

# Segmentación de objetos parcialmente ocluidos utilizando cámaras TOF

Luciano Lorenti<sup>1</sup>, Cristina Manresa-Yee<sup>2</sup>, Javier Giacomantone<sup>1</sup>

<sup>1</sup> Instituto de Investigación en Informática (III-LIDI),  
Facultad de Informática - Universidad Nacional de La Plata - Argentina.  
{llorenti,jog}@lidi.info.unlp.edu.ar

<sup>2</sup> Universitat de les Illes Balears. Unidad de Gráficos, Visión por Computador e  
Inteligencia Artificial.  
Ed. Anselm Turmeda, Crta Valldemossa km 7.5, 07122 Palma, España.  
cristina.manresa@uib.es

**Resumen.** El presente artículo propone un método de segmentación para imágenes obtenidas con una cámara de tiempo de vuelo. El objetivo principal es segmentar objetos parcialmente ocluidos evitando que sean subdivididos como dos o más partes de objetos diferentes. El método propuesto explota la capacidad de la cámara utilizada para obtener simultáneamente imágenes de intensidad y de rango de la misma escena. Aplicando adecuadamente técnicas de detección de bordes en ambas imágenes, el método propuesto cierra adecuadamente el contorno de los objetos de la escena 3D. La estimación de las superficies se realiza minimizando la suma de los errores al cuadrado para finalmente estimar las superficies correspondientes a los mismos objetos. Se presentan resultados experimentales del método propuesto y se derivan conclusiones a partir de los mismos.

**Palabras clave:** Segmentación, Imágenes de Rango, Tiempo de Vuelo

## 1 Introducción

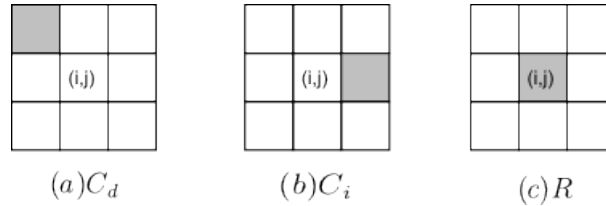
El objetivo de un método de segmentación aplicado a una imagen de intensidad 2D es dividir una imagen en sus partes constitutivas u objetos que la componen. La división depende del nivel de detalle requerido por el problema que se intenta resolver. Una imagen de intensidad 2D brinda información limitada con respecto a la escena 3D que contiene los objetos a segmentar. Algoritmos de visión por computador, en particular de segmentación, que han sido utilizados con éxito en ambientes industriales, con colores e iluminación controlada, no obtienen resultados similares en contextos diferentes y en condiciones de contorno naturales [1]. Una alternativa para abordar problemas en que las condiciones de contorno no permiten una segmentación adecuada es incorporar información de profundidad, la distancia de los distintos objetos que conforman la escena respecto a la cámara. Los principales métodos para obtener mediciones de rango de una

escena son, medición por barrido laser, medición con luz estructurada y técnicas estéreo mediante dos cámaras. Hay distintos tipos de escáners 3D que se pueden dividir entre activos y pasivos en función de la técnica de iluminación utilizada. En particular en este trabajo utilizamos una cámara de tiempo de vuelo, "Time of Flight" (TOF), que nos permite obtener imágenes de rango y de intensidad simultáneamente, la cámara utilizada es la MESA SR 4000 [2]. La SR 4000 es una cámara activa, utiliza su propia fuente de iluminación mediante una matriz de diodos emisores de luz infrarroja modulada en amplitud. Los sensores de la cámara detectan la luz reflejada en los objetos iluminados y la cámara genera dos imágenes. La imagen de intensidad es proporcional a la amplitud de la onda reflejada y la imagen de rango o distancia es generada a partir de la diferencia de fase entre la onda emitida y reflejada en cada elemento de la imagen [3][4]. Las principales ventajas con respecto a otras técnicas de medición 3D es la posibilidad de obtener imágenes a velocidades compatibles con aplicaciones en tiempo real y la posibilidad de obtener nubes de puntos 3D desde un solo punto de vista [5]. Técnicas de procesamiento de imágenes de intensidad han sido utilizadas en imágenes de rango para distintas aplicaciones [6][7]. Distintos métodos aplicando operadores derivativos se utilizan en imágenes de intensidad para detección de bordes [8]. En imágenes de intensidad como de rango afectadas por ruido es necesario considerar el efecto del ruido presente en la imagen al aplicar operadores basados en derivadas de primer y segundo orden. Entre los detectores de bordes clásicos mas utilizados, que consideran la influencia del ruido en el proceso de segmentación podemos mencionar el de Marr y Hildreth [9] y un método robusto como el algoritmo de Canny [10]. El primer detector de bordes mencionado utiliza la propiedad de cruce por cero del Laplaciano con un precursor Gaussiano para atenuar los efectos del ruido, en el segundo se minimiza la inclusión de falsos bordes, se minimiza la distancia entre los bordes reales y los detectados y se establece un criterio que integre respuestas múltiples que correspondan a un mismo borde. En las imágenes de intensidad generadas por una cámara de tiempo de vuelo predominan los detalles generados por la textura de los objetos y en las de rango los detalles y ruido asociados al fondo de la escena 3D. Recientemente han sido propuestas técnicas para segmentar objetos que operan sobre imágenes de rango e intensidad con el objetivo de definir bordes mas precisos en presencia, tanto de ruido como de oclusiones en distintos planos [11][12][13]. Estos métodos intentan determinar los bordes de los objetos en forma mas precisa, no siempre obteniendo contornos cerrados, y evitar que oclusiones parciales eviten una segmentación correcta. En este artículo se propone una solución integral para los dos problemas anteriores utilizando las dos imágenes generadas por la cámara TOF, la de intensidad y la de rango. El artículo está organizado del siguiente modo, en la sección 2,3 y 4 se presentan los fundamentos y el método propuesto. En la sección 5 se describen los resultados obtenidos. Finalmente en la sección 6 se presentan las conclusiones.

## 2 Extracción de contornos usando operaciones lógicas

Danciu, et ál. [13] propusieron un método mejorado de extracción de contornos en imágenes adquiridas con cámaras TOF, que combina información de las aristas obtenidas a partir de la imagen de distancia y de la imagen de intensidad. Para esto definen y utilizan una operación basada en la operación lógica AND, denominada VAND, que considera cada pixel y su vecindad para determinar la presencia de contornos. Sean  $D(x, y)$  e  $I(x, y)$  las imágenes de distancia e intensidad respectivamente, que provee la cámara TOF, ambas de dimensión  $n \times m$ . El método propuesto obtiene una imagen resultante  $R(x, y)$  de igual dimensión que contiene información mejorada de los contornos de los objetos con respecto a la que se podría obtener utilizando un algoritmo de detección de bordes a  $D(x, y)$  o  $I(x, y)$  por separado. El método es descrito a continuación:

1. Se aplica el método de Canny a las imágenes  $D(x, y)$  e  $I(x, y)$ .  
Sea  $B_d(x, y)$  la información de bordes extraída a partir de la imagen de distancia y sea  $B_i(x, y)$  la información de bordes extraída a partir de la imagen de intensidad.
2. Por cada pixel  $(p, k)$  se inspecciona su entorno delimitado por un círculo de radio 2 en  $B_i(x, y)$  y  $B_d(x, y)$  simultáneamente.  
Sea  $C_{B_d(p,k)}$  el entorno del pixel  $(p, k)$  de  $B_d(x, y)$  y  $C_{B_i(p,k)}$  el entorno del pixel  $(p, k)$  de  $B_i(x, y)$ .
3. Si al menos un pixel del contorno es encontrado en  $C_{B_d(p,k)}$  y en  $C_{B_i(p,k)}$  entonces el pixel  $(p, k)$  en  $R(x, y)$  es considerado como un pixel de contorno.



**Fig. 1.** Operación VAND aplicada a los vecinos 8 de  $(i, j)$  a) contornos de la imagen de distancia, b) contornos de la imagen de intensidad, c) imagen resultante

## 3 Estimación de un plano a partir de datos de distancia

Las nubes de puntos obtenidas a partir de los sensores de distancia representan muestras relativamente ruidosas de superficies que existen en el mundo real. Es posible conocer la orientación de estas superficies utilizando métodos de minimización del error [16]. Un conjunto de puntos no correctamente segmentado puede incluir la superficie plana deseada pero también superficies no

pertenecientes al plano que causarán errores residuales en la estimación. Una expresión general de un plano en un espacio cartesiano de tres dimensiones está dada por

$$\tilde{A}x + \tilde{B}y + \tilde{C}z = \tilde{E} \quad (1)$$

y normalizada por la siguiente ecuación

$$Ax + By + Cz = 1 \quad (2)$$

Dado el conjunto de muestras  $\{(x_i, y_i, z_i)\}_{i=1}^m$  se calcula  $A$ ,  $B$  y  $C$  para determinar el plano,  $Ax + By + Cz = 1$ , que minimice la suma de los errores al cuadrado. Se define a la suma de los errores cuadráticos medios (SSE) como:

$$SSE = E(A, B, C) = \sum_{i=1}^m [(Ax_i + By_i + Cz_i) - 1]^2 \quad (3)$$

Esta función no toma valores negativos y su gráfica es un hiperparaboloide cuyo vértice ocurre cuando el gradiente satisface  $\vec{\nabla} \mathbf{E} = (0, 0, 0)$ . Esto lleva a un sistema de 3 ecuaciones lineales que puede ser resueltas fácilmente.

$$\vec{\nabla} \mathbf{E} = (0, 0, 0) \quad (4)$$

$$2 \sum_{i=1}^m [(Ax_i + By_i + Cz_i) - 1](x_i, y_i, z_i) = (0, 0, 0) \quad (5)$$

y por lo tanto

$$\begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i y_i & \sum_{i=1}^m x_i z_i \\ \sum_{i=1}^m x_i y_i & \sum_{i=1}^m y_i^2 & \sum_{i=1}^m y_i z_i \\ \sum_{i=1}^m x_i z_i & \sum_{i=1}^m y_i z_i & \sum_{i=1}^m z_i^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i \\ \sum_{i=1}^m y_i \\ \sum_{i=1}^m z_i \end{bmatrix}$$

La solución expresa la estimación de mínimos cuadrados de  $1 = Ax + By + Cz$ .

**Vector Normal** Los parámetros  $A$ ,  $B$ ,  $C$  y  $D$  pueden ser relacionados al vector normal  $\vec{\mathbf{n}}$  y a un punto de referencia  $\vec{\mathbf{p}}_0 = (x_0, y_0, z_0)$  considerando que cualquier vector del plano es perpendicular al vector normal, es decir:

$$\vec{\mathbf{n}}(\vec{\mathbf{x}} - \vec{\mathbf{p}}_0) = 0 \quad (6)$$

$$\vec{\mathbf{n}} \vec{\mathbf{x}} - \vec{\mathbf{n}} \vec{\mathbf{p}}_0 = 0 \quad (7)$$

$$\vec{\mathbf{n}} \vec{\mathbf{x}} = \vec{\mathbf{n}} \vec{\mathbf{p}}_0 \quad (8)$$

$$n_x x + n_y y + n_z z = \tilde{E} \quad (9)$$

Donde  $\tilde{E}$  es ahora la distancia desde el origen a la superficie plana.

## 4 El método propuesto

El método propuesto se divide en dos etapas principales: La primera consiste en extraer correctamente las fronteras de los objetos capturados y la segunda en determinar cuáles de esos conjuntos de puntos delimitados por los contornos eran originalmente un solo objeto pero fueron segmentados como dos debido a una oclusión parcial.

### Extracción de Bordes Inicial

1. Se utiliza el método descrito en la sección 2 para obtener una imagen de bordes  $R(x, y)$ .

### Clausura de Contornos

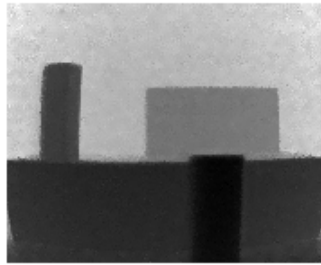
2.  $NR = NOT(R(x, y))$
3. Utilizando  $NR$  se localizan los componentes conectados  $W_i$  en la imagen de rango.
4. Por cada componente conectado  $W_i$  se estima el plano que mejor describe a ese conjunto y se calcula el error cuadrático medio  $SSE_{W_i}$  como se describió en la sección 3.
5. Si  $\sum_{i=1}^n SSE_{W_i} > \text{umbral}$  entonces se dilatan los bordes de  $R$  y se continúa en 2, sino se continúa en 6.

### Análisis de Objetos Ocluidos

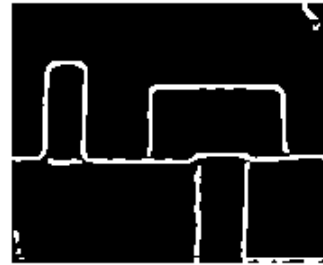
6. Se calcula para cada conjunto  $W_i$ 
  - La media de intensidad.
  - La media de distancia.
  - El centroide de la nube de puntos.y se estima
  - El vector normal del plano que mejor que describe a la superficie utilizando el método descrito en la sección 3.
7. Se comparan los componentes conectados entre sí buscando la similitud entre las características calculadas en el paso anterior. Si dos conjuntos resultan similares se unen.

## 5 Resultados Experimentales

El método propuesto ha sido evaluado en tres escenas 3D constituidas por objetos geométricos simples. Para cada escena se realizaron cuatro experimentos variando la posición relativa, la distancia con respecto a la cámara y el grado de oclusión del objeto a segmentar. En todos los casos se utilizó la cámara TOF SwissRanger SR4000.



(a) Imagen de Intensidad



(b) Extracción de Bordos Inicial



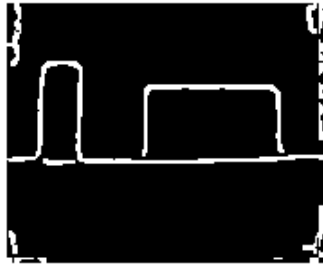
(c) Contornos Clausurados



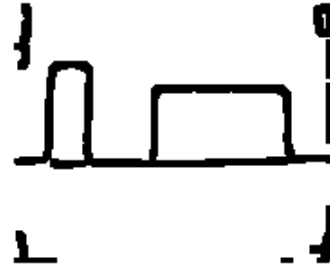
(d) Imagen Segmentada

**Fig. 2.** Segmentación utilizando el método propuesto

En la figura 2 se pueden observar las imágenes obtenidas para una escena y configuración específica, en cada etapa del método propuesto. La imagen en 2(a) es la imagen de intensidad y 2(b) muestra los bordes extraídos mediante el método descrito en la sección 2. En la figura 2(c) se muestran las fronteras de los objetos correctamente delimitados y en la figura 2(d) se observa la segmentación de los objetos, identificando al objeto parcialmente ocluido como uno, sin dividirlo en dos objetos diferentes.



(a) Extracción de Bordes



(b) Contornos Cerrados

**Fig. 3.** Clausura de Contornos

En figura 3 se puede observar como el método propuesto mejora la definición de la frontera de los objetos, permitiendo la posterior segmentación adecuada de objetos parcialmente ocluidos. En la figura 3(a) solo se aplica el método descrito en la sección 2. En ella se pueden apreciar objetos no correctamente delimitados debido a la presencia de irregularidades en las fronteras de los mismos. La imagen en la figura 3(b) muestra los bordes correctamente cerrados utilizando el método propuesto.



(a) Componentes Conectados



(b) Método Propuesto

**Fig. 4.** Segmentación de Objetos Ocluidos

La figura 4 muestra dos imágenes que comparan la segmentación de los objetos delimitados por los contornos basándose solamente en los componentes conectados de a 8 en la figura 4(a) y los objetos segmentados utilizando el método propuesto en la figura 4(b). El objeto parcialmente ocluido en la parte inferior de la imagen es segmentado como uno solo puesto que su valor de intensidad, su vector normal y su posición en el espacio se encuentra dentro de un umbral determinado.

## 6 Conclusiones

En este artículo se presenta un método para segmentar objetos parcialmente ocluidos utilizando imágenes de rango y de intensidad obtenidas con cámaras de tiempo de vuelo. Los resultados preliminares obtenidos permiten observar el cierre correcto de los contornos de los objetos y la segmentación de objetos ocluidos de características simples. Trabajos de investigación futuros permitirán mejorar el método y así poder segmentar adecuadamente superficies curvas y objetos ocluidos con direcciones variables.

## 7 Agradecimientos

Esta investigación ha sido parcialmente financiada por el proyecto A1/037910/11 Formación de Recursos Humanos e Investigación en el Área de Visión por Computador e Informática Gráfica, del Programa de Cooperación Interuniversitaria e Investigación Científica entre España e Iberoamérica del MAEC-AECID.

## Referencias

1. A. A. Dorrington, C. D. Kelly, S. H. McClure, A. D. Payne, M. J. Cree. Advantages of 3D Time of Flight Range Imaging Cameras in Machine Vision Applications. 16<sup>th</sup> Electronics New Zealand. Dunedin, New Zealand, 95-99, 2009.
2. M. Cazorla, D. Viejo, C. Pomares. Study of the SR 4000 camera. XI Workshop de Agentes Físicos, Valencia, 2010.
3. N. Blanc, T. Oggier, G. Gruener, J. Weingarten, A. Codourey, P. Seitz. Miniaturized smart cameras for 3D imaging in real time. IEEE Sensors, Vienna, Austria, 471-474, 2004.
4. J. Mure-Dubois, H. Hugli. Real-Time Sattering compensation for Time of Flighth cameras. International Conference of Vision Systems, 117-122, 2007.
5. F. Chiabrando, D. Piatti, F. Rinaudo. SR-4000 TOF Camera: Further Experimental Tests and First Applications to Metric Surveys. V Symposium on Remote Sensing and Spatial Information Sciences, Newcastle, UK, 38(5):149-154, 2010.
6. R. Tanner, M. Studer, A. Zanoli, A. Hartmann. People detection and tracking with tof sensor. IEEE Conference on Advanced Video and Signal Based Surveillance, 356-361, 2008.
7. S. Caraian, N. Kirchner. Robust Manipulalability-Centric Object Detection in Time-of-Flighth Camera Point Clouds. Australian Conference on Robotics and Automation, 1-9, Brisbane, Australia, 2010.



8. D. Ziou, S. Tabbone. Edge Detection Techniques - An Overview. *International Journal of Pattern Recognition and Image Analysis*, 8:537-559, 1998.
9. D. Marr, E. Hildreth. Theory of Edge Detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 207 (1167):187-217, 1980.
10. J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8 (6):679-698, 1986.
11. R. Benlamri. Range image segmentation of scenes with occluded curve objects. *Pattern Recognition Letters*, 21: 1051-1060, 2000.
12. S. Oprisescu, C. Burlacu, A. Sultana. A new contour extraction algorithm for ToF images. *10<sup>th</sup> International Symposium on Signals, Circuits and Systems (ISSCS)*, Bucharest, Romania, 1-4, 2011.
13. G. Danciu, M. Ivannovici, V. Buzuloiu. Improved Contours for ToF Cameras based on Vicinity Logic Operations. *12<sup>th</sup> International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, 989-992, 2010.
14. V. Don, H. Maarten Uijt. Near real-time extraction of planar features from 3d flash-ladar video frames. *SPIE: Optical Pattern Recognition*, 6977:69770N-69770N-12, 2008.
15. N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9 (1):61-66, 1979.
16. K. Klasing, D. Althoff, D. Wollherr, M. Buss. Comparison of surface normal estimation methods for range sensing applications. *IEEE International Conference on Robotics and Automation (ICRA)*, 3206-3211, 2009.