Fusión de información de geometría e intensidad para segmentación de imágenes TOF

Luciano Lorenti, Javier Giacomantone, Oscar Bria, Armando De Giusti

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

Resumen. Las cámaras de tiempo de vuelo (TOF) generan dos imágenes simultáneas, una de intensidad y una de rango. Esto permite abordar problemas de segmentación donde la información de intensidad o de rango separadamente es insuficiente para extraer los objetos de interés de la escena 3D. A su vez, la información de rango permite obtener una aproximación del vector normal de cada punto de las superficies capturadas. En este artículo se presenta un método de segmentación espectral, que combina la información de intensidad, de rango y las orientaciones de los vectores normales para mejorar los resultados de la segmentación. Los agrupamientos obtenidos suponen una estructura subyacente común entre todas las fuentes de información, llamadas vistas. Se utilizan técnicas de clustering espectral co-regularizado para obtener agrupamientos que sean consistentes de acuerdo a todas las vistas. La evaluación del método propuesto fue realizado sobre imágenes reales. El rendimiento obtenido al combinar las tres fuentes de información presenta mejoras en los agrupamientos resultantes.

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

1. Introducción

El problema de segmentación de imágenes es uno de los principales problemas en el campo de la visión automática. Su objetivo consiste en la extracción de los elementos que constituyen una imagen [3][15]. Para poder lograrlo, estos métodos agrupan pixeles de acuerdo a algún criterio de semejanza. Tradicionalmente, el problema de segmentación de imágenes es abordado utilizando la información de color o intensidad de los objetos presentes en la escena. Los progresos recientes en segmentación de imágenes han mostrado que incorporar la profundidad de los objetos como característica adicional mejora la precisión de los métodos de segmentación [9]. A su vez, a partir de una nube de puntos es posible obtener los vectores normales locales de las superficies. Estos vectores permiten discriminar con mayor precisión los objetos presentes en la escena [7] [6].

Desarrollos recientes en *hardware* permiten estimar la geometría de la escena y posibilitan la utilización de nuevos enfoques para segmentar imágenes. Con esta perspectiva el problema de segmentación puede ser formulado como la búsqueda de formas efectivas para particionar adecuadamente un conjunto de muestras con información de intensidad, distancia e información acerca de la geometría de los objetos presentes en la escena.

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 [4]. 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].

El método propuesto, en una primera etapa, obtiene la información geométrica de la escena a partir de la nube de puntos organizada que provee la cámara TOF mediante el cálculo de los vectores normales locales a las superficies. Luego combina la información de intensidad, de distancia y geométrica para mejorar la calidad de la segmentación utilizando técnicas de agrupamiento espectral coregularizado [8]. Las técnicas de agrupamiento espectral han mostrado resultados prometedores en el campo de la segmentación de imágenes [14] [5]. Estás técnicas requieren la construcción de un grafo de afinidad entre los pixeles de las imágenes y la resolución de un problema generalizado de autovalores. El método utiliza un mecanismo que permite la construcción del grafo de afinidad que se ajusta a las características particulares de cada imagen. Además, con el objetivo de reducir la demanda computacional obtiene una solución aproximada del problema generalizado de autovalores.

El método propuesto es evaluado comparando 4 métricas de evaluación supervisadas [13] sobre un conjunto de imágenes reales.
El artículo está organizado del siguiente modo, en la sección 2 se muestra la técnica utilizada para la obtención del vector normal a las superficies, en la sección 3 se presenta una revisión de los conceptos de agrupamiento espectral utilizados en el método propuesto. En la sección 4 se expone el método. En la sección 5 se presentan resultados experimentales. Finalmente en la sección 6 se presentan las conclusiones.

2. Estimación del vector normal

Una de las características más importantes para la interpretación de datos de rango es el vector normal de la superficie. Sin embargo, ya que no se puede medir de forma directa, tiene que ser estimado para cada elemento de la nube de puntos. Los nuevos sensores de profundidad desarrollados obtienen nubes de puntos organizadas, en donde los puntos de la nube 3D son muestreados a partir

de una grilla regular 2D. Esto permite la utilización de imágenes integrales para optimizar los tiempos de procesamiento [7].

Una imagen integral I_o correspondiente a una imagen O se define como la suma de todos los elementos que se encuentran dentro de un área rectangular entre $(0,0)$ y (m,n): $I_O(m,n) = \sum_{i=0}^{m}\sum_{j=0}^{n} O(i,j)$. Es posible calcular una imagen integral con una sola pasada por la imagen.

El valor promedio de una región puede ser calculado como:

$$S(I_O,m,n,r) = \frac{1}{4r^2}(I_O(m+r,n+r) - I_O(m-r,n+r) \\ - I_O(m+r,n-r) + I_O(m-r,n-r))$$

En donde (m,n) es el centro de la región y r el radio interno de la región rectangular.

Un modo tradicional de estimar el vector normal a la superficie $\overrightarrow{n}_p$ en un punto p en la ubicación en la imagen $(m,n)^T$ es calcular el vector 3D $\overrightarrow{v}_{p,h}$ entre el vecino izquierdo y derecho; calcular el vector $\overrightarrow{v}_{p,v}$ entre el vecino superior e inferior de p y luego calcular el producto externo entre los dos vectores: $\overrightarrow{n}_p = \overrightarrow{v}_{p,h} \times \overrightarrow{v}_{p,v}$. A partir de la utilización de imágenes integrales es posible calcular el vector normal para cada pixel de la nube de puntos organizada de forma eficiente teniendo en cuenta que:

$$\overrightarrow{v}_{p,h,x} = \frac{P_x(m+r,n) - P_x(m-r,n)}{2} \qquad \overrightarrow{v}_{p,v,x} = \frac{P_x(m,n+r) - P_x(m,n-r)}{2}$$

$$\overrightarrow{v}_{p,h,y} = \frac{P_y(m+r,n) - P_y(m-r,n)}{2} \qquad \overrightarrow{v}_{p,v,y} = \frac{P_x(m,n+r) - P_y(m,n-r)}{2}$$

$$\overrightarrow{v}_{p,h,z} = \frac{S(I_{P_z},m+1,n,r-1)}{2} - \frac{S(I_{P_z},m-1,n,r-1)}{2} \qquad \overrightarrow{v}_{p,v,z} = \frac{S(I_{P_z},m,n+1,r-1)}{2} - \frac{S(I_{P_z},m,n-1,r-1)}{2}$$

Donde P_x, P_y y P_z son mapas de las coordenadas x, y y z de la nube de puntos organizada, I_{P_Z} es la imagen integral del componente z de la nube de puntos y r el radio de suavizado a utilizar.

3. Agrupamiento espectral

Dado un conjunto de patrones $X = \{x_1, x_2, ...x_n\} \in \mathbb{R}^m$, y una función de semejanza $d : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, es posible construir una matriz de afinidad W tal que $W(i,j) = d(x_i, x_j)$. Los algoritmos de agrupamiento espectral obtienen una representación de los datos en un espacio de dimensión inferior resolviendo el siguiente problema de optimización:

$$\begin{array}{ll} \underset{U \in \mathbb{R}^{n \times k}}{\text{máx}} & Tr\left(U^T L U\right) \\ \text{s.a.} & U^T U = I \end{array} \tag{1}$$

donde $L = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$ es la matriz laplaciana de W de acuerdo a [12] y D es una matriz diagonal con la suma de las filas de W ubicadas en su diagonal principal. Una vez obtenido U sus filas son consideradas como las nuevas coordenadas de los patrones. En esta nueva representación es más sencillo aplicar un algoritmo de clustering tradicional [14].

Es posible obtener una aproximación a las coordenadas en este nuevo espacio calculando las afinidades de un pequeño conjunto de píxeles y aproximando las afinidades restantes.

Sea $A \subset X$ un subconjunto de patrones muestreados y $B = V - A$, el resto de los patrones no muestreados. W_A es la matriz de semejanza derivada de los datos de A y L_A es la matriz laplaciana de W_A. W_B y L_B son las matrices correspondientes de las afinidades de los puntos de A y B. Es posible definir a L como:

$$W = \begin{bmatrix} W_A & W_B \\ W_B^T & W_C \end{bmatrix} \qquad L = \begin{bmatrix} L_A & L_B \\ L_B^T & L_C \end{bmatrix}$$

Es posible obtener una aproximación de W, denominada $\hat{W}$, solamente a partir de A y B:

$$\hat{W} = \bar{U} \Lambda \bar{U}^T = \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix}$$

Con el objetivo de obtener los autovectores de la matriz laplaciana aproximada, $\hat{L} = \hat{D}^{\frac{1}{2}} \hat{W} \hat{D}^{\frac{1}{2}}$, es necesario calcular $\hat{L_A}$ y $\hat{L_B}$:

$$\hat{L_A}_{ij} = \frac{W_{A_{ij}}}{\sqrt{\hat{d}_i \hat{d}_j}} \qquad \hat{L_B}_{ij} = \frac{W_{B_{ij}}}{\sqrt{\hat{d}_i \hat{d}_{j+|A|}}} \tag{2}$$

donde $\hat{d} = \hat{W}\mathbf{1}$. Si $\hat{L_A}$ es positiva definida, es posible hallar los autovectores ortogonales aproximados en un solo paso. Sea $S = \hat{L_A} + \hat{L_A}^{-\frac{1}{2}} \hat{L_B} \hat{L_B}^T \hat{L_A}^{-\frac{1}{2}}$ y su diagonalización $S = U_S \Lambda_S U_S^T$, Fowkles et al. [5] demostraron que si la matriz V se define como

$$V = \begin{bmatrix} \hat{L_A} \\ \hat{L_B}^T \end{bmatrix} \hat{L_A}^{-\frac{1}{2}} U_S \Lambda_S^{-\frac{1}{2}} \tag{3}$$

$\hat{L}$ es diagonalizada por V y por Λ_S y $V^T V = I$

3.1. Co-regularización

Cuando el conjunto de datos tiene más de una representación, a cada una de ellas se la denomina vistas. En el contexto de agrupamiento espectral las técnicas de co-regularización intentan fomentar la semejanza de los ejemplos en la nueva representación generada a partir de los autovectores de cada una de las vistas.

Sean $X^{(v)} = \{x_1^{(v)}, x_2^{(v)}, ..., x_m^{(v)}\}$ los ejemplos para la vista v y $L^{(v)}$ la matriz laplaciana creada a partir de X para la vista v. Definimos $U^{(v)}$ a la matriz formada por los primeros k autovectores correspondientes como la matriz $L^{(v)}$ de acuerdo con (1). En [8] fue propuesto un criterio que mide el desacuerdo entre dos representaciones:

$$D(U^{(v)}, U^{(w)}) = \left|\left| \frac{K_{U^{(v)}}}{||K_{U^{(v)}}||_F} - \frac{K_{U^{(w)}}}{||K_{U^{(w)}}||_F} \right|\right|_F^2$$

Donde $K_{U^{(v)}}$ es la matriz de semejanza generada a partir de los patrones en la nueva representación $U^{(v)}$ y $||\cdot||_F$ es la norma Frobenius. Si se utiliza como medida de semejanza el producto interno entre los vectores se obtiene $K_{U^{(v)}} = U^{(v)}{U^{(v)}}^T$. Ignorando las constantes aditivas y de escalado, la ecuación anterior puede ser formulada de la siguiente manera:

$$D(U^{(v)}, U^{(w)}) = -Tr\left(U^{(v)}{U^{(v)}}^T U^{(w)}{U^{(w)}}^T\right) \tag{4}$$

El objetivo es minimizar el desacuerdo entre las representaciones obtenidas a partir de cada una de las vistas. Por lo tanto, si se disponen de m vistas, se obtiene el siguiente problema de optimización que combina los objetivos de agrupamiento espectral individuales y el objetivo que determina el desacuerdo entre las representaciones:

$$\begin{aligned} \max_{\substack{U^{(i)} \in R^{n\times k}, \\ 1 \le i \le m}} \quad & \sum_{v=1}^{m} Tr\left({U^{(v)}}^T L^{(v)} U^{(v)}\right) + \lambda \sum_{\substack{1 \le v,\, w \le m \\ v \ne w}} Tr\left({U^{(w)}}^T L^{(w)} U^{(w)}\right) \\ \text{s.a.} \quad & {U^{(v)}}^T U^{(v)} = I \quad \forall 1 \le v \le m \end{aligned} \tag{5}$$

El parámetro λ balancea el objetivo de agrupamiento espectral y el de desacuerdo entre las representaciones. El problema de optimización conjunta puede ser resuelto utilizando maximización alternante. Dados $U^{(w)}, 1 \le w \le m$ dado se obtiene el siguiente problema de optimización para $U^{(v)}, v \ne w$:

$$\begin{aligned} \max_{U^{(v)} \in R^{n\times k}} \quad & Tr\left({U^{(v)}}^T \left(L^{(v)} + \lambda \sum_{\substack{1 \le w \le m \\ v \ne w}} U^{(w)}{U^{(w)}}^T\right) U^{(v)}\right) \\ \text{s.a.} \quad & {U^{(v)}}^T U^{(v)} = I \end{aligned} \tag{6}$$

Lo que resulta en un algoritmo de clustering tradicional con la matriz laplaciana modificada $L^{(v)} + \lambda \sum_{\substack{1 \le w \le m \\ v \ne w}} U^{(w)}{U^{(w)}}^T$

4. Método propuesto

A partir de la imagen de intensidad I y la imagen de rango R provistas por la cámara de tiempo de vuelo se obtiene el mapa con los vectores normales de cada punto de la superficie, N, según lo descripto en la sección 2.

Para determinar la semejanza W_{ij} entre cada elemento de una imagen Img $\in \{I, R, N\}$ se utiliza una función que combina la distancia de los pixeles en el plano de la imagen y la semejanza entre sus valores:

$$W(\text{Img})_{ij} = \exp\left(\frac{-||\text{pos}_i - \text{pos}_j||_2^2}{2(sx)^2} - \frac{d(\text{Img}(i), \text{Img}(j))^2}{2(sy)^2}\right)$$

En donde pos_i es la ubicación espacial (x, y) del pixel i-ésimo; Img(i) son los elementos i-ésimos de la imagen. El parámetro sx determina la importancia otorgada a la ubicación espacial en la función de semejanza y sy determina la importancia otorgada a la diferencia entre los valores de cada pixel. d es una función de distancia entre los elementos de la imagen.

En lugar de seleccionar un solo parámetro sy para toda la imagen en [16] proponen calcular un parámetro de escalado local para cada punto teniendo en cuenta las estadísticas locales de su vecindad. La escala local para un punto i de una imagen P utilizando una distancia d se define como máx $d(P(i), P(j))\forall j \in N(i)$, en donde $N(i)$ son todos los vecinos dentro de un radio de r pixeles.

Sean p, r dos elementos de la imagen I,$d_I(p, r) = |p - r|$. La misma función se utiliza para la imagen de rango. Si p, r sos dos elementos de N, $d_N(p, r) = \mathbf{p}^T\mathbf{r}$.

El método propuesto consiste en las siguientes etapas:

1. A partir de I, R y N se obtienen las matrices laplacianas aproximadas $\hat{L_1}$ y $\hat{L_2}$, $\hat{L_3}$, respectivamente según lo descripto en (2). Utilizando para cada imagen su función de semejanza correspondientes: $W(I), W(R)$ y $W(N)$.
2. Se obtiene $\hat{V_2}$, $\hat{V_3}$ los autovectores aproximados de $\hat{L_2}$ y $\hat{L_3}$ calculados de acuerdo a (3)
3. Se resuelve el problema de optimización 6 para $\hat{V_1}$ dadas $\hat{V_2}$ y $\hat{V_3}$.
4. La optimización es llevada a cabo de forma cíclica por todas las vistas manteniendo fijas las obtenidas previamente.
5. Se evalúa $\sum_{i=1}^{3}\sum_{j=1}^{3} D(V_i, V_j)$. Si el desacuerdo disminuye ir a 4.
6. Se aplica un algoritmo de agrupamiento sobre $\hat{V_1}$

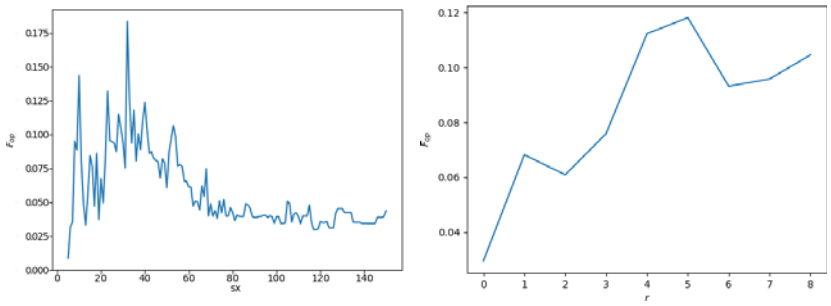
5. Resultados experimentales

El rendimiento del algoritmo de segmentación propuesto fue evaluado sobre 13 imágenes capturadas utilizando la cámara de tiempo de vuelo MESA Swiss-Ranger SR4000 [4]. La cámara de tiempo de vuelo proporciona dos imágenes: una imagen de amplitud y una imagen de rango ambas de 144×176 pixeles.

Se utilizan las siguientes métricas evaluación supervisadas para determinar la calidad de la segmentación obtenida con el método propuesto: la medida de precisión-exhaustividad para objetos y partes, F_{op}, propuesto en [13], la medida de precisión-exhaustividad para contornos [10], F_b, la Cobertura de Segmentación [1] , SegCov y la variación de información [11], VoI . En este contexto una segmentación obtenida es mejor cuanto menor es su VoI y cuando mayor es su SegCov, F_b y su F_{op}.

Se utilizó un muestreo de 150 pixeles para construir las matrices de afinidad y se obtuvieron los 6 autovectores correspondientes a los autovalores de mayor magnitud para generar el espacio en donde se aplica k-medias como método de agrupamiento. Como término de co-regularización se utilizó $\lambda = 0{,}001$ según lo sugerido en [8].

Con el objetivo de determinar el parámetro sx, se evaluó la métrica F_{op} para el conjunto de las imágenes TOF. La figura 1a muestra el promedio de los resultados para cada imagen con respecto a la variación de la influencia de la ubicación espacial en la función de semejanza. Teniendo en cuenta los resultados se estableció $sx = 35$. La figura 1b muestra la influencia del tamaño de la ventana a considerar al momento de seleccionar el escalado local de sy. Se utilizó $r = 5$.



(a) Influencia del parámetro sx (b) Influencia del parámetro r

Figura 1: Influencia de los parámetros sx y r con respecto a la métrica de precisión-exhaustividad de objetos y partes

La tabla 1 muestra el análisis de rendimiento del método propuesto en comparación con utilizar el método de Nystrom tradicional [5] sobre la imagen de intensidad, de rango, y normal por separado, Co-Regularización entre intensidad y rango [9] y el método propuesto. Los resultados representan el promedio de 10 ejecuciones sobre las 13 imágenes del conjunto de datos. Es posible observar que el método propuesto mejora las métricas F_{op}, F_b y VoI, indicando que los agrupamientos obtenidos tienen mayor coincidencia con los agrupamientos del agrupamiento de referencia. La medida SegCov es mayor para las imágenes de rango e intensidad. Esto puede deberse a que los agrupamientos obtenidos a partir de estas imágenes tienden a agrupar al fondo junto con los objetos. Es-

ta métrica de evaluación no penaliza en gran medida las segmentaciones poco precisas.

	F_b	VoI	F_{op}	Seg. Cov.
Intensidad	0.07935	0.23629	0.32422	2.80678
Rango	0.05077	0.22466	0.27684	**3.05854**
Normal	0.05283	0.17645	0.32911	2.95473
IR Co-Reg	0.08162	0.23582	0.32879	2.70953
Método Propuesto	**0.10558**	**0.16610**	**0.35124**	2.79449

Tabla 1: Evaluación de rendimiento del método propuesto

La figura 2 presenta resultados experimentales del método propuesto aplicado a dos capturas del conjunto de datos. Es posible observar de forma cualitativa que los segmentos obtenidos a partir de la imagen recobran los objetos presentes en la escena.

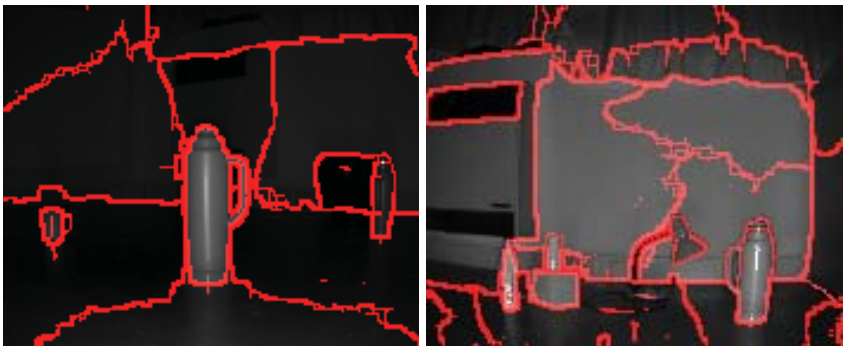


Figura 2: Segmentación realizada utilizando el método propuesto

6. Conclusiones

En este trabajo presentamos un método de agrupamiento aplicado a segmentación de imágenes capturadas con cámaras de tiempo vuelo. Los datos presentan resultados preliminares satisfactorios. El algoritmo extrae información de la geometría de la escena y los combina con los datos de intensidad y rango mejorando los resultados de la segmentación. El rendimiento resultante al utilizar las orientaciones de los vectores normales en el contexto de aprendizaje semi-supervisado

presenta mejoras en los casos probados de acuerdo a las métricas utilizadas. Una etapa futura de este trabajo prevé la obtención de los autovectores aproximados con métodos más eficientes y menos sensibles al ruido, la utilización de otras técnicas de agrupamiento en la etapa final del método y la incorporación de información adicional a la función de semejanza.

Referencias

1. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence 33(5), 898–916 (2011)
2. Blanc, N., Oggier, T., Gruener, G., Weingarten, J., Codourey, A., Seitz, P.: Miniaturized smart cameras for 3d-imaging in real-time [mobile robot applications]. In: Sensors, 2004. Proceedings of IEEE. pp. 471–474 vol.1 (Oct 2004)
3. Canny, J.: A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-8(6), 679–698 (Nov 1986)
4. Cazorla, M., Viejo, D., Pomares, C.: Study of the sr 4000 camera. In: XI Workshop de Agentes FÃsicos (2004)
5. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the nyström method. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(2), 214–225 (Feburary 2004)
6. Holz, D., Behnke, S.: Fast range image segmentation and smoothing using approximate surface reconstruction and region growing. Intelligent autonomous systems 12 pp. 61–73 (2013)
7. Holzer, S., Rusu, R.B., Dixon, M., Gedikli, S., Navab, N.: Adaptive neighborhood selection for real-time surface normal estimation from organized point cloud data using integral images. In: Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on. pp. 2684–2689. IEEE (2012)
8. Kumar, A., Rai, P., Daume, H.: Co-regularized multi-view spectral clustering. In: Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K. (eds.) Advances in Neural Information Processing Systems 24, pp. 1413–1421. Curran Associates, Inc. (2011), http://papers.nips.cc/paper/4360-co-regularized-multi-view-spectral-clustering.pdf
9. Lorenti, L., Giacomantone, J.: Time of flight image segmentation through co-regularized spectral clustering. In: XX Congreso Argentino de Ciencias de la Computación (CACIC 2014) (2014)
10. Martin, D.R., Fowlkes, C.C., Malik, J.: Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE transactions on pattern analysis and machine intelligence 26(5), 530–549 (2004)
11. Meila, M.: Comparing clusterings: an axiomatic view. In: Proceedings of the 22nd international conference on Machine learning. pp. 577–584. ACM (2005)
12. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: Analysis and an algorithm. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS. pp. 849–856. MIT Press (2001)
13. Pont-Tuset, J., Marques, F.: Measures and meta-measures for the supervised evaluation of image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2131–2138 (2013)
14. Shi, J., Malik, J.: Normalized cuts and image segmentation. In: Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. pp. 731–737 (Jun 1997)

15. Wu, Z. y Leahy, R.: An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 15(11), 1101–1113 (Nov 1993)
16. Zelnik-Manor, L., Perona, P.: Self-tuning spectral clustering. In: Advances in neural information processing systems. pp. 1601–1608 (2005)