- ORIGINAL ARTICLE -

Unsupervised TOF Image Segmentation through Spectral Clustering and Region Merging

Segmentación no Supervisada de Imágenes TOF vía Clustering Espectral y Unión de Regiones

Luciano Lorenti [1], Javier Giacomantone[1], and Oscar Bria[1]

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

Abstract

Time of Flight (TOF) cameras generate two simultaneous images, one of intensity and one of range. This allows to tackle segmentation problems in which the separate use of intensity or range information is not enough to extract objects of interest from the 3D scene. In turn, range information allows to obtain a normal vector estimation of each point of the captured surfaces. This article presents a semi-supervised spectral clustering method which combines intensity and range information as well as normal vector orientations to improve segmentation results. The main contribution of this article consists in the use of a statistical region merging as a final step of the segmentation method. The region merging process combines adjacent regions which satisfy a similarity criterion. The performance of the proposed method was evaluated over real images. The use of this final step presents preliminary improvements in the metrics evaluated.

Keywords: Spectral Clustering, TOF images, Unsupervised image segmentation

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 clustering espectral no supervisado que combina la información de intensidad, de rango y las orientaciones de los vectores normales para mejorar los resultados de la segmentación. La principal contribución de este artículo consiste en la utlización de un proceso estadístico de unión de regiones como paso final de método de segmentación. El proceso de union de regiones combina regiones adjacentes que satisfacen un criterio de semejanza. El rendimiento del método propuesto fue evaluado sobre imágenes reales. El uso de este paso final presenta mejoras preliminares en las métricas evaluadas.

Palabras claves: Agrupamiento espectral, Imágenes TOF, Segmentación de imágenes no supervisada

1 Introduction

Image segmentation is one of the main challenges in automatic vision field. Its aim is to extract the elements that constitute an image [1][2]. To achieve this, these methods group pixels according to some similarity criterion. Traditionally, the problem of image segmentation has been tackled using color or intensity information of the objects in the scene. Recent developments in image segmentation have shown that adding the depth of objects as an additional feature improves precision in segmentation methods [3]. From a point cloud, in turn, it is possible to obtain local normal vectors of the surfaces. These vectors allow to discriminate objects in the scene with greater precision [4] [5]. Actual developments in hardware allow to estimate the geometry of the scene and to use new approaches to segment images. With this perspective, the challenge of image segmentation can be posed as the search for effective ways to adequately partition a set of samples with intensity and distance information, as well as information regarding the geometry of the objects in the scene.

In this work we use a TOF camera allowing us to simultaneously obtain range and intensity images. TOF

Citation: L. Lorenti, J. Giacomantone and O. Bria. "*Unsupervised TOF Image Segmentation through Spectral Clustering and Region Merging*". Journal of Computer Science & Technology, vol. 18, no. 2, pp. 97 - 104, 2018.

DOI: 10.24215/16666038.18.e11

Received: February 9, 2018 **Revised:** July 8, 2018 **Accepted:** July 23, 2018.

cameras illuminate the scene with amplitude modulated infrared light-emitting diodes. The sensors of the camera detect the light reflected in the illuminated objects and two images are generated. The intensity image is proportional to the amplitude of the reflected wave and the range or distance image is generated from the phase difference between the emitted and reflected wave in each element of the image [6].

Recently there have been proposed techniques to segmentate objects presents in range and intensity images with the aim of define more precise contours. In [7] a method was proposed that combines range and intensity information via feature product [8]. The method proposed in [3] uses semi-supervised learning to fuse range and intensity information. In [9] surface normals were added to the semi-supervised scheme in order to improve the segmentation quality.

The method proposed in this article incorporates a variation of the region merging stage used in [10] to the algorithm proposed in [9] that refines the segmentation obtained. In the first step, the proposed method uses an optimized co-regularization technique to obtain an over-segmentation of the image. The last step involves merging the over-segmented regions according to a predicate that takes planar property of each region into account.

The proposed method is evaluated by comparing 4 supervised evaluation metrics [11] over a set of real images.

This article is organized as follows: in section 2 we present a revision of spectral clustering that is used in the proposed method; in section 3 we describe the process of region merging; in section 4 we explain the proposed method; in section 5 we present the experimental results; and finally, in section 6 we present the conclusions.

2 Spectral Clustering

Given a set of patterns $X = \{x_1, x_2, ...x_m\} \in \mathbb{R}^m$, and a similarity function $d : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, it is possible to construct an affinity matrix W such that $W(i,j) = d(x_i, x_j)$. Algorithms of spectral clustering obtain a data representation in a lower dimensional space solving the following optimization problem:

$$\begin{array}{rcl} \max\limits_{U \in \mathbb{R}^{n\times k}} & Tr\left(U^T L U\right) & \\ \text{s.t.} & U^T U \;=\; I \end{array} \tag{1}$$

where $L = D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$ is the Laplacian matrix of W according to [12], and D is a diagonal matrix with the sum of the rows of W placed in its main diagonal. Once U has been obtained, its rows are considered the new pattern coordinates. In this new representation it is easier to apply a traditional clustering algorithm [13].

It is possible to obtain an approximation to the pattern coordinates in this new space calculating the affinities of a small set of pixels and approximating the remaining affinities.

Let $A \subset X$ be a subset of sampled patterns and $B = V - A$, the remaining not sampled patterns. W_A is the similarity matrix derived from A data and L_A is the Laplacian matrix of W_A. W_B and L_B are the corresponding affinity matrices of points of A and B. It is possible to define L as:

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

It is possible to obtain an approximation of W, named $\hat{W}$, only from A and B:

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

With the aim of obtaining the eigenvectors of the approximate Laplacian matrix, $\hat{L} = \hat{D}^{\frac{1}{2}}\hat{W}\hat{D}^{\frac{1}{2}}$, it is necessary to calculate $\hat{L_A}$ y $\hat{L_B}$:

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

where $\hat{d} = \hat{W}\mathbf{1}$. If L_A is positive-definite, it is possible to find the approximate orthogonal eigenvectors in just one step. Let S be a matrix defined as $S = \hat{L_A} + \hat{L_A}^{-\frac{1}{2}}\hat{L_B}\hat{L_B}^T\hat{L_A}^{-\frac{1}{2}}$ and its diagonalization $S = U_S\Lambda_S U_S^T$, Fowkles et al. [14] demonstrated that if matrix V is defined as

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

$\hat{L}$ is diagonalized by V and by Λ_S y $V^T V = I$.

2.1 Co-regularization

When the dataset has more than one representation, each of them is named view. In the context of spectral clustering, co-regularization techniques attempt to encourage the similarity of the examples in the new representation generated from the eigenvectors of each view.

Let $X^{(v)} = \{x_1^{(v)}, x_2^{(v)}, ..., x_m^{(v)}\}$ be the samples for view v and $L^{(v)}$ the Laplacian matrix created from X for view v. $U^{(v)}$ is defined as the matrix formed by the first k eigenvectors corresponding to $L^{(v)}$ according to Eq (1). A criterion was proposed in [15] that measures the disagreement between two representations:

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

where $K_{U^{(v)}}$ is the similarity matrix generated from the patterns of the new representation $U^{(v)}$ and $||\cdot||_F$ is the Frobenius norm. If the inner product among the vectors is used as similarity measure, $K_{U^{(v)}} = U^{(v)}{U^{(v)}}^T$ is obtained. Ignoring the constant additive and scaling terms, the previous equation can be formulated as follows:

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

The objective is to minimize the disagreement among the representations obtained from each view. Therefore, if we have m views, we obtain the following optimization problem that combines the individual spectral clustering objectives and the objective that determines the disagreement among the representations:

$$\begin{array}{lrcl} \max\limits_{\substack{U^{(i)} \in R^{n\times k},\\ 1\le i\le m}} & \multicolumn{3}{l}{\sum\limits_{v=1}^{m} Tr\left({U^{(v)}}^T L^{(v)} U^{(v)}\right) + \lambda \sum\limits_{\substack{1\le v,w\le m\\ v\neq w}} Tr\left({U^{(w)}}^T L^{(w)} U^{(w)}\right)} \\ \text{s.t.} & {U^{(v)}}^T U^{(v)} & = & I \quad \forall 1\le v\le m \end{array} \quad (5)$$

The λ parameter balances the spectral clustering objective and the disagreement among the representations. The problem of joint optimization can be solved using alternating maximization. Given $U^{(w)}, 1 \le w \le m$, the following problem of optimization is obtained for $U^{(v)}, v \neq w$:

$$\begin{array}{lrcl} \max\limits_{U^{(v)}\in R^{n\times k}} & \multicolumn{3}{l}{Tr\left({U^{(v)}}^T \left(LM^{(v)}\right) U^{(v)}\right)} \\ \text{s.t.} & {U^{(v)}}^T U^{(v)} & = & I \end{array} \quad (6)$$

resulting in a traditional clustering algorithm with the Laplacian matrix modified $LM^{(v)} = L^{(v)} + \lambda \sum\limits_{\substack{1\le w\le m\\ v\neq w}} U^{(w)}{U^{(w)}}^T$

3 Region Merging

Spectral clustering methods are usually used as a first step as image over-segmentation [16]. Therefore, it is necessary to merge the over-segmented regions to obtain the constitutive elements of the image. The iterative scheme proposed in [10] requires, first of all, the construction of an adjacency graph $RAG = (V,E)$ for region merging. This graph takes each segmented region $v_i \in V$ as a node and each of them is connected with their adjacent regions. Each node is characterized by parameters μ and κ of a Watson distribution, shown in Appendix A, associated to the region, and by the 3D points obtained from the range image corresponding to the pixels of the node. Each edge $e_{ij} \in E$ consists of two weights: w_d, based on statistical dissimilarity, and w_b, based on the similarity of the boundary shared by regions v_i and v_j.

Weight w_d is given by the Bregman divergence among two Watson distributions [17]:

$$D(\theta_i,\theta_j) == F(\theta_i) - F(\theta_j) - <\theta_i - \theta_j, \nabla F(\theta_2)>$$

where θ is the natural parameter of a Watson distribution $W_d(x;\mu,\kappa)$. Given $\mathbf{v} = [\mu_1^2,...,\mu_d^2,\sqrt{2}\mu_1\mu_2,...,\sqrt{2}\mu_{d-1}\mu_d]$, θ is defined as $\theta = \kappa\mathbf{v}$, $F(\theta) = log(M(\frac{1}{2},\frac{d}{2},\kappa))$ and $\nabla F(\theta) = g(\frac{1}{2},\frac{d}{2};\kappa)\frac{\theta}{\kappa}$. Here $M(a,b,\kappa)$ is the confluent hypergeometric Kummer's function and $g(\frac{1}{2},\frac{d}{2};\kappa)$ is the Kummer's ratio.

Then w_d is defined as:

$$w_d(v_i,v_j) = \max(D(\theta_i,\theta_j), D(\theta_j,\theta_i)) \quad (7)$$

Weight w_b based on the boundary shared between two regions v_i and v_j is calculated from the normalized magnitude of the image gradient along the limit of its corresponding regions r_i and r_j as:

$$w_b(v_i,vj) = \frac{1}{|r_i \cap r_j|} \sum_{b\in r_i\cap r_j} I(b) \quad (8)$$

where $r_i \cap r_j$ is a set of boundary pixels between two regions, $|.|$ indicates the cardinality and $I(b)$ is the normalized magnitude of image gradient calculated from the intensity image.

3.1 Merging strategy

The strategy to merge the regions proposed in [10] consists of an iterative procedure that evaluates a merging predicate between the adjacent nodes.

The candidacy of a region defines if the region is valid to be merged with its neighboring nodes. For each node, the candidature criterion proves the planar property of the corresponding region. The planar property can be proved analyzing the concentration parameter κ associated to node v_i. The predicate of the planar property is defined as:

$$candidacy(v_i) = \begin{cases} \text{T} & \text{If } \kappa_i > \kappa_p \\ \text{F} & \text{otherwise} \end{cases}$$

κ_i is the concentration parameter calculated for the region v_i and k_p is the threshold that defines if a region is considered planar.

The eligibility criterion to decide if two regions should be merged evaluates the dissimilarity in the weights of edges w_d and w_b:

$$\text{eligibility}(v_i,v_j) = \begin{cases} \text{T} & \text{if} \quad w_b(v_i,v_j) < th_b \ \text{ and} \\ & \qquad w_d(v_i,v_j) < th_d \\ \text{F} & \qquad \text{otherwise} \end{cases}$$

where th_b and th_d are thresholds associated to the weight based on the contours w_b and the weight based on the distance between regions w_d respectively.

Given two regions r_i and r_j, the consistency criterion evaluates if the two merged regions are still one planar region calculating the plane inlier ratio [18] of the new region. It fits a plane to the 3D point cloud which is the result of combining both regions, and then it calculates the proportion of anomalous points and of points belonging to the plane according to a distance threshold. RANSAC algorithm is used to fit the plane to the point cloud. Therefore, the consistency between r_i and r_j is defined as:

$$\text{consistency}(v_i, v_j) = \begin{cases} \text{T} & \text{If} \quad pl\_i\_r(v_i, v_j) > th_r \\ \text{F} & \quad\text{otherwise} \end{cases}$$

pl_i_r is calculated dividing the total number of inliers, i.e., of the 3D points fitted within the plane based on a distance threshold, by the total number of 3D points of the combined regions. th_r is the threshold associated to the proportion of points within the plane.

The merging predicate involves evaluating the candidacy of each node of the graph, as well as the eligibility of the adjacent nodes to be merged, and verifying the consistency in merging both regions. Two regions are merged if the merging predicate P_{ij} is true. Predicate P_{ij} is defined as:

$$P_{ij} = \begin{cases} \text{T} \quad \text{If} & \begin{array}{ll} (1)\ \text{candidacy}(v_i) & = T \\ (2)\ \text{eligibility}(v_i, v_j) & = T \\ (3)\ \text{consistency}(v_i, v_j) & = T \end{array} \\ \\ \text{F} \quad \text{otherwise.} \end{cases}$$

4 Proposed method

The method proposed in this article incorporates a variation of the region merging stage used in [10] to the algorithm proposed in [9] that refines the segmentation obtained. In the first step, the proposed method performs an spectral embedding of the pixels in both images in a lower-dimensional space more suitable to perform clustering. After that the both embedding are co-regularized to acoord betweenm them. One the embedding converge in an agreement, the subspace is clusterized in orted to obtain an oversegmentation of the image. The last step involves merging the oversegmented regions according to a predicate that takes planar property of each region into account.

From intensity image I and range image R provided by the TOF camera, we obtain the image N with the surface normal vectors of each point . To do this, the library provided in [4] was used.

A function that combines the distance of the pixels in the image plane and the similarity among their values was used to determine the similarity W_{ij} among each element of an image $\text{Img} \in \{I, R, N\}$:

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

where pos_i is the spatial location (x,y) of the i-th pixel, $Img(i)$ are the I-th elements of the image. Parameter sx determines the importance given to the spatial location in the similarity function and sy determines the importance given to difference among the values of each pixel. d is a distance function among the elements of the image.

Instead of selecting only one parameter sy for all the image, what is proposed in [19] is to calculate a local scale parameter for each point considering the local statistics of their vicinity. The local scale for a point i of an image P using a distance d is defined as $\max d(P(i), P(j)) \forall j \in N(i)$, where $N(i)$ are all the neighbors within a ratio r of pixels .

Let p and r be two elements of the image I, $d_I(p,r) = |p - r|$. The same function is used for the range image. If p and r are two elements of N, $d_N(p,r) = \mathbf{p}^T\mathbf{r}$.

The proposed method involves the following steps:

1. The approximate Laplacian matrices $\hat{L_1}$, $\hat{L_2}$ and $\hat{L_3}$ are obtained from I, R and N respectively, as described in (2), using its corresponding similarity function for each image: $W(I)$, $W(R)$ and $W(N)$.

2. The approximate eigenvectors $\hat{V_2}$ and $\hat{V_3}$ are obtained from $\hat{L_2}$ and $\hat{L_3}$ calculated according to Eq (3).

3. The optimization problem in Eq (6) is solved for $\hat{V_1}$ given $\hat{V_2}$ and $\hat{V_3}$.

4. Optimization is cycled over all the views keeping fixed the ones previously obtained.

5. $\sum_{i=1}^{3}\sum_{j=1}^{3} D(V_i, V_j)$ is evaluated. If the disagreement is reduced, go to 4.

6. Algorithm k-means is applied over $\hat{V_1}$ to obtain M regions.

7. The algorithm of region merging described in section 3 is used to obtain $N < M$ regions.

5 Experimental results

5.1 Experimental setup

The performance of the proposed segmentation algorithm was evaluated over a set of real images captured with a TOF camera MESA SwissRanger SR4000 [6]. The TOF camera provides two images: an amplitude image and a range image both of 144 x 176 pixels.

Image segmentation was evaluated through the following parameters: precision and recall measure for

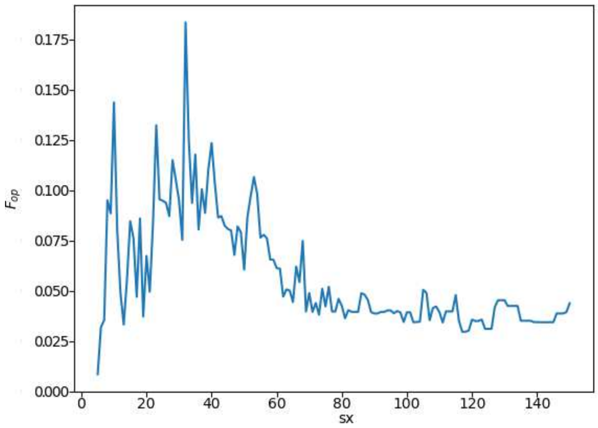

Figure 1: Influence of parameter *sx* in relation to the precision and recall metrics of objects and parts.

objects, F_{op} [11]; precision and recall measure for contours, F_b [20]; segmentation covering, SegCov [21] and variation of information, VoI [22]. An effective image segmentation method decrease the parameter value of VoI, and increase the values of SegCov, Fb and Fop.

As reported by [14] the Nyström method used to approximate the eigenvectors needs sampling only less than one percent of the pixels in the image to obtain performance comparable with the traditional spectral clustering algorithms. For this reason the 0.005% of the total number of pixels were sampled, resulting in a set of 250 evenly spaced pixels to build the affinity matrices. Taking into account that the images have an average of 6 segments, we compute the leading 6 eigenvectors to generate the space where k-means is applied as clustering method. The co-regularization parameter was established as $\lambda = 0.01$ conforming to [15]. According to [10], parameters $\kappa_p = 3$, $th_d = 3.0$, $th_b = 0.2$ and $th_r = 0.9$ were used.

In order to determine the parameter s_x, F_{op} was evaluated for the set of TOF images. Figure 1 shows the average result among all the images in relation to the variation of the influence of the spatial location in the similarity function. Considering the results, s_x = 35 was established. Figure 2 shows the influence of the size of the window to be considered when selecting the local scale of s_y, $r = 5$ was used.

5.2 Results and discussion

Table 1 shows the performance analysis of the proposed method and Nystrom method [14] over the intensity, range and normal image, separately, the use of co-regularization between intensity and range [7] and the use of co-regularization between intensity, range and normal images [9]. Results present the average of 10 runs over 13 images of the dataset. It is possible to observe that the proposed method improves metrics F_{op}, F_b and *Seg.Cov.*, indicating that the obtained clusters have a greater coincidence with the ground truth. In Figure 3, we show results of the proposed method on real images from the dataset. Qualitatively, the segments obtained from the image recover the objects present in the scene.

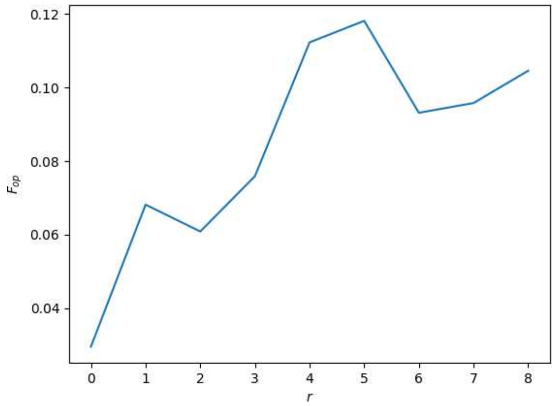

Figure 2: Influence of parameter *r* in relation to the precision and recall metrics of objects and parts.

Table 1: Performance evaluation of the proposed method

	F_b	VoI	F_{op}	Seg.Cov.
I [14]	0.268	0.126	0.309	2.61
R [14]	0.205	**0.054**	0.249	3.192
N [14]	0.228	0.059	0.256	3.043
IR Coreg [7]	0.257	0.106	0.313	2.717
IRN Coreg [9]	0.266	0.096	0.316	2.683
IRN Coreg-RM (Proposed Method)	**0.293**	0.104	**0.482**	**3.24**

Comparing IRN Coreg with IRN Coreg-RM, it is possible to analyze the contributions of the region merging process. The added step improves clustering output in all metrics except for VoI. The metric most improved was F_{op} indicating that the the region merging process helps to recover more adequately the objects presents in the scene.

6 Conclusions

In this work we presented a clustering method applied to segmentation of images captured with TOF cameras, leading to satisfactory preliminary results. The algorithm uses information from the geometry of the scene, as well as intensity and range information, thus improving segmentation results. The use of a region merging process that exploits the planar statistics of the image regions improves the results obtained according to the metrics used. The region merging algorithm find a good balance between preserving the segments obtained and the risk of overmerging for the remaining regions.

A future step of this work foresees the use of more efficient techniques to obtain an approximate eigenvector embedding space, the use of other co-

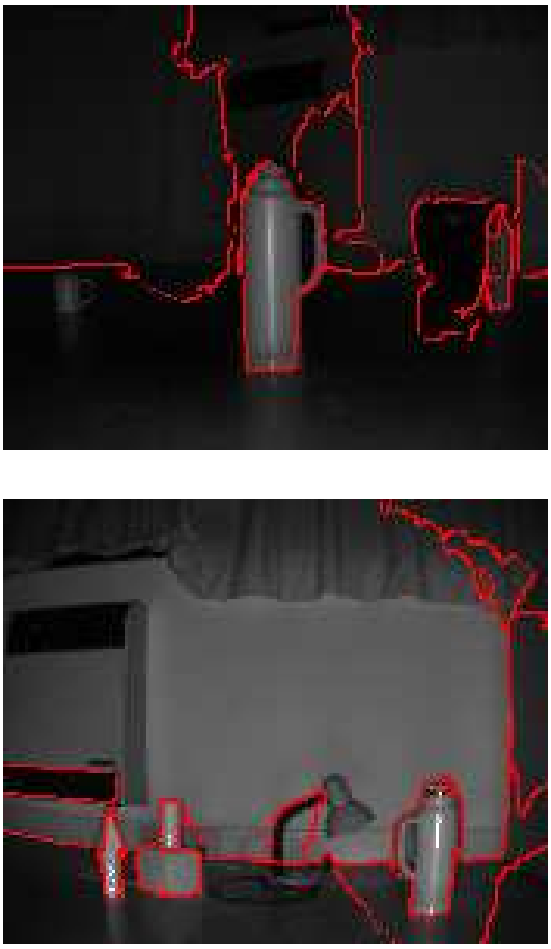

Figure 3: Segmentation obtained using the proposed method

regularization techniques and enhancing the region merging method with intensity information.

Appendix A Watson Distribution

Multivariate Watson distribution is a distribution that models axially symmetric directional data. For a m dimensional unit vector symmetrically axial $\pm x = [x_1, ..., x_m] \in R^m$, multivariate Watson distribution is defined as:

$$W_d(\mathbf{x}|\mu,\kappa) = M(a,c,\kappa)^{-1}\exp(\kappa(u^T\mathbf{x})^2) \quad (9)$$

$$W_d(-\mathbf{x}|\mu,\kappa) = W_d(\mathbf{x}|\mu,\kappa) \quad (10)$$

Where μ is the mean direction (con $|\mu|_2 = 1$), $\kappa \in \mathbb{R}$ is the concentration parameter, $a = \frac{1}{2}$, $c = \frac{m}{2}$, and $M(a,c,\kappa)$ is the confluent hypergeometric Kummer's function. Watson distribution is rotationally symmetric around μ and its shape depends upon the value of concentration parameter κ.

A.1 Maximum likelihood estimation

Let $\mathbf{X} = \{x_1, .., x_n\}, x_i \in \mathbb{R}^m, 1 \leq i \leq n$ be n pointsi.i.d. sampled from a Watson distribution $W_d(x;\mu\kappa)$, with mean μ and concentration κ. The logarithm of the likelihood function is given by:

$$l(\mu,\kappa;\mathbf{X}) = n(\kappa\mu^T S\mu - \log M(\frac{1}{2},\frac{p}{2},\kappa)+\gamma) \quad (11)$$

where $\mathbf{S} = n^{-1}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^T$ is the sample scatter matrix and γ is a constant that can be ignored. It is possible to obtain parameter μ that maximizes Eq (11) as follows:

$$\hat{\mathbf{u}} = \mathbf{s_1} \text{ si } \hat{k} > 0, \hat{\mathbf{u}} = \mathbf{s_p} \text{ si } \hat{k} < 0 \quad (12)$$

where $\mathbf{s_1}, \mathbf{s_2}, ..., \mathbf{s_p}$ are the normalized eigenvectors of the scatter matrix S corresponding to the eigenvalues $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_p$. The estimation of concentration parameter $\hat{\kappa}$ is obtained solving:

$$g(\tfrac{1}{2},\tfrac{p}{2},\hat{\kappa}) = \frac{M'(\frac{1}{2},\frac{p}{2},\hat{\kappa})}{M(\frac{1}{2},\frac{p}{2},\hat{\kappa})} = \hat{\mu}^T\mathbf{S}\hat{\mu} := r$$

It is possible to calculate $\hat{\kappa}(r)$ by means of the bounds proposed by [23].

- $\hat{k}(r) \approx U(r)$ for $0 < r < \dfrac{a}{2c}$
- $\hat{k}(r) \approx B(r)$ for $\dfrac{a}{2c} \leq r < \dfrac{2a}{\sqrt{c}}$
- $\hat{k}(r) \approx L(r)$ for $\dfrac{2a}{\sqrt{c}} \leq c < 1$

where $L(r), B(r)$ and $U(r)$ are defined as follows:

- $L(r) = \dfrac{rc-a}{r(1-r)}\left(1+\dfrac{1-r}{c-a}\right)$
- $B(r) = \dfrac{rc-a}{2r(1-r)}\left(1+\sqrt{1+\dfrac{4(c+1)r(1-r)}{a(c-a)}}\right)$
- $U(r) = \dfrac{rc-a}{r(1-r)}\left(1+\dfrac{r}{a}\right)$

Given that Eq (12) has two possible solutions, the easiest way to obtain parameter μ is by solving both equations and selecting the one with greater log-likelihood.

Competing interests

The authors have declared that no competing interests exist.

References

[1] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, pp. 679–698, Nov 1986.

[2] R. Wu, Z. y 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, pp. 1101–1113, Nov 1993.

[3] L. Lorenti and J. Giacomantone, "Time of flight image segmentation through co-regularized spectral clustering," in *XX Congreso Argentino de Ciencias de la Computación (CACIC 2014)*, pp. 101–110, 2014.

[4] S. Holzer, R. B. Rusu, M. Dixon, S. Gedikli, and N. Navab, "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.

[5] D. Holz and S. Behnke, "Fast range image segmentation and smoothing using approximate surface reconstruction and region growing," *Intelligent autonomous systems 12*, pp. 61–73, 2013.

[6] M. Cazorla, D. Viejo, and C. Pomares, "Study of the sr 4000 camera," in *X I Workshop de Agentes Físicos*, pp. 88–97, 2004.

[7] L. Lorenti and J. Giacomantone, "Segmentación espectral de imágenes utilizando cámaras de tiempo de vuelo," in *XVIII Congreso Argentino de Ciencias de la Computación*, pp. 430–439, 2013.

[8] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[9] L. Lorenti, J. Giacomantone, O. N. Bria, and A. E. De Giusti, "Fusión de información de geometría e intensidad para segmentación de imágenes tof," in *XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017).*, pp. 508–517, 2017.

[10] M. A. Hasnat, O. Alata, and A. Trémeau, "Joint color-spatial-directional clustering and region merging (jcsd-rm) for unsupervised rgb-d image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 11, pp. 2255–2268, 2016.

[11] J. Pont-Tuset and F. Marques, "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.

[12] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, pp. 849–856, MIT Press, 2001.

[13] J. Shi and J. Malik, "Normalized cuts and image segmentation," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp. 731–737, Jun 1997.

[14] C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the nyström method," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 214–225, Feburary 2004.

[15] A. Kumar, P. Rai, and H. Daume, "Co-regularized multi-view spectral clustering," in *Advances in Neural Information Processing Systems 24* (J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, eds.), pp. 1413–1421, Curran Associates, Inc., 2011.

[16] G. Pagnutti and P. Zanuttigh, "Joint color and depth segmentation based on region merging and surface fitting.," in *VISIGRAPP (4: VISAPP)*, pp. 93–100, 2016.

[17] M. A. Hasnat, O. Alata, and A. Trémeau, "Unsupervised clustering of depth images using watson mixture model," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pp. 214–219, IEEE, 2014.

[18] B. Peng, L. Zhang, and D. Zhang, "Automatic image segmentation by dynamic region merging," *IEEE Transactions on image processing*, vol. 20, no. 12, pp. 3592–3605, 2011.

[19] L. Zelnik-Manor and P. Perona, "Self-tuning spectral clustering," in *Advances in neural information processing systems*, pp. 1601–1608, 2005.

[20] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 5, pp. 530–549, 2004.

[21] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 5, pp. 898–916, 2011.

[22] M. Meila, "Comparing clusterings: an axiomatic view," in *Proceedings of the 22nd international conference on Machine learning*, pp. 577–584, ACM, 2005.

[23] S. Sra and D. Karp, “The multivariate watson distribution: Maximum-likelihood estimation and other aspects,” *Journal of Multivariate Analysis*, vol. 114, pp. 256–269, 2013.