

Segmentación espectral de imágenes utilizando cámaras de tiempo de vuelo

Luciano Lorenti, Javier Giacomantone

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

Resumen. En este artículo se presenta un método de segmentación aplicable a imágenes adquiridas con cámaras de tiempo de vuelo (TOF). Las cámaras TOF generan dos imágenes simultáneas, una de intensidad y una de rango. El método propuesto modela ambas imágenes mediante una matriz de afinidad independiente para cada imagen. Transformando la imagen de rango y utilizando el criterio de cortes normalizados se optimiza la segmentación de los objetos de interés de la escena. El método mejora la segmentación en escenas donde la información de intensidad o de rango es insuficiente para obtener una separación adecuada de los objetos de interés. Se presentan resultados experimentales del método propuesto sobre imágenes simuladas e imágenes reales adquiridas por una cámara específica, y se derivan conclusiones a partir de los mismos.

Palabras clave: Segmentación, Imágenes de Rango, Cámaras de Tiempo de Vuelo, Agrupamiento Espectral

1 Introducción

El objetivo de un método de segmentación 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. Cuando una imagen de intensidad 2D brinda información limitada con respecto a la escena 3D que contiene los objetos a segmentar una alternativa posible es incorporar información de profundidad, la distancia de los distintos objetos que conforman la escena respecto a la cámara. 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 [1]. 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

[2][3]. 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[4][5]. Han sido utilizadas técnicas clásicas de segmentación directamente sobre imágenes de rango considerando distintas condiciones de ruido [6][7][8], [9] en distinto tipo de aplicaciones [10][11]. 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 [12][13][14]. Han sido utilizados diversos algoritmos de agrupamiento tanto jerárquicos como particionales para abordar el problema de segmentar objetos en imágenes de intensidad. En particular el método propuesto, en primer término, asocia el problema de segmentación en el dominio de la imagen con la partición de un grafo. En segundo término optimiza el corte del grafo formulando la minimización necesaria para determinar el corte como un problema de autovalores generalizado [15][16][17]. Finalmente el método combina los resultados obtenidos del procesamiento de la imagen de intensidad y de la de rango mejorada [18], para obtener la segmentación final de los objetos de interés de la escena. La evaluación del método propuesto se realiza mediante dos tipos de datos, imágenes simuladas con todos los parámetros y factores de ruido controlados e imágenes reales adquiridas por la cámara SR 4000 en condiciones de operación real.

El artículo está organizado del siguiente modo, en la sección 2 se describe el método de partición mediante cortes normalizados y en la sección 3 se expone el método propuesto. En la sección 4 se presentan los resultados experimentales obtenidos. Finalmente en la sección 5 se presentan las conclusiones.

2 Cortes Normalizados en Segmentación de Imágenes

2.1 Particionado de Grafos

Un grafo $G=(V,E)$ está formado por un conjunto de vértices V y un conjunto de aristas E que relacionan elementos de V . La teoría de grafos es usada, por lo general, en el modelado de problemas como tráfico de redes, circuitos eléctricos y redes de internet.

Con el objetivo de construir grafos a partir de imágenes, los vértices son generados a partir de los pixeles que la constituyen. Como cada uno de los elementos de la imagen contiene información de intensidad y posición, agrupar los pixeles de acuerdo a su semejanza y desemejanza puede lograr una correcta segmentación de la imagen. El conjunto de las aristas E está constituido por elementos que denotan la semejanza y desemejanza entre los pixeles. Como paso previo a la segmentación, es necesario construir un grafo pesado asignando a cada arista del conjunto E un peso $w(i, j)$, que resulta de evaluar la semejanza entre el pixel i y el pixel j . El valor de $w(i, j)$ aumenta con el grado de semejanza entre el pixel i y el pixel j . Denominamos corte de un grafo a la partición del mismo que se consigue removiendo las aristas de menor peso. Un arista con peso pequeño

indica un bajo grado de semejanza entre los pixeles que conecta. Por lo tanto para particionar el grafo en dos sub-grafos, el corte mínimo esta dado por:

$$cut(A, B) = \sum_{i \in A, j \in B} w(i, j)$$

donde A y B son dos subgrafos tales que $A \cup B = V$, $A \neq \emptyset$, $B \neq \emptyset$ y $A \cap B = \emptyset$. El criterio de corte mínimo, que fue introducido por Wu y Leahy[15], minimiza los posibles cortes máximos a través de los subgrafos. Para segmentar correctamente una imagen, se debe buscar el valor mínimo de corte entre todos los posibles subconjuntos del grafo, removiendo las aristas que involucren al corte mínimo y luego repetir el proceso de forma recursiva hasta que el valor de corte no supere cierto umbral.

2.2 Cortes Normalizados

La segmentación de imágenes utilizando el criterio de corte anterior otorga buenos resultados, pero favorece la aparición de sub-grafos de pocos pixeles aislados. Shi y Malik [16] propusieron un nuevo criterio, llamado cortes normalizados, que evita la aparición de estos conjuntos aislados. Definieron el criterio propuesto de la siguiente forma:

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)}$$

donde $assoc(A, V) = \sum_{i \in A, j \in B} w(i, j)$ y $assoc(B, V)$ se define de manera análoga.

El criterio de cortes normalizados, a diferencia del criterio de corte mínimo, no considera la sumatoria de los pesos de las aristas que conectan los dos conjuntos sino que considera la proporción que esas aristas representan con respecto a la suma de los pesos de todas las aristas de los nodos del sub-grafo. Shi y Malik mostraron también que existe una relación entre el grado de asociación y disociación de los conjuntos, como se puede ver en la siguiente expresión:

$$\begin{aligned} Ncut(A, B) &= \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \\ &= \frac{assoc(A, V) - assoc(A, A)}{assoc(A, V)} + \frac{assoc(B, V) - assoc(B, B)}{assoc(B, V)} \\ &= 2 - \left(\frac{assoc(A, A)}{assoc(A, V)} + \frac{assoc(B, B)}{assoc(B, V)} \right) \\ &= 2 - Nassoc(A, B) \end{aligned}$$

Por lo tanto minimizar la disociación entre las particiones es equivalente a maximizar la asociación entre ellas.

2.3 Resolviendo los cortes normalizados como un problema de autovectores

Una de las ventajas más importantes para usar el criterio de cortes normalizados es que se puede obtener una buena aproximación de la partición óptima de forma muy eficiente.

Sea $W_{ij} = w(v_i, v_j)$ la matriz de pesos del grafo y sea D la matriz diagonal de forma que $D_{ii} = \text{grado}(v_i) = \sum_{v_j \in V} w(v_i, v_j)$

Shi y Malik demostraron que una partición óptima se puede obtener calculando:

$$y = \arg \min Ncut = \arg \min_y \frac{y^T (D - W)y}{y^T D y}$$

donde y es un vector indicador binario que especifica a que grupo pertenece cada pixel.

Calcular el vector indicador es un problema NP-Completo para algunos tipos de grafos particulares, pero eliminando la restricción que y sea un vector indicador y resolviendo el problema en el dominio real, se puede hallar una aproximación a la solución discreta eficientemente. La ecuación anterior puede ser optimizada resolviendo el sistema de autovalores generalizado:

$$(D - W)y = \lambda D y$$

Shi y Malik demostraron que el segundo autovector de este sistema es la solución real del problema de cortes normalizados.

2.4 Ncut simultáneo de k-vías

Shi y Malik definieron el corte simultáneo de k-vías que da como resultado k segmentos en una sola iteración de la siguiente forma:

$$Ncut(A_1, A_2, \dots, A_k) = \frac{cut(A_1, A_1)}{assoc(A_1, V)} + \frac{cut(A_2, A_2)}{assoc(A_2, V)} + \dots + \frac{cut(A_n, A_n)}{assoc(A_n, V)}$$

Dado un vector indicador v de un sub-grafo A_j tal que

$$v_i = \begin{cases} \frac{1}{\sqrt{assoc(A_n, V)}} & \text{si } i \in A_j \\ 0 & \text{si } i \notin A_j \end{cases}$$

resulta que

$$v_i^T L v_i = \frac{cut(A_j, A_j)}{assoc(A_j, V)}$$

Sea H la matriz formada por k vectores indicadores puestos en columnas, minimizar $Ncut(A_1, A_2, \dots, A_k)$ es equivalente a minimizar:

$$\min_{A_1, A_2, \dots, A_k} \text{Tr}(H^T L H) \text{ sujeto a } H^T D H = I$$

Relajando el carácter discreto de la restricción y substituyendo $P = D^{\frac{1}{2}} H$ se obtiene el siguiente problema

$$\min_{P \in R^{x \times k}} \text{Tr}(P^T D^{-1/2} L D^{-1/2} P) \text{ sujeto a } P^T P = I$$

Este problema de minimización de traza estándar es resuelto por la matriz P cuando esta contiene los primeros k autovectores, puestos en columnas, correspondientes a los autovalores mas pequeños de la matriz $LN = D^{-1/2} L D^{-1/2}$, llamada matriz laplaciana normalizada [19]. Resustituyendo $H = D^{-1/2} P$, se puede ver que estos autovectores son los autovectores generalizados correspondientes a los autovalores mas pequeños de $(D - W)u = \lambda D u$.

3 Método Propuesto

Sea $I(i)$ una imagen de intensidad y $R(i)$ una imagen de rango, ambas de dimensión $n \times m$.

1. A partir de la imagen $R(i)$ se genera una imagen de rango mejorada $NERI(i)$ [18]
2. (a) Se construye la matriz de afinidad W_I de forma que

$$W_I(i, j) = e^{\frac{-\|F(i)-F(j)\|_2}{\alpha_I}} * \begin{cases} e^{\frac{-\|X(i)-X(j)\|_2}{\alpha_X}} & \text{si } \|X(i) - X(j)\|_2 < r \\ 0 & \text{c. c.} \end{cases}$$

donde $X(i)$ es la locación espacial del nodo i y $F(i) = I(i)$.

- (b) Se construye la matriz de afinidad W_R de forma que

$$W_R(i, j) = e^{\frac{-\|F(i)-F(j)\|_2}{\alpha_I}} * \begin{cases} e^{\frac{-\|X(i)-X(j)\|_2}{\alpha_X}} & \text{si } \|X(i) - X(j)\|_2 < r \\ 0 & \text{c. c.} \end{cases}$$

donde $X(i)$ es la locación espacial del nodo i y $F(i) = NERI(i)$.

3. (a) Se calcula la matriz laplaciana normalizada asociada a W_I

$$L_{NI} = D_I^{-1/2} (D_I - W_I) D_I^{-1/2}$$

- (b) Se calcula la matriz laplaciana normalizada asociada a W_R

$$L_{NR} = D_R^{-1/2} (D_R - W_R) D_R^{-1/2}$$

4. Se genera la matriz $H_I \in R^{k \times n}$ que contiene como columnas primeros k autovectores del sistema $L_{NI}u = \lambda u$. Se genera la matriz $H_R \in R^{l \times n}$ que contiene como columnas los primeros l autovectores del sistema $L_{NR}u = \lambda u$.
5. Se obtienen los autovectores correspondientes a las matrices laplacianas no normalizadas. $PI = D_I^{-1/2}H_I$ y $PR = D_R^{-1/2}H_R$
6. Se forma la matriz P_{IR} concatenando las matrices PI y PR
7. Para $i = 1, \dots, n$, sea $y_i \in R^m$ los vectores correspondientes a la i -ésima fila de P_{IR} , se segmentan los puntos $y_i \in R^m$ con algoritmo k-medias en clusters C_1, \dots, C_k .

4 Resultados Experimentales

Se presentan resultados experimentales del método propuesto aplicado a imágenes artificiales y a imágenes reales. Las capturas reales fueron obtenidas utilizando la cámara de tiempo de vuelo MESA SwissRanger SR4000 [1] y las capturas reales fueron simuladas utilizando el software Blensor [20].

En la figura 1 se puede visualizar el resultado de aplicar el algoritmo propuesto sobre una captura artificial.

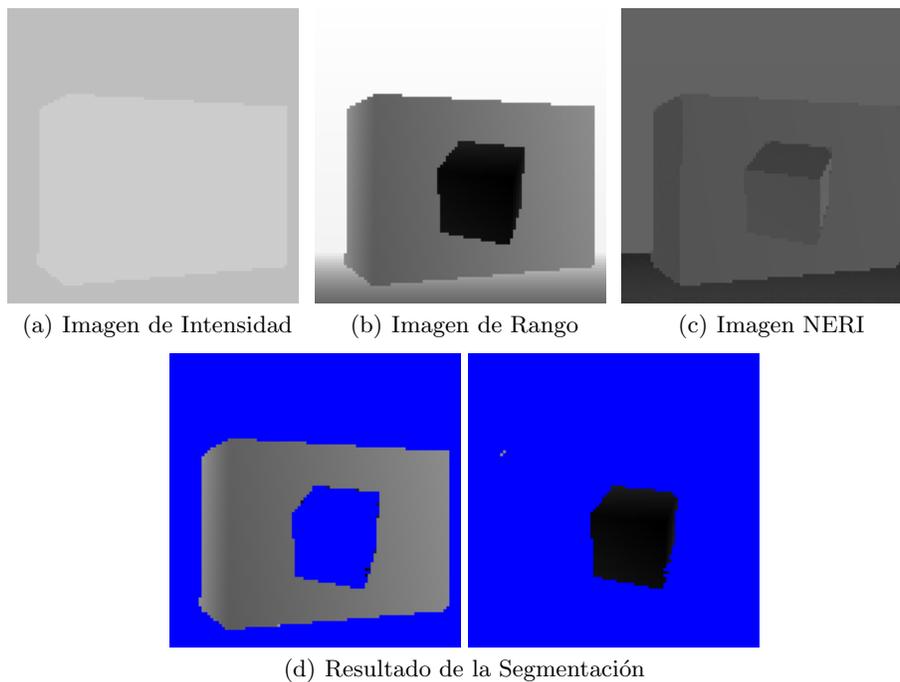


Fig. 1. Segmentación utilizando el método propuesto

La imagen 1(a) muestra dos objetos de niveles de intensidad iguales y ambos con niveles de intensidad próximos al del fondo. En la imagen 1(b) se puede observar que la información de distancia es útil para realizar una segmentación correcta. La imagen 1(c) muestra como la imagen de rango mejorada, incorpora información sobre las orientaciones de los objetos, conservando la información de distancia. Los parámetros utilizados para generar las matrices de afinidad y la cantidad de autovectores utilizados para la segmentación están detallados en la Tabla 1. Aplicando k-medias sobre el espacio generado por los autovectores, el algoritmo permite extraer correctamente las partes constitutivas de la imagen como se puede observar en las imágenes 1(d).

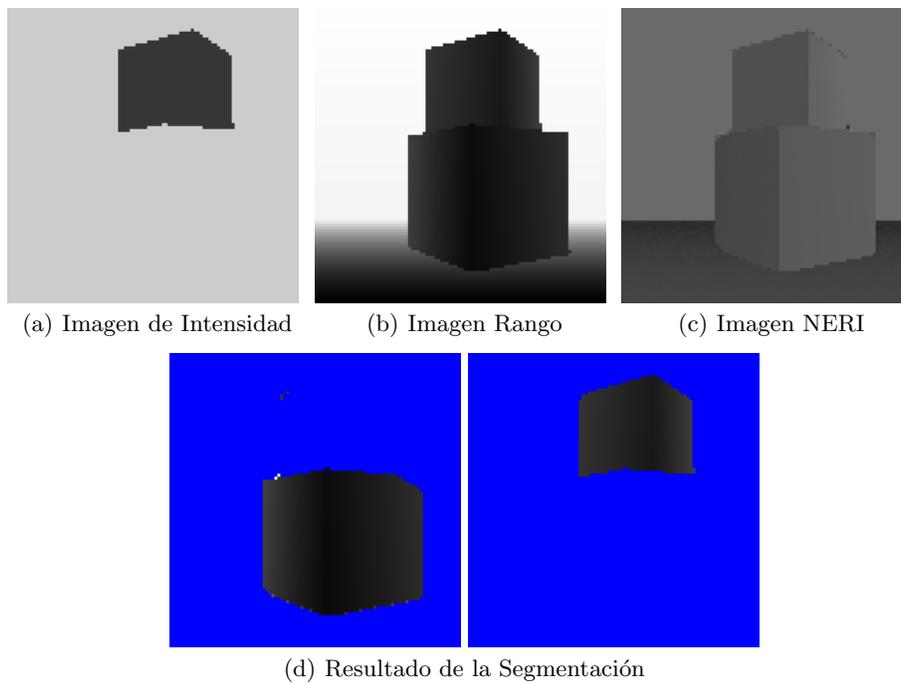


Fig. 2. Segmentación utilizando el método propuesto.

En la figura 2 se puede visualizar el resultado de aplicar el algoritmo propuesto sobre otra captura artificial. La imagen 2(a) muestra como en la imagen de intensidad el segundo objeto tiene niveles de intensidades muy próximos a los del fondo y en la imagen 2(b) como los dos objetos que integran la imagen son difíciles de separar uno del otro. Al utilizar los parámetros especificados en la Tabla 1, el algoritmo puede combinar correctamente la información de las dos imágenes. Utilizando correctamente los datos de distancia para separar el fondo del objeto situado en la parte inferior de la imagen, y los datos de intensidad

para separar un objeto del otro como presenta la figura 2(d).

La figura 3 muestra el resultado de aplicar el algoritmo propuesto sobre una captura real. La imagen de amplitud 3(a) presenta 3 objetos sobre un fondo negro, todos a la misma distancia. Uno de los objetos tiene un nivel de intensidad similar al del fondo, lo que dificulta su segmentación. En la imagen de rango 3(b) los objetos se distinguen claramente del fondo pero no uno del otro. El método combina correctamente la información de ambas imágenes ruidosas para segmentar los tres objetos presentes en la escena, como se muestra en la figura 3.

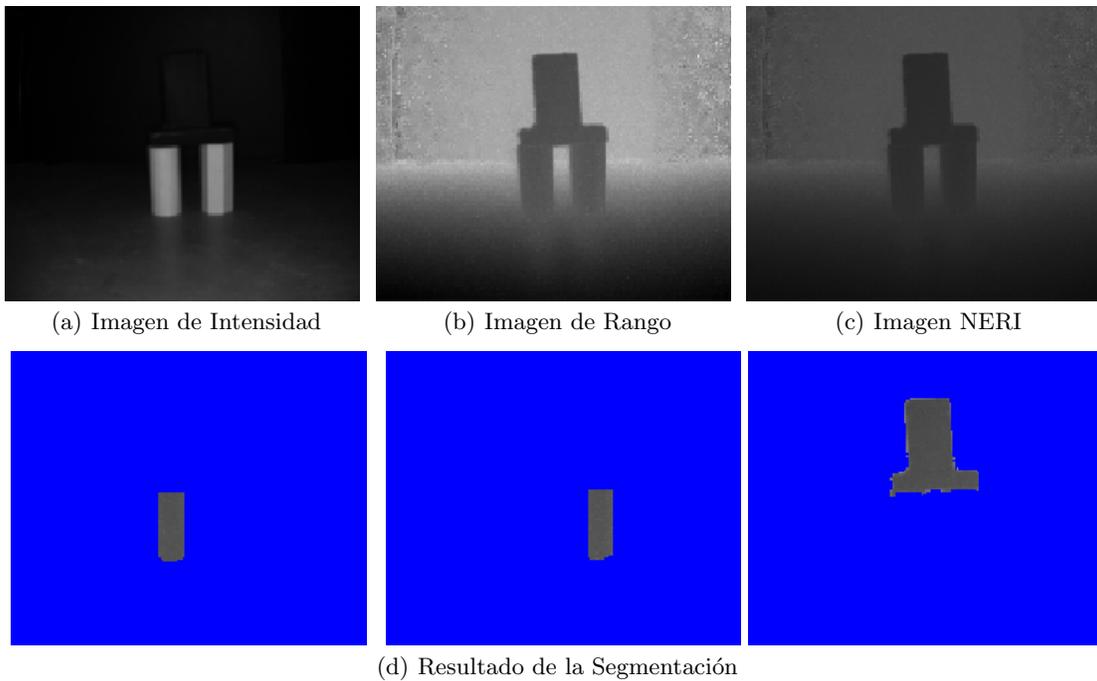


Fig. 3. Segmentación utilizando el método propuesto sobre una captura real

	Imagen	Autovectores	r	α_I	α_X
Prueba 1	Intensidad	3	3	0.5	4
	Rango	3	3	2	4
Prueba 2	Intensidad	2	4	0.2	3
	Rango	2	4	12	2
Prueba 3	Intensidad	2	4	2	4
	Rango	3	4	190	2

Tabla 1. Parámetros

5 Conclusiones

En este artículo presentamos un método de agrupamiento aplicado a la segmentación de imágenes de rango. La descripción formal del método se plantea utilizando herramientas conocidas de álgebra lineal, simplificando así su implementación. Los resultados obtenidos tanto sobre imágenes de intensidad y rango simuladas como reales presentan resultados preliminares satisfactorios. El algoritmo combina adecuadamente la información provista por ambas imágenes incluso en presencia de ruido. Permite segmentar objetos con niveles de intensidad próximos, ubicados a distintas distancias, u objetos cercanos con niveles de intensidad diferentes o con orientaciones diferentes. Como trabajo futuro se propone un análisis detallado de la influencia de los parámetros de la función de pesos $w(i, j)$ en los autovectores de las matrices laplacianas. Otro aspecto para un trabajo de investigación futuro es modificar el método propuesto combinando la información de intensidad y profundidad en la función de pesos $w(i, j)$.

Referencias

1. M. Cazorla, D. Viejo, C. Pomares. Study of the SR 4000 camera. XI Workshop de Agentes Físicos, Valencia, 2010.
2. 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.
3. J. Mure-Dubois, H. Hugli. Real-Time Sattering compensation for Time of Flight cameras. International Conference of Vision Systems, 117-122, 2007.

4. 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. 16th Electronics New Zealand. Dunedin, New Zealand, 95-99, 2009.
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. D. Ziou, S. Tabbone. Edge Detection Techniques - An Overview. International Journal of Pattern Recognition and Image Analysis, 8:537-559, 1998.
7. N. Otsu. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man and Cybernetics, 9 (1):61-66, 1979.
8. D. Marr, E. Hildreth. Theory of Edge Detection. Proceedings of the Royal Society of London. Series B, Biological Sciences, 207 (1167):187-217, 1980.
9. J. Canny. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (6):679-698, 1986.
10. 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.
11. S. Caraian, N. Kirchner. Robust Manipulability-Centric Object Detection in Time-of-Flight Camera Point Clouds. Australian Conference on Robotics and Automation, 1-9, Brisbane, Australia, 2010.
12. R. Benlamri. Range image segmentation of scenes with occluded curve objects. Pattern Recognition Letters, 21: 1051-1060, 2000.
13. S. Oprisescu, C. Burlacu, A. Sultana. A new contour extraction algorithm for ToF images. 10th International Symposium on Signals, Circuits and Systems (ISSCS), Bucharest, Romania, 1-4, 2011.
14. G. Danciu, M. Ivannovici, V. Buzuloiu. Improved Contours for ToF Cameras based on Vicinity Logic Operations. 12th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), 989-992, 2010.
15. Z. Wu, R. Leahy, An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.15, no.11, 1101-1113, 1993.
16. J. Shi, J. Malik, Normalized cuts and image segmentation. Proceedings of 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 731 - 737, 1997.
17. J. Shi, J. Malik, Normalized Cuts and Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, 888-905, 2000.
18. K. Pulli; M. Pietikainen. "Range image segmentation based on decomposition of surface normals". Proceedings of 8th Scandinavian Conference on Image Analysis. pp. 893-899, 1993.
19. R.K. Chung. Spectral Graph Theory. Am. Math. Soc, 1997.
20. M. Gschwandtner; R. Kwitt; A. Uhl, BlenSor: Blender Sensor Simulation Toolbox. Proceedings of 7th International Symposium In Advances in Visual Computing. ISVC 2011, 2011.