Sistema de posicionamiento global de un objeto en el espacio usando visión estéreo de alta resolución

Global positioning system of an object using high resolution stereo vision

D. Acevedo, J. Meneses^(*)

Escuela de Física, Facultad de Ciencias, Universidad Industrial de Santander, Cra. 27 Calle 9 Bucaramanga-Colombia.

^(*) Email: jaimen@uis.edu.co

Recibido / Received: 15/05/2012. Revisado / Revised: 27/07/2012. Aceptado / Accepted: 28/07/2012.

DOI: <http://dx.doi.org/10.7149/OPA.45.3.307>

RESUMEN:

En el presente trabajo se evalúa una estrategia basada en el procesamiento digital de imágenes para determinar el desplazamiento de un objeto en el espacio calibrado de un sistema de visión estéreo. Un patrón conformado por franjas alternas blancas y negras ubicadas en la periferia de un cono truncado define la mira cónica, que al ser adherida a la superficie del objeto permite determinar su trayectoria. Con el fin de emplear el método de visión estéreo, es necesario encontrar el vértice de la mira cónica en cada imagen. Se identifican las líneas presentes usando la transformada de Hough. Se recurre al algoritmo RANSAC para identificar las líneas y el punto de congruencia de la mira cónica, que corresponde al vértice. El procedimiento es empleado en un dispositivo de medida 3D a manos libres.

Palabras clave: Reconstrucción 3D, Sistema de Posicionamiento Global, Procesamiento Digital de Imágenes.

ABSTRACT:

This paper shows a strategy based on digital image processing to determine 3D displacement of an object in a calibrated space of a stereo vision system. An artificial target with alternating white and black fringes located on the periphery of a truncated cone defines a conical target, when attached to object surface allows determine its trajectory. In order to calculate the 3D position with stereo vision, the vertex of the conical target must be calculated. Hough transform is used in line detection procedure. RANSAC algorithm is implemented in order to identify lines belonging to conical target and the congruency point that define their vertex. This strategy is employed in a 3D hand-held reconstruction system.

Key words: 3D Reconstruction, Global Positioning System, Digital Image Processing.

REFERENCIAS Y ENLACES / REFERENCES AND LINKS

- [1]. S. S. Gorthi, G. Rajshekhar, P. Rastogi, "Three dimensional shape measurement using high-order instantaneous moments based fringe projection method", *Opt. Laser Technol.* **43**, 40-44 (2011).
- [2]. Y. Fujun, D. Meiling, H. Xiaoyuan, D. Xiaolei, "Single fringe projection profilometry based on sinusoidal intensity normalization and subpixel fitting", *Opt. Laser Eng.* **49**, 3, 465-472 (2011).
- [3]. M. Suodong, Q. Chenggen, J. T. Cho, Z. Rihong, L. Bo, "Phase error correction for digital fringe projection profilometry", *Physics Procedia* **19**, 227-232 (2011).
- [4]. A. Cigada, F. Mancosu, S. Manzoni, E. Zappa, "Laser triangulation device for in-line measurement of road texture at medium and high speed", *Mech. Syst. Signal Pr.* **24**, 2225-2234 (2010).
- [5]. D. Asahina, M. A. Taylor, "Geometry of irregular particles: Direct surface measurements by 3-D laser scanner", *Powder Technol.* **213**, 70-78 (2011).

- [6]. I. Brilakis, H. Fathi, A. Rashidi, "Progressive 3D reconstruction of infrastructure with videogrammetry", *Automat. Constr.* **20**, 884-895 (2011).
- [7]. J. Jiang, J. Cheng, X. Chen, "Registration for 3-D point cloud using angular-invariant feature", *Neurocomputing* **72**, 3839-3844 (2009).
- [8]. Z. Xie, S. Xu, X. Li, "A high-accuracy method for fine registration of overlapping point clouds", *Image Vision Comput.* **28**, 563-570 (2010).
- [9]. Q. Zhang, G. E. Sobelman, T. He, "Gradient-based target localization in robotic sensor networks", *Pervas. Mobile Comput.* **5**, 37-48 (2009).
- [10]. K. Doğançay, "Bearings-only target localization using total least squares", *Signal Process.* **85**, 1695-1710 (2005).
- [11]. G. Moriya, H. Chikatsu, "Practical target location and accuracy indicator in digital close range photogrammetry using consumer grade cameras", *Proc. SPIE* **8085**, 808503-1 (2011).
- [12]. R. O. Duda, P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures", *Commun. ACM* **15**, 11-15 (1972).
- [13]. M. Kalantari, F. Jung, "Détection entièrement automatique de points de fuite dans des scènes architecturales urbaines", *Revue XYZ* **107**, 41-46 (2006).
- [14]. M. A. Fischler, R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", *Commun. ACM* **24**, 381-395 (1981).

1. Introducción

En muchos procedimientos industriales y médicos se debe conocer con exactitud la posición, orientación y forma de un objeto con respecto a un sistema coordenado fijo. Aunque existe una gran variedad de dispositivos de medida que recuperan información tridimensional (3D), los métodos ópticos son ampliamente utilizados y adaptados a las condiciones industriales y médicas por sus propiedades no invasivas, resoluciones y campos de exploración ajustables a las necesidades del usuario y posibilidad de sistematizar la información. Por su facilidad de implementación, bajo costo e inmunidad a vibraciones y condiciones inhóspitas, métodos ópticos de triangulación láser, visión estéreo y proyección de franjas permiten generar dispositivos de medida de exploración 3D adaptado a diferentes entornos industriales. Sin embargo aunque los principios básicos remontan de varias décadas, aun continúan apareciendo en la literatura especializada artículos basados en estos métodos [1-6]. Para muchos grupos de investigación llama la atención los inconvenientes como: La dificultad de obtener información 3D de objetos con regiones ocultas al sistema de proyección u observación, la necesidad de obtener información a 360° de observación y la necesidad de medir con resolución pequeña para un campo de observación grande. Básicamente cuando se

desea recuperar información 3D de objetos extendidos con alta resolución se recurre al uso de un dispositivo de medida clásico de resolución adecuada, que tradicionalmente tiene campo de exploración pequeño, y se introducen desplazamientos relativos a valores conocidos del objeto con respecto al dispositivo o se emplean varios dispositivos de medida 3D ubicados en posiciones estratégicas o se usa un dispositivos de medida 3D y se desplaza a mano libre por el usuario. En este último caso se debe conocer el desplazamiento 3D realizado al dispositivo de medida con el fin de unificar a un único sistema coordenado global las diferentes nubes de datos 3D obtenidos en cada posición relativa del dispositivo.

Aunque el dispositivo de medida 3D ofrece una resolución adecuada a las condiciones de medida exigidas, la determinación del desplazamiento global realizado al dispositivo no exige alta resolución, algoritmos de registro de información 3D pueden ser utilizados para afinar la unificación de múltiples nubes de datos 3D [7,8]. Tradicionalmente los algoritmos para el registro de información 3D requieren de 2 nubes de datos de un mismo objeto, con puntos comunes y ubicados en posiciones relativas diferentes. El objetivo es acercar las dos nubes de datos minimizando la distancia entre puntos comunes a partir de un valor inicial de acercamiento entre las nubes. Este valor puede ser obtenido sin alta resolución empleando

estrategias de posicionamiento global [9,10], adhiriendo patrones artificiales, tradicionalmente planos, y empleando algoritmos para el reconocimiento automático diseñados en función de los rasgos geométricos característicos del patrón [11].

En este trabajo se evalúa una estrategia basada en el procesamiento digital de imágenes para determinar el desplazamiento de un objeto en el espacio calibrado de un sistema de visión estéreo. La estrategia de posicionamiento global emplea un patrón no plano conformado por una serie de franjas alternas blancas y negras ubicadas en la periferia de un cono truncado. Este patrón, llamado mira cónica, al ser adherido a un dispositivo de reconstrucción 3D y ubicado en el espacio calibrado de un sistema de visión estéreo permite la determinación del desplazamiento global introducido al dispositivo cuando se desplaza libremente alrededor del objeto que se desea reconstruir. La medida calculada del desplazamiento global es empleada como valor inicial en un algoritmo de registro con el fin de recuperar información 3D de un objeto a campo extendido y con regiones de difícil acceso. En este artículo se muestra el procesamiento digital implementado a una imagen de la mira cónica con el fin de determinar el vértice de la mira. La utilización de una mira cónica en un sistema de visión estéreo a múltiples cámaras permitirá la observación de la mira a cualquier ángulo de observación.

2. Esquema experimental

La Fig. 1 muestra una imagen de la mira cónica. El objetivo consiste en determinar la posición del vértice del cono en píxeles sobre una imagen. La mira cónica se ha construido usando mecanizado numérico con radios 20 mm y 9 mm y una altura de 35 mm. Las líneas que separan zonas blancas y negras convergen al vértice del cono truncado. La Fig. 2 muestra una imagen de la mira cónica en condiciones experimentales. El procedimiento digital de la imagen tiene las siguientes etapas: 1. Extracción de bordes, 2. Detección de líneas a partir de la transformada de Hough, 3. Identificación de líneas de la mira 4. Cálculo del vértice por mínimos cuadrados.

Fig. 1: Imagen de la mira cónica.

Fig. 2: Imagen experimental de la mira cónica.

3. Procesamiento digital de imágenes

3.a. Extracción de bordes

La Fig. 3 muestra la imagen binaria obtenida de aplicar el operador Prewitt para el cálculo aproximado del gradiente de la imagen. Sobre la figura se pueden identificar los segmentos de líneas que corresponden a las transiciones de las franjas negras y blancas. De igual forma se observan segmentos rectos y curvos de otras transiciones detectadas como bordes por el operador Prewitt, y que corresponde a ruido en la detección del vértice del cono.

3.b. Detección de líneas

Con el fin de determinar el vértice del cono, se emplea un algoritmo de detección de líneas a partir de la transformada de Hough. Las variables (ρ, θ) del espacio de Hough están asociadas a las coordenadas en píxeles (x, y) según la definición de la transformada de Hough [12]:

$$y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(-\frac{\rho}{\sin\theta}\right), \quad (1)$$

$$\rho = x \cos\theta + y \sin\theta. \quad (2)$$

Fig. 3: Imagen binaria de bordes obtenida aplicando el operador Prewitt.

Fig. 4: Variables empleadas en la transformada de Hough.

La Ec. (1) indica que para puntos (x, y) de la imagen ubicados en una línea recta con pendiente $-1/\tan(\theta)$ e intercepto $-\rho/\sin(\theta)$ corresponde al punto (ρ, θ) en el espacio de Hough. Y la Ec. (2) indica que un punto (x, y) en la imagen corresponde a una función sinusoidal en el espacio de Hough. Muestreando el espacio de Hough para valores de ρ entre $-D$ a $+D$ a intervalos de 0.5 píxeles, siendo D la diagonal de la imagen de 640×480 píxeles y para valores de θ entre $-\pi/2$ a $+\pi/2$ a intervalos de 0.01° , se obtiene una matriz (llamada acumulador) al evaluar la función sinusoidal para cada punto de la imagen de bordes. La superposición de varias curvas sinusoidales en un punto (ρ, θ) indica la presencia de una línea recta en la imagen. Es decir identificando los N picos o valores agudos más fuertes de la matriz acumulador se obtienen las N rectas más notorias de la imagen de bordes. Cada pico posee un valor de ρ y θ específicos, que al emplear la Ec. (1) para valores de x e y dentro de la imagen se obtiene la recta correspondiente. La Fig. 5 muestra las $N = 45$

Fig. 5: $N = 45$ líneas encontradas de la imagen de bordes a partir de la transformada de Hough.

rectas identificadas en la imagen de la Fig. 3. De la Fig. 5 se observa que se identifica visualmente el punto de congruencia de las líneas que provienen de la mira cónica, pero también se identifican más líneas que impiden calcular las coordenadas del vértice con precisión.

3.c. Cálculo del vértice de la mira cónica

Conociendo la pendiente y el intercepto de cada línea que corresponde a la mira cónica, se podrían estimar las coordenadas del punto de intersección como el punto que minimice la distancia a las líneas. Desafortunadamente la presencia de líneas que no pertenecen a la mira cónica afecta notoriamente el cálculo. Observando la Fig. 4 que relaciona (ρ, θ) con (x, y) , cada línea identificada posee un único punto de intersección H. Como se conocen los parámetros de la línea identificada (línea roja en la Fig. 4) y sabiendo que la línea OH es perpendicular a la línea identificada y pasa por el origen, se pueden emplear relaciones geométricas para calcular las coordenadas del punto de intersección H. La Fig. 6 muestra los 45 puntos H de las líneas identificadas.

Realizando un análisis geométrico sencillo se demuestra que los puntos de intersección H que corresponde a líneas congruentes definen un círculo que pasa por el origen, ver Fig. 7. El punto de congruencia de las líneas V en la Fig. 7 pertenece al círculo y une el centro con el origen. El ajuste por mínimos cuadrados a un círculo es utilizado en tratamiento digital de imágenes para determinar puntos de fuga en escenas

Fig. 6: Puntos de intercepción H encontrados de las líneas identificadas de la imagen de bordes.

Fig. 7: Círculo que define los puntos de intercepción H.

urbanas [13]. El problema se plantea de la siguiente manera: conociendo los puntos de intercepción H de la Fig. 6, ¿cómo interpolar el mejor círculo que pase por el origen teniendo en cuenta solamente puntos que corresponden a las líneas del cono? Si se realiza el ajuste por mínimos cuadrados a un círculo con la condición que pase por el origen, empleando todos los puntos H encontrados, afectaría fuertemente las líneas que no son congruentes. Matemáticamente esta situación se resuelve empleando el algoritmo RANSAC (por sus siglas en inglés *R*ANdom *S*amples *A*nd *C*onsensus).

El algoritmo RANSAC fue inicialmente propuesto por Fischler y Bolles [14] en 1981 como un método para estimar los parámetros de un modelo a partir de una serie de datos que son

Fig. 8: Clasificación de datos en *Inliers* (puntos azules) y *Outliers* (puntos rojos) empleados en el algoritmo RANSAC.

Fig. 9: Clasificación de datos en *Inliers* (puntos azules) y *Outliers* (puntos rojos) empleados en el algoritmo RANSAC.

contaminados con una gran cantidad de datos ruidosos llamados *Outliers*, indicados en rojo en la Fig. 8. Los datos aunque ruidosos que definen correctamente el modelo son llamados *Inliers*. La estadística de los datos *Inliers* determina correctamente los parámetros del modelo matemático y define un error umbral; valor que depende de la máxima desviación atribuida a efectos del ruido. Es decir, a partir del error umbral se pueden clasificar desde un punto de vista estadístico los datos como *Inliers* y *Outliers*, que es el objetivo del algoritmo RANSAC.

La Fig. 9 muestra el algoritmo RANSAC. S corresponde al menor número de datos empleados para calcular los parámetros $\theta(S)$ que definen el modelo. Así, para una línea recta $S = 2$ y para un círculo $S = 3$. La función de error se calcula como la distancia Eucladiana entre los puntos del modelo y el dato. Inicialmente se escogen aleatoriamente S datos y se evalúan los parámetros $\theta(S)$. Para todos los datos se calcula el error con los parámetros encontrados y se determina el número de datos con error inferior al error umbral. Esto define el número de *Inliers* para los parámetros $\theta(S)$. Si el número es el más grande encontrado se almacenan los parámetros $\theta(S)$. El algoritmo se repite escogiendo aleatoriamente otra grupo de S datos. El ciclo se repite hasta completar el máximo número de iteraciones, o hasta completar un porcentaje mínimo de *Inliers* definido por un valor umbral dependiendo de la estadística de los datos.

La adaptación del algoritmo RANSAC para el ajuste de un círculo usando los puntos de intercepción H se realiza de la siguiente manera: S se define como 2 y se emplea el punto de coordenadas $(0,0)$, sabiendo que el círculo pasa por el origen, para completar tres puntos que definen un círculo; El error umbral se define como 0.5 píxeles; el algoritmo termina al completar 100 iteraciones o al identificar máximo 16 puntos, hay 16 líneas en la mira cónica. El algoritmo reporta las coordenadas del centro (x_o, y_o) y radio R que minimiza el error cuadrático medio. La Fig. 10 muestra el círculo interpolado y el grupo de *Inliers* encontrado. Con el centro y el radio se encuentra la coordenada del punto de congruencia o vértice de la mira cónica.

El procedimiento anterior permitió encontrar un valor de coordenadas de (334.8, 163.4) píxeles para el vértice de la mira cónica mostrada en la Fig. 2. Una serie de imágenes de la mira desplazada en posiciones secuenciales permite generar un video donde se calcula la trayectoria del vértice del cono. La precisión en el cálculo de las coordenadas del vértice se puede determinar al evaluar el error que se introduce en la reconstrucción de objetos después de hacer el registro de nubes de datos 3D, situación que actualmente está en proceso de evaluación.

Fig. 10: Círculo interpolado empleando el algoritmo RANSAC

Fig. 11: Ubicación del vértice de la mira cónica encontrada para la Fig. 2.

4. Conclusión

En el presente trabajo se evalúa una estrategia basada en el procesamiento digital de imágenes para determinar la trayectoria de un objeto en el espacio 3D. Un patrón conformado por franjas alternas blancas y negras ubicadas en la periferia de un cono truncado define la mira cónica, que al ser adherida a la superficie del objeto permite determinar su trayectoria. Con el fin de emplear el método de visión estéreo, es necesario encontrar el vértice de la mira cónica en cada imagen. Inicialmente se determinan los bordes de la imagen empleando el operador Prewitt, y se identifican las líneas presentes usando la transformada de Hough. Se recurre al algoritmo RANSAC para identificar las líneas que pertenecen a la mira cónica. El procedimiento es empleado en la construcción de un dispositivo de medida 3D a manos libres, que no requiere de alta precisión en el cálculo de desplazamientos globales, si se emplea algoritmo de registro de

información 3D. Los resultados presentados corresponden a la validación del método propuesto. El paso siguiente en la investigación corresponde a la optimización de las dimensiones de la mira cónica para adherirla a un dispositivo de medida 3D y su correspondiente validación experimental del error que se introduce en la obtención de la medida.

Agradecimientos

Este trabajo ha sido financiado por la Vicerrectoría de Investigaciones y Extensión de la Universidad Industrial de Santander a través del proyecto N° 5184: Posicionamiento Global de alta resolución a campo extendido por visión estéreo: Aplicaciones en metrología óptica.