

PERCEPCIÓN DE PROFUNDIDAD USANDO VISIÓN ESTEREOSCÓPICA

Lurgo G. , Neira R. , Musso D. y Daniele F.

Universidad tecnológica Nacional – Facultad Regional San Francisco
Av. De la Universidad 50- San Francisco (Pcia. de Córdoba) – CP (2400)
E-mail: slurgo@arnet.com.ar

INTRODUCCIÓN

La visión estereoscópica es el proceso de percibir las distancias relativas a los objetos basándose en el desplazamiento lateral entre dos imágenes retínales correspondiente a una escena del mundo real. Este desplazamiento lateral se conoce como disparidad, siendo esta inversamente proporcional a la distancia de los objetos percibidos.

La triangulación utiliza la distancia conocida de dos puntos separados mirando el mismo punto de la escena. A partir de estos parámetros, se puede calcular la distancia al punto de escena. La Figura 1 a continuación muestra gráficamente la geometría.

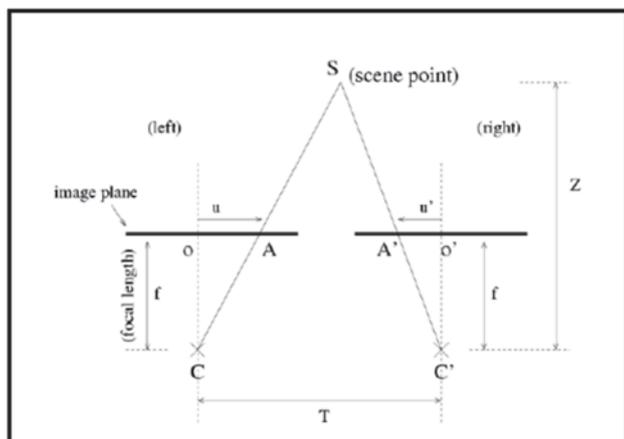


Figura 1 - Triangulación en un sistema Estéreo Visión

En la disposición anterior, dos cámaras (C, C') ven el mismo punto de característica (S). La ubicación del punto en los dos planos de imagen se denota por A y A'. Cuando las cámaras están separadas por una distancia T, la ubicación de A y A' del eje normal de las cámaras será diferente (denotado por U, U'). Usando estas diferencias, la distancia (Z) al punto se puede calcular a partir de la siguiente fórmula:

$$Z = f \frac{T}{u-u'} \quad (1)$$

La obtención de información de profundidad se logra a través de un proceso de cuatro pasos. En primer lugar, las cámaras deben calibrarse. Después de calibrar las cámaras, se supone que las diferencias en las imágenes están en la misma línea horizontal o epipolar (Foggia et al., 1997). El segundo paso es la decisión de qué método se utilizará para encontrar las diferencias entre las dos imágenes. Una vez que se toma esta decisión, es necesario diseñar o decidir un algoritmo para

obtener el mapa de disparidad. El tercer paso es implementar el algoritmo para obtener la información de disparidad. El paso final es utilizar la información de disparidad, junto con la calibración de la cámara establecida en el paso uno, para obtener una vista tridimensional detallada de la escena.

A pesar de haber muchos algoritmos disponibles, hay dos de uso común que se utilizan actualmente para encontrar la disparidad. El primer método está basado en características. El segundo método es un método estadístico basado en áreas. Debido a que son ampliamente utilizados, nos centraremos en estos dos métodos en este informe

El primer método examina las características en una imagen e intenta encontrar la característica correspondiente en la otra. Las características pueden ser bordes, líneas, círculos y curvas. (Nasrabadi, 1992) aplica un algoritmo de coincidencia basado en segmentos curvos. Los segmentos de curva se usan como bloque de construcción en el proceso de correspondencia. Los segmentos de curva se extraen de los puntos de borde detectados. El centro de cada curva extraída se usa como la característica en el proceso de correspondencia.

(Medioni y Nevatia, 1985) utilizan segmentos de puntos de borde conectados como primitivas coincidentes. La correspondencia estereó se logra minimizando la medida de disparidad diferencial para la correspondencia global, teniendo en cuenta cosas como los puntos finales y la orientación del segmento. Para cada característica en la imagen de la izquierda (QL) debe haber una característica similar en la imagen de la derecha (QR). Se necesita una medida de similitud para asociar las dos características. Esta medida viene dada por la siguiente fórmula adaptada de (Candocia y Adjouadi, 1997)

$$\varphi(L \rightarrow R) = \frac{1}{NL} \sum_{q=1}^h \frac{1}{(D_q+1)} \quad (2)$$

donde NL = número total de características en la imagen izquierda. 1/ NL es el peso asociado con una característica coincidente. h es el número mínimo de características que se encuentran en cualquier imagen, es decir, $h = \min(NL, NR)$. D_q es la distancia mínima entre una entidad coincidente en las imágenes izquierda y derecha.

Se asocian las características de la imagen correcta, dentro de un área de búsqueda restringida, con el coeficiente de similitud más alto sobre un umbral. Estos se comparan globalmente con otras características asociadas para verificar la coherencia. La diferencia en la ubicación de estas características da la disparidad.

Tabla 1 - Resultado de la técnica de coincidencia de funciones estereó

Scene	Features	Features Matched	% Matched	Features Corrected	% Corrected	Time
Pentagon	13491	10928	81.0%	1223	11.2%	4h 9m 36s
Fruit	7525	6159	84.9%	655	10.6%	3h 37m 40s
Renault	2459	2093	85.1%	226	10.8%	1h 48m 20s
Stereogram	6800	6257	92.0%	0	0	15m 56s

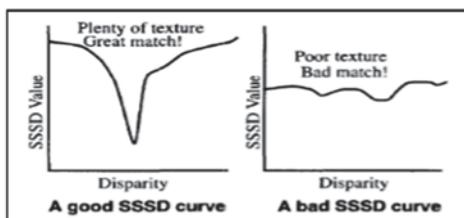
En el segundo método basado en áreas hay dos técnicas que utilizan este algoritmo. En ambos métodos, se coloca una ventana en una imagen. La otra imagen se escanea usando la misma ventana de tamaño. Los píxeles en cada ventana son comparados y operados. Estos se suman para dar un coeficiente para el píxel central. Estas técnicas han sido desarrolladas por (Okutomi et al., 1991).

La primera operación descrita es la correlación y coeficiente de correlación viene dado por:

$$C_{LR} = \sum_{[i,j] \in Window} L(i,j)R(i,j) \quad (3)$$

El segundo método usa el mismo principio de ventana, pero usa la suma de las diferencias al cuadrado (SSD). Esto examina los valores de píxeles en ambas ventanas y estima la disparidad calculando los coeficientes SSD. En este método, el coeficiente SSD debe minimizarse. La fórmula para SSD es:

$$SSD = \sum_{[i,j] \in Window} [L(i,j) - R(i,j)]^2 \quad (4)$$



Valores de SSD sobre el área de escaneo tomada de (Ross, 1993)

MÉTODO

OpenCV tiene una biblioteca predefinida que incluye un método para obtener mapas de disparidad para un par estéreo dado. (Munro y Gerdelan, 2006) utilizando esta biblioteca obtuvieron el siguiente resultado:



Mapa de disparidad

Después de probar muchas opciones diferentes, este es el mejor resultado obtenido. El mapa de disparidad muestra errores de rayas. Esto se atribuye al hecho de que OpenCV pierde saltos importantes de disparidad en los bordes de los objetos y, por lo tanto, asume las disparidades incorrectas en las etapas de búsqueda posteriores (Sunyoto et al., 2004).

CONCLUSIONES

La principal ventaja del algoritmo basado en características es su velocidad. El proceso de encontrar características en ambas imágenes y luego calcular la disparidad se puede hacer fácilmente en tiempo real. El método basado en áreas proporciona mapas densos de disparidad. Se pueden obtener otras mejoras procesando previamente las imágenes antes de implementar el algoritmo. (Zitnick y Kanade, 1999).

Si bien el método basado en características es rápido, la mayor desventaja es que produce mapas de disparidad dispersos. Incluso con pares estéreo ricos en funciones, este método no puede producir mapas tan detallados como el método basado en el área. Por esta razón, la mayoría de las implementaciones actuales de visión estéreo utilizan enfoques basados en áreas

REFERENCIAS

- Candocia, F. and M. Adjouadi: "A similarity measure for stereo feature matching". IEEE Transaction on Image Processing, Vol. 6, pp. 1460-1464, 1997.
- Foggia P., A. Limongiello, and M. Vento, "A real-time stereo-vision system for moving object and obstacle detection in AVG and AMR applications," IEEE Transactions on Image Process, Vol. 6, No. 10, October 1997
- Medioni G., and R. Nevatia, "Segment-based stereo matching," Comput. Vis., Graph., Image Process., vol. 31, pp. 2-18, July, 1985.
- Munro P. y Gerdelan A. Stereo Vision Computer Depth Perception (2006)
- Nasrabadi N. "A Stereo vision technique using curve segments and relaxation matching," IEEE Trans. Pattern Anal. Machine Intell., vol. 14 no. 5, pp. 566-572, May 1992
- Okutomi, Masatoshi and Kanade, Takeo (1991) A MultipleBaseline Stereo. CVPR proceedings, 1991
- Ross B, "A Practical Stereo Vision System" Computer Vision and Pattern Recognition, IEEE Computer Society Conference June 1993 Page(s):148 - 153.
- Sunyoto H, W van der Mark, and D. Gavrilu. "A Comparative Study of Fast Dense Stereo Vision Algorithms" IEEE Intelligent Vehicles Symposium June 2004
- Zitnick C. and T. Kanade, tech. A Cooperative Algorithm for Stereo Matching and Occlusion Detection report CMU-RI-TR-99-35, Robotics Institute, Carnegie Mellon University, October, 1999.