realiza. Deben de trasladarse las condiciones (2) y (3) a las coordenadas corticales (ξ, γ) . El primer supuesto trasladado a coordenadas log-polares es directamente:

$$\frac{\partial^2 E}{\partial \xi^2} = \frac{\partial^2 E}{\partial \gamma^2} = 0 \quad (5)$$

Esta condición viene a expresar una cierta suavidad en la imagen, es decir no deben de haber bordes dentro de un mismo objeto. Esta condición se consigue mediante la aplicación de un filtro pasa baja en la forma de una máscara de convolución. Esto hace una linearización en el plano cortical para conseguir que se cumpla la condición (5). Este método no es óptimo en absoluto pero si que es mas fácil de implementar por hardware y por eso ha sido escogido.

El segundo supuesto (3) en coordenadas log-polares es:

$$\frac{\partial^2 \xi}{\partial t^2} = \frac{\partial^2 \gamma}{\partial t^2} = 0 \quad (6)$$

Estas ecuaciones son pasadas a coordenadas cartesianas mediante (1) para saber que tipo de movimiento es el descrito obteniéndose:

$$\begin{cases} (x\ddot{y} + \dot{x}\dot{y} - \dot{x}\dot{y} - \ddot{x}y)(x^2 + y^2) = (x\dot{y} - \dot{x}y)(2x\dot{x} + 2y\dot{y}) \\ (\dot{x}^2 + \dot{y}^2 + x\ddot{x} + y\ddot{y})(x^2 + y^2) = (x\dot{x} + y\dot{y})(2x\dot{x} + 2y\dot{y}) \end{cases} \quad (7)$$

La solución a estas ecuaciones diferenciales son funciones que expresan un movimiento log-spiral de la forma:

$$\begin{cases} x = Ae^{b+vt} \cos(w_0 + wt) \\ y = Ae^{c+vt} \sin(w_0 + wt) \end{cases} \quad (8)$$

Donde A, b, v, w_0 y w son constantes de integración. Donde w expresa la velocidad angular de rotación y v la velocidad exponencial de traslación.

Este resultado parece extraño, pues no detectar sólo objetos que se muevan de esa manera nada habitual no parece muy útil. Para encontrar alguna similitud con algún movimiento conocido se va a ver como crece un objeto en el plano imagen cuando la cámara se aproxima hacia él. Para ello el movimiento de la cámara quedará restringido a una rotación sobre el eje óptico y una traslación sobre este mismo eje. Tal como ocurriría en un misil, o (sin rotación) en un robot móvil con la cámara apuntando en la dirección de avance. Observando la figura 2, se obtiene que la expresión para el tamaño en el plano imagen de la flecha de tamaño d que se aproxima con velocidad v_t desde z_0 viene dada por:

$$d' = f \frac{d}{z_0 - v_t t} \quad (9)$$

Donde f es la distancia focal, d el tamaño del objeto que se aproxima, z_0 la distancia en $t=0$, v_t la velocidad de traslación del objeto (o cámara) y del tamaño en el plano retínico.

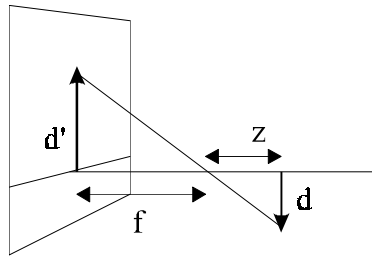


Figura 2: Tamaño de un objeto que se aproxima

En el caso de rotación alrededor del eje de traslación añadiríamos para cada componente la función seno o coseno. Por simplicidad la dependencia angular se ha obviado. Desarrollando en series de Taylor la función (9) alrededor del origen y haciendo lo mismo con la parte no angular de las ecuaciones (8) se obtiene:

$$\frac{1}{(1-X)} = \sum_{n=0}^{\infty} X^n \quad (10)$$

$$e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}$$

Se puede comprobar que ambas expresiones empiezan a diferir en el término de tercer orden. Por tanto es lícito realizar la aproximación, tal como los resultados experimentales mostrarán, de que las soluciones (8) a las condiciones (6) son exactamente objetos que se aproximan a la cámara en la dirección del eje óptico. Con esta aproximación se obtiene que no se considerará como movimiento la variación en la imagen debida a la aproximación de los objetos estáticos con respecto a la cámara. De la misma manera si la cámara tiene una rotación alrededor del eje óptico tampoco será tenida en cuenta como movimiento siempre que sea a velocidad angular constante. Esto se puede conseguir tomando imágenes muy rápidamente y suponiendo que la variación entre 3 imágenes es aproximadamente lineal.

El método aplicado viene a ser el mismo que el desarrollado por Cheng y Nandhakumar con las restricciones de que la traslación y/o rotación de la cámara debe de ser a través del eje óptico, y es:

1. Adquisición de las imágenes $E(\xi, \gamma, t)$.
2. Aplicación del filtro pasa-baja para suavizar los bordes.
3. Cálculo de la derivada de segundo orden temporal.
4. Binarización de la imagen con un umbral

5. Resultados experimentales

El algoritmo ha sido probado con imágenes reales tomadas en el laboratorio con la cámara log-polar. La cámara fue montada en una cinta y sobre un soporte que permitía la rotación sobre el eje de avance, que es el eje óptico. La secuencia de imágenes de la figura 3 muestra la cámara avanzando en el laboratorio y rotando a su vez. Al fondo de la imagen se puede observar un pequeño polígono que se desliza de izquierda a derecha. Siguiendo la secuencia podemos observar como hay objetos cuyo tamaño aumenta debido al movimiento de la cámara, además de la rotación de la cámara.

Después de la adquisición se ha aplicado una máscara de convolución de 5×5 para todo los pixels excepto los de los 2 anillos mas exteriores, por carecer de los pixels vecinos necesarios. El hecho de que no se haya aplicado una linearización ideal y de que se haya realizado la aproximación para el movimiento no representa un problema si se ajusta el valor umbral para la binarización. Este valor umbral será menor cuanto las condiciones (5) y (6) sean mas ciertas.

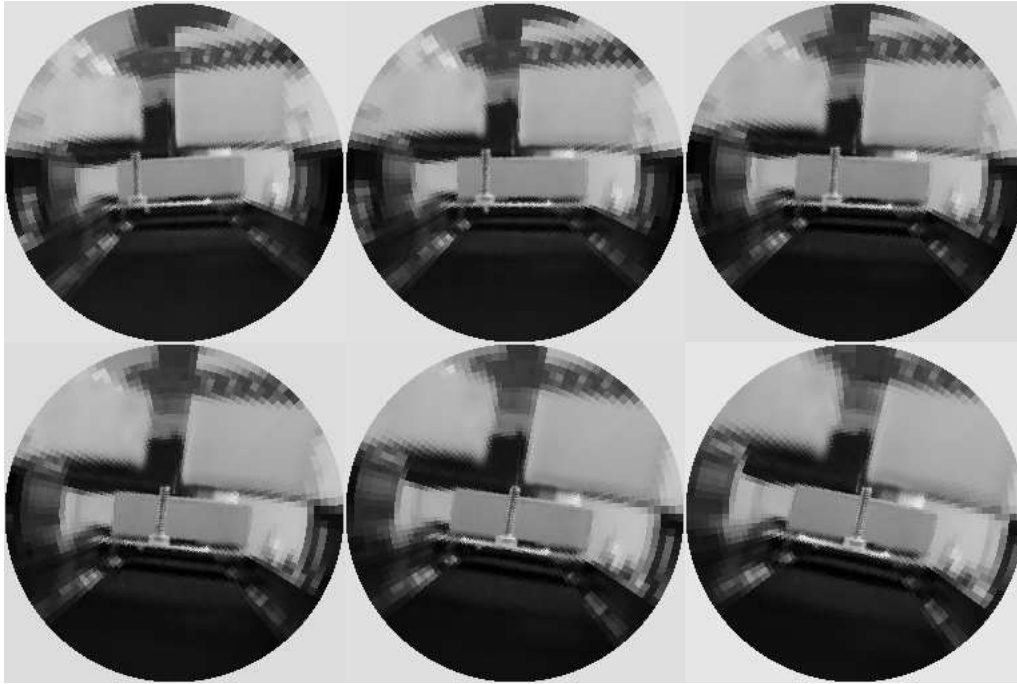


Figura 3: Secuencia de imágenes reales

En la figura 4 se puede observar el resultado de la aplicación del algoritmo con un valor umbral de 65. Si este valor se disminuye aparecen pixels seleccionados en la periferia de la imagen debido a que la linearización ahí no es perfecta y que se está lejos del origen, con lo cual la aproximación que se hace en (10) no es cierta.

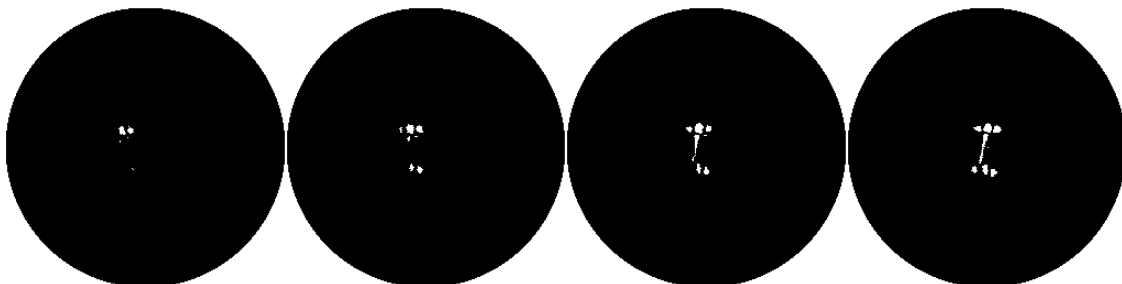


Figura 4: Resultado del algoritmo aplicado a la secuencia de imágenes de la figura 3

Se puede observar que el algoritmo sigue el objeto que se mueve independientemente del movimiento de la cámara. El ajuste del umbral para la binarización hace que las aproximaciones sean correctas. Un valor muy alto de este umbral puede hacer, como se ve en la primera imagen binarizada, que apenas se detecte el movimiento si este es muy lento.

6. Conclusiones

Se ha realizado una aplicación para la detección de movimiento desde una cámara que a su vez se mueva. La aplicación del método desarrollado por Cheng y Nandhakumar a coordenadas log-polares a mostrado ser muy útil cuando el movimiento de la cámara se realiza en la dirección del eje óptico. Incluso la rotación de la cámara está incluida y no dificulta la detección de objetos. Las aproximaciones realizadas con la idealidad del filtrado de la imagen y de la solución a las condiciones (6) se pueden resolver con el ajuste del valor umbral.

La validez práctica del método es todavía mayor por estar muy orientado a la implementación por hardware. Es fácil y rápido aplicar la máscara de convolución y restar imágenes para calcular las imágenes derivadas. De la misma manera la binarización es sencilla y paralelizable. Si se añade la característica de reducción selectiva de información debida a la utilización del formalismo log-polar (aproximadamente 8Kb/imagen) se obtiene un algoritmo con poco coste computacional.

Actualmente se está realizando el módulo hardware que, situado en una plataforma móvil y reorientable, suministre a un robot móvil información sobre la posición en tiempo real de objetos que se muevan.

7. Agradecimientos

Este trabajo ha sido parcialmente subvencionado por la CICYT con el proyecto TAP95-1086-C02-02 . Este trabajo también ha sido parcialmente desarrollado en el Machine Vision Lab de la Universidad de Virginia.

8. Referencias

- [1]. Thomson, V. A. Berzins and K. M. Mutch. dynamic occlusion analysis in optical flow fields. *IEEE Trans. Patt. Anal. Machine Intelligence.*,7 N° :374-383, 1985
- [2]. Beauchemin and J. L. Barron. The computation of the optical flow. *ACM Computing Surveys*, Vol. 27, N° 3. September 1995.
- [3]. Torr and D. W. Murray. Statistical detection of independent movement from a moving camera. *Image and Vision Computing*, 11 (4):180-187, 1993.
- [4]. Daniilidis and V. Krüger. Optical flow computation in the log-polar plane. *Proc. Inter. Conf. on Computer Analysis of Images and Patterns*, Prague, Czech Republic, September 1995
- [5]. Chen and N. Nandhakumar. A Simple scheme for motion boundary detection. *Proc. IEEE Inter Conf. on Systems, Man and Cybernetics.* 1994.
- [6]. Pardo, J.A. Boluda, J.J Pérez, S. Felici, B. Dierickx and D. Scheffer. Response properties of a foveated space-variant CMOS image sensor. *International IEEE Symposium on Circuits and Systems.* Atlanta, USA; May 1996.
- [7]. Boluda, F. Pardo, T. Kayser, J.J. Pérez and J. Pelechano. A new foveated space-variant camera for robotic applications. *The Third IEEE International Conference on Electronics, Circuits and Systems.* Rodos, Greece, October 1996.