International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

An Adaptive Recursive Reconstruction Technique for Segmentation of Images

Peram Nagamani[1], Gunna Kishore[2]

[1, 2]Computer Science and Engineering, Rise Krishna Sai Prakasam Group of Institutions, Ongole, India

Abstract: ***Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels).Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). Among the various approaches proposed for this task, unsupervised methods have the advantage of being able to segment images without any assistance from the user. However, such methods often suffer from long runtimes and tend to be sensitive to the choice of parameters. Because of these problems, users will often prefer semi-supervised methods, which provide a more controllable output in the same amount of time. This paper proposes a new unsupervised approach, based on random walks, which maps each pixel to the most probable label in a local neighbourhood. To make this approach more robust to the choice and learning of the parameters, we propose an efficient computational technique, in which the parameters and the segmentation probabilities are recomputed alternatively. We also describe a refinement strategy that improves the speed and accuracy of the segmentation by applying random walks at different scales.***

Keywords: Segmentation, Partitioning, reconstruction, random walks, accuracy.

1. Introduction

Image segmentation is an important step in many applications, sometimes even the ultimate goal of the analysis. It remains a challenging problem which requires the search for an optimal partition of the pixel grid of an image. Even under simple model assumptions, this problem is NP-hard. The task of image segmentation has thus been addressed by various constructive algorithms, e.g., watershed segmentation as well as by spectral methods such as normalized cuts. All segmentation algorithms have design parameters which need to be tuned for each specific application. Hence, a quantitative validation of segmentations is all the more important. Given segmentation, there are two possible types of errors: (i) under-segmentation – a segment contains parts which belong to different regions and should be split; (ii) over-segmentation, two adjacent segments in fact belong to the same region and should be merged. Most image segmentations suffer from at least one if not both types of errors.

The automated extraction of regions in an image, known as image segmentation, is a challenging yet important task. In medical imaging, for instance, the segmentation of images or 3D volumes is essential to several key applications, such as the diagnosis and monitoring of pathologies, the selection of optimal treatment plans, as well as image-guide procedures. The segmentation of natural images is also an important step in the identification and classification of persons and objects, and has applications in various domains such as surveillance.

While supervised and semi-supervised segmentation methods like [1] and [2] offer a greater level of control, the manual process of collecting or annotating images can be very tedious and time-consuming. Unsupervised approaches, on the other hand, can segment images without any human assistance. Such segmentation approaches can be divided in two broad categories: region-based methods and edge-based methods [3]. Region-based methods focus on grouping pixels into regions sharing common properties like color, texture or gradient. This category includes methods based on clustering [4], [5], Gaussian Mixture Models (GMM) [6], mode-seeking methods like the Mean-Shift algorithm [7], and Markov Random Fields (MRFs) [8]. Edge-based segmentation approaches, on the other hand, attempt to describe the regions in an image using their boundary. These boundaries can be found by detecting discontinuities in the image, for instance, using edge detection filters [9].

Although recent unsupervised methods like [3] have considerably reduced the gap in terms of quality between automatic segmentation and human-assisted segmentation, such methods typically suffer from two problems: 1) they often need to learn a parametric model, which makes them slower than semi-supervised methods, and 2) they are also more sensitive to the choice of parameters. Because of these problems, users will often prefer semi-supervised methods, which provide a more controllable output in the same amount of time.

This paper proposes a new unsupervised segmentation approach, based on random walks, which offers an accuracy comparable to state-of-the-art methods, while being much faster than these methods. The proposed approach is similar to mode-seeking methods like Mean-Shift [7], where pixels are mapped to the most probable label in a local neighbourhood (i.e., the modes). While these methods typically use a geometric distance measure to define this neighborhood, our approach uses a diffusion process to define the neighborhood, which allows it to better fit the regions in the image. Moreover, our approach uses a Gaussian Mixture Model (GMM) to learn the class distributions in the feature space. Because the segmentation results highly depend on the initial number of classes (i.e., mixture components) and the ability to properly learn the distribution parameters, we propose a strategy to dynamically adjust these parameters. While inspired by a similar technique used in Hidden Markov Random Fields (HMRF) [8], this strategy is more efficient and has the advantage of guaranteeing a global optimum. Finally, our

segmentation approach also uses an original refinement technique, in which random walks are applied at different scales. As well as making the approach faster, this technique also helps obtain a higher accuracy.

2. The Proposed Approach

Given an image $\mathcal{I}$ to segment, we denote as i a pixel of $\mathcal{I}$ and write $xi \in R^d$ the d-dimensional feature vector of i. Although more complex features like texture could be considered, to have a fast segmentation, we only use the CIE-LAB color coordinates as features. The unsupervised segmentation problem consists in assigning a label from a set C to each pixel i, without any assistance from a human and such that the regions defined by pixels with the same label correspond to the true regions in the image.

Our objective is to develop a new unsupervised segmentation approach that is both fast and accurate. To achieve this objective, we propose a model composed of three steps. In the first step, we learn the class distribution parameters (i.e., class priors) using a Gaussian Mixture Model (GMM) approach. As we show in our experiments, segmenting an image using only these priors leads to inaccurate results due to the fact that the spatial coherence of class labels is not considered. To avoid this problem, in the second step, we use a stochastic model based on random walks to obtain a more coherent segmentation. To find global regions in the image and speedup the segmentation, we apply our model on a downscaled version of the image. Since the resulting segmentation is very coarse, in the third step, we then refine it at the original image scale. This refinement step is composed of three sub-steps. First, we extract the connected components in the image as individual regions. We then apply a few random walk iterations at full resolution to smooth the edges of the segmented regions. Finally, we reduce the number of regions by merging those whose size is smaller than a threshold. In the following subsections, we explain each of these steps in details and provide information on how to implement them efficiently.

2.1. Class Parametersestimation

The first step is to learn the class priors that will be used during the segmentation process. As in many segmentation methods, we suppose that the feature probability distribution of each class y_k follows a Gaussian model:

$$p(x_i|y_k = G(x_i; \mu_k, \sum k)) \quad (1)$$

We use GMM to estimate distribution parameters (μ_k, Σ_k) and the class a priori probabilities $p(y_k)$. In GMM segmentation, each pixel is assigned to the class maximizing the posterior probability:

$$p(y_k|x_i) = \frac{p(x_i|y_k)p(y_k)}{\sum_{k' \in c} p(x_i|y_k')p(y_k')} \quad (2)$$

While simple, this method requires the user to specify the number of classes (regions) K in the image, a parameter which varies from one image to the next and can have a significant impact on the segmentation results. To be more robust to the initial value of K and the class parameters, in SectionII-B2, we present a technique that re computes these parameters iteratively

2.2. Coarse Segmentation using Random Walks

Once the class parameters have been obtained, we then apply a random walk segmentation model on a downscaled version of the image (1:4 ratio in our experiments), obtained using bicubic interpolation. We use a downscaled image to accelerate the computation as well as to find larger regions that better correspond to the true regions in the image.

1) Basic model: Our proposed segmentation model is based on a generative process where a random walker moves along the nodes of a graph corresponding to the pixels of the target image. At each step, the walker can either terminate the walk or generate a class label based on the features of the current pixel, or move to a neighbor pixel. Since these movements are more probable between similar pixels, the pixels within a region will have a greater chance of generating the same label.

Each pixel $i \in \mathcal{I}$ is connected to a set N_i of neighbour pixels, defined as the 8-neighborhood in this work, and the weight of an edge between neighbor pixels i and j is given by

$$w_{ij} = \exp\left(-\sum_{i=1}^{d} \gamma_l (x_{il} - x_{jl})^2\right) \quad (3)$$

In this formulation, γ_l control the relative influence of each feature on the weights computation. Since our features are CIE-LAB coordinates, we have three parameters: γ_L, γ_A and γ_B.

The transition probabilities q_{ij} from pixel i to pixel j are defined using the edge weights

$$q_{ij} = \frac{w_{ij}}{\sum_{j' \in N_i} w_{ij}'} \quad (4)$$

Denote by W the matrix of elements w_{ij} and let D be a diagonal matrix such that $d_{ij} = \sum_j w_{ij}$. The transition probabilities can be expressed as $Q = D^{-1}W$.

At each step, the walker can stop its walk with probability $\alpha > 0$ and emit a label based on the intensity of the current pixel i. The probability of emitting label y_k at pixel i is defined as $e_{ik} = p(y_k|x_i)$, and we denote by E the matrix containing these emission probabilities. Otherwise, if the walk continues, the walker chooses a neighbor $j \in N_i$ with probability qij and moves to this neighbor

Let $S_{ik}^{(t)}$ be the probability of emitting label y_k in a random walk of length t *starting* at pixel i. These probabilities can be expressed as a matrix

$$S^{(t)} = \alpha(1-\alpha)^t Q^t E \quad (5)$$

The total probability S_{ik} of producing label y_k starting at pixel i, considering all possible walk lengths, is thus

$$S = \sum_{t=0}^{\infty} S^{(t)} = \alpha(I - (1-\alpha)Q)^{-1}E \quad (6)$$

Finally, image I is segmented by mapping each pixel i to the label y_k with maximum probability s_{ik}.

This basic model is related to other segmentation methods. Thus, as Mean-Shift [7], our model can be considered as a mode seeking method, where the most probable class of a pixel is obtained by considering the pixels within a small local region. However, while Mean-Shift uses a geometric distance measure to find the modes, our model uses a diffusion distance based on random walks. Parameter α controls the size of the local region. Also, when $\alpha = 1$ is used, the proposed formulation corresponds to GMM segmentation [10]. Finally, if seeds (i.e., manually labeled pixels) are used instead of class priors, our basic model is then equivalent to the Random Walks with Restart (RWR) segmentation algorithm [2].

2) Adaptive model: Because our method uses GMM to estimate the class priors, it is sensitive to the number of classes K. Using a too big value for K will lead to the over segmentation of the image, and a too small value to its under segmentation. To avoid this problem, we use a strategy similar to the Expectation Maximization (EM) process of Hidden Markov Random Fields (HMRF) methods [8]. Starting with initial parameters obtained using GMM, we alternate between computing the segmentation probabilities S and re-estimating the distribution parameters, until convergence. Given probabilities S, and let $M_k = \sum_{i \in I} s_{ik}$, the distribution parameters of label yk can be recomputed as follows:

$$\mu_{k=\frac{1}{M_k}} \sum_{i \in I} s_{ik}\, x_i \quad (7)$$

$$\sum k_{=\frac{1}{M_k}} \sum_{i \in I} s_{ik}\, (x_i \ -\mu_k)(x_i \ -\mu_k)^T \quad (8)$$

$$P(y_k) = \frac{1}{|I|} M_K \quad (9)$$

For the pixel-based method, computing the segmentation probabilities of Eq. (6) at each iteration can be expensive. To speed up this process, we perform the following Cholesky decomposition $LL^T = D-(1-\alpha)W$ in a pre-processing step, and then recompute the probabilities at each iteration as

$$S=\alpha(L^T)^{-1}L^{-1}DE \quad (10)$$

Since L is triangular, this computation can be done efficiently by using forward (back) substitution. Moreover, we use the fact that the decomposed matrix $D-(1-\alpha)W$ is sparse to further accelerate the computation, by setting to 0 the values of L lower than a given threshold (a value of $10-8$ was used in our experiments). Thresholding these values can greatly accelerate the computation without affecting the results. Since the sparsity of L is similar to that of $D - (1 - \alpha)W$, which has only $|Ni|+1$ non-zero elements in each row i, the time complexityof solving Eq. (10) is in $O(|I| \cdot K)$. Finally, since the linear system can be solved independently for each column of E, the computation can easily be parallelized.

In MRF-based methods, finding the most probable label assignment is a NP-hard problem when the number of classes is greater than two (as is often the case in unsupersived segmentation). Hence, these methods generally use heuristic techniques like the Iterative Conditional Mode (ICM) algorithm for the inference process, which do not guarantee a global optimum. An advantage of our proposed model over MRF-based approaches is that it guarantees to find a unique global solution efficiently.

2.3. Full-resolution segmentation refinement

Since we apply the random walk method on a downscaled version of the image, the obtained segmentation can be very coarse. To refine it, we first resize the segmented image to the original scale, once more using interpolation, and apply the three following refinement steps. First, we extract the connected components (pixels of same label connected through an 8-neighborhood) in the image and map each of them to an individual class. This step is necessary to separate the regions that have similar properties but correspond to different persons or objects.

Next, we do a few random walk iterations to remove the blockiness of the region boundaries. Let S_0 be the rescaled segmentation obtained from the random walk model. We update the pixel-to-class probabilities iteratively as follows:

$$S'=(1-\alpha)QS+\alpha S_0 \quad (11)$$

In our experiments, this process typically converges after 10 to 15 iterations. Since it only involves the multiplication of sparse matrices, the total runtime of this process is not significant.

Finally, the last refinement step consists in merging the regions smaller than a given number of pixels to larger neighbor regions. Let y_k be the class of a region that does not satisfy the minimum size constraint. We compute a merging score h_{kl} for every candidate region yl, defined as the sum of the weights between the pixels of class y_k and class y_l, and merge y_k to the candidate region with the highest score. These class-to-class merging scores can be evaluated efficiently as $H = S^T WS$.

3. Conclusion

We presented a new method for the unsupervised segmentation problem. This method uses a generative model based on random walks to map each pixel in an image to the most probable label in a local neighborhood. A GMM approach is used to estimate the initial class priors. To make our method more robust to this step, we proposed an iterative process in which the class parameters and segmentation probabilities are recomputed alternatively. We also presented a refinement strategy to improve the coarse segmentation obtained at the previous step. Due to its efficiently, this strategy can be applied at the original image scale without affecting the runtimes. In future works, we would like to investigate the use of other pixel features, and its impact on the segmentation efficiency and accuracy.

References

[1] C. A. Cocosco, A. P. Zijdenbos, and A. C. Evans, "A fully automatic and robust brain MRI tissue classification method," *Medical Image Analysis*, vol. 7, no. 4, pp. 513–527, 2003.
[2] T. Kim, K. Lee, and S. Lee, "Generative image segmentation using random walks with restart," in *European Conference on Computer Vision (ECCV)*, 2008, pp. 264–275.
[3] M. Magnate, "MDS-based segmentation model for the fusion of contour and texture cues in natural images," *Computer Vision and Image Understanding*, vol. 116, no. 9, pp. 981–990, 2012.
[4] "Segmentation by fusion of histogram-based K-means clusters in different color spaces," *IEEE Transactions on image processing*, vol. 17, no. 5, pp. 780–787, 2008.
[5] "A de-texturing and spatially constrained K-means approach for image segmentation," *Pattern Recognition Letters*, vol. 32, no. 2, pp. 359–367, 2011.
[6] G. Sfikas, C. Nikou, and N. Galatsanos, "Edge preserving spatially varying mixtures for image segmentation," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–7.
[7] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, 2002.
[8] Q. Wang, "HMRF-EM-image: Implementation of the hidden markov random field model and its expectation-maximization algorithm," *CoRR*, vol. abs/1207.3510, 2012.
[9] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–156, 1998.
[10] M. Magnate, "A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation," *Image Processing, IEEE Transactions on*, vol. 19, no. 6, pp. 1610–1624, 2010.
[11] "MDS-based multiresolution nonlinear dimensionality reduction model for color image segmentation," *Neural Networks, IEEE Transactions on*, vol. 22, no. 3, pp. 447–460, 2011.
[12] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.
[13] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 888–905, 2000.