DOMAIN CONVERSION WITH LOCAL POSTERIORS FOR IMAGE SEGMENTATION

EunSang Bak[†] *and Kayvan Najarian*[‡]

[†]Electrical and Computer Engineering Department, [‡]College of Information Technology,
University of North Carolina at Charlotte
9201 University City Blvd, Charlotte, NC 28223, U.S.A.
Email: ebak@uncc.edu, knajaria@uncc.edu

ABSTRACT

The estimates of the posterior probabilities of the attributes in the image are widely used as criteria for image segmentation. The methods using this measure, however, suffer from intrinsic errors that occur around the boundary between regions. The errors are caused by estimating the posterior probabilities over the entire image. To resolve this problem, we define novel local posterior probabilities to better capture the local characteristics and then use them in an iterative segmentation process. Furthermore, the image itself is converted to another image in a new domain by a domain conversion method. It is shown that the converted image in the new domain is less susceptible to intrinsic errors.

1. INTRODUCTION

There have been numerous attempts to take advantage of the local information to separate inhomogeneous regions in an image. In such approaches, several different criteria have been introduced to measure the similarity or dissimilarity of pixels. From the statistical point of view, two of types of probabilities are often used to determine how likely a particular attribute is associated with a particular region.

One of these approaches considers the local conditional probabilities [2,3] and updates the segmented image by maximizing this measure. The other approach [1] makes use of so called contextual information. On the other hand, spatial thresholding method [4] was introduced to overcome the drawbacks of simple thresholding methods. The main idea in this approach is also based on exploiting the local spatial information to define segmentation criterion without excessive computations.

In this paper, we define a local likelihood approach that takes advantage of both global and local statistical properties. The presented method incorporates the knowledge of local priors and provides a relatively simple local posterior function. Furthermore, we introduce a domain conversion method by which the actual image is converted to another form of two-dimensional domain that is not in the gray level range. The image created in the new domain is generated by a linear combination of the neighboring pixels, and as a result, the local information around a site in the image is merged into a single site in the converted image. This further improves the results of our proposed segmentation method.

The paper is organized as follows: In Section 2, the local likelihoods and local posteriors are defined. In Section 3, a domain conversion method is proposed and it is shown how this method is adapted with the local posterior function in an iterative segmentation process. The iterative segmentation method is briefly explained in Section 4. The experimental results are presented in Section 5, which is followed by the conclusions in Section 6.

2. LOCALIZED POSTERIORS

Two types of images are often considered in image segmentation. One category is the observable images denoted by $\boldsymbol{Y} = \{\boldsymbol{Y}_s, s \in \boldsymbol{S}\}$ where $\boldsymbol{Y}_s$ is the intensity at a site s. The other is the hidden label images denoted by $\boldsymbol{X} = \{\boldsymbol{X}_s, s \in \boldsymbol{S}\}$ where $\boldsymbol{X}_s$ is the label at a site s.

In order to relate $\boldsymbol{X}$ and $\boldsymbol{Y}$ images statistically, we make the assumptions made in [1]. First, the random variables $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n$ in $\boldsymbol{Y}$ are assumed to be conditionally independent where each $\boldsymbol{Y}_i$ has the same known conditional density function. Second, we assume that $\boldsymbol{Y}_i$ depends only on $\boldsymbol{X}_i$. These assumptions simplify the joint conditional probability (2.1) over the given neighborhood system. The neighborhood N_s represents a set of sites in the 3×3 window without the center site while the $\widetilde{N}_s$ includes the center site.

$$p\left(\boldsymbol{Y}_{\widetilde{N}_s} \mid \boldsymbol{X}_{\widetilde{N}_s}\right) = \prod_{s \in \widetilde{N}_s} p\left(\boldsymbol{Y}_s \mid \boldsymbol{X}_s\right) \tag{2.1}$$

2.1. Local Likelihoods

In a typical likelihood function, the likelihood of each pixel is measured by the probability density function estimated over the entire image. It means that the likelihood ignores the location of each pixel and hence disregards the local information. Assuming the spatial continuity, the pixels close to each other tend to have large probabilities (close to 1) of being included in the same region. In this paper, we take the local information into consideration and produce a new function, "local likelihood" hereafter, to address this issue. The local area where the local likelihood is computed is defined as the same area of the given neighborhood system. Eq. (2.2) defines the local likelihood.

$$\begin{aligned} l\left(\boldsymbol{X}_s;\boldsymbol{Y}_{\tilde{N}_s}\right) &= p\left(\boldsymbol{Y}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right) \\ &= \sum_{\boldsymbol{X}_{\tilde{N}_s}\in\tilde{\boldsymbol{N}}} p\left(\boldsymbol{Y}_{\tilde{N}_s} \mid \boldsymbol{X}_{\tilde{N}_s},\boldsymbol{X}_s\right)P\left(\boldsymbol{X}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right) \end{aligned} \tag{2.2}$$

Eq. (2.1) being considered, the local likelihood function can be decomposed into weight term and conventional likelihood function as:

$$\begin{aligned} l\left(\boldsymbol{X}_s;\boldsymbol{Y}_{\tilde{N}_s}\right) &= P\left(\boldsymbol{Y}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right) \\ &= \sum_{\boldsymbol{X}_{\tilde{N}_s}\in\boldsymbol{N}}\left[\prod_{i\in\tilde{N}_s} p\left(\boldsymbol{Y}_i \mid \boldsymbol{X}_i\right)P\left(\boldsymbol{X}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right)\right] \\ &= \sum_{\boldsymbol{X}_{\tilde{N}_s}\in\tilde{\boldsymbol{N}}}\left[P\left(\boldsymbol{X}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right)\prod_{i\in N_s} p\left(\boldsymbol{Y}_i \mid \boldsymbol{X}_i\right)p\left(\boldsymbol{Y}_s \mid \boldsymbol{X}_s\right)\right] \\ &= \tau\left(\boldsymbol{X}_s\right)l\left(\boldsymbol{X}_s;\boldsymbol{Y}_s\right) \end{aligned} \tag{2.3}$$

where

$$\tau\left(\boldsymbol{X}_s\right) = \sum_{\boldsymbol{X}_{N_s}\in\boldsymbol{N}} P\left(\boldsymbol{X}_{N_s} \mid \boldsymbol{X}_s\right)\prod_{i\in N_s} p\left(\boldsymbol{Y}_i \mid \boldsymbol{X}_i\right)$$

As can be seen, the local likelihood contains the likelihood function. This means that the local likelihood can be interpreted as the weighted version of the likelihood and the weight term $\tau(\boldsymbol{X}_s)$ adjusts the likelihood appropriately to represent the local characteristics. Consequently, the local likelihood is the likelihood multiplied by $\tau(\boldsymbol{X}_s)$ which is the probability of $\boldsymbol{Y}_{N_s}$ given $\boldsymbol{X}_{N_s}$ averaged over all possible configurations of $\boldsymbol{X}_{N_s}$.

When the image is severely degraded by noise, the label field of each site may not necessarily show the most possible label map. However, as the iterative segmentation process continues, the label image is getting updated and the label field of each site is more likely to be the most possible label map. This intuitive fact gives rise to another local likelihood as in (2.4).

$$\begin{aligned} l\left(\boldsymbol{X}_s;\boldsymbol{Y}_{\tilde{N}_s}\right) &= p\left(\boldsymbol{Y}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right) \\ &\approx P\left(\boldsymbol{X}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right)p\left(\boldsymbol{Y}_s \mid \boldsymbol{X}_s\right)\prod_{i\in N_s} p\left(\boldsymbol{Y}_i \mid \boldsymbol{X}_i\right) \\ &= P\left(\boldsymbol{Y}_{\tilde{N}_s},\boldsymbol{X}_{\tilde{N}_s} \mid \boldsymbol{X}_s\right) \end{aligned} \tag{2.4}$$

The local likelihood in (2.4) indicates how likely $\boldsymbol{X}_s$ is to be a current label when not only its intensity at site s is $\boldsymbol{X}_s$, but also its intensity field of the neighborhood is $\boldsymbol{Y}_{\tilde{N}_s}$, taking into consideration the label field $\boldsymbol{X}_{\tilde{N}_s}$. This results in significant improvement in the overall segmentation process, especially after some iterations (when each of the label field begins to have a reliable label map.)

2.2. Local Posteriors

As a posterior probability is generated in association with a prior probability, the local likelihood can be extended to the corresponding local posterior by defining the local prior. Evidently, the accuracy of the prior knowledge is crucial to make a good segmentation.

The determination of the size of the local area which could possibly contain enough information to estimate the local prior is to the most part at our disposal. More specifically, this size is not necessarily the same as the size of the neighborhood system. Therefore, the local posterior probability can be represented as:

$$\begin{aligned} P\left(\boldsymbol{X}_s \mid \boldsymbol{Y}_{\tilde{N}_s}\right) &= \frac{l\left(\boldsymbol{X}_s;\boldsymbol{Y}_{\tilde{N}_s}\right)P\left(\boldsymbol{X}_s, s\in\hat{N}_s\right)}{\sum_{\boldsymbol{X}_s\in\Lambda_{\boldsymbol{X}_s}} l\left(\boldsymbol{X}_s;\boldsymbol{Y}_{\tilde{N}_s}\right)P\left(\boldsymbol{X}_s, s\in\hat{N}_s\right)} \\ &\propto W\left(\boldsymbol{X}_s\right)l\left(\boldsymbol{X}_s;\boldsymbol{Y}_s\right)P\left(\boldsymbol{X}_s, s\in\hat{N}_s\right) \end{aligned} \tag{2.5}$$

where $\hat{N}_s$ is a set of indices of local area in which the prior probabilities are estimated.

3. DOMAIN CONVERSION ANALYSIS

Two different regions can be discriminated from each other based on the ratio of their probability distributions. In particular, when the variance stationary process is concerned, the ratio is described by a simple linear equation.

3.1. Linear Discrimination Function

Suppose there are two different multivariate normal distributions. Each distribution represents a specific region in the image. During the segmentation, a pixel is assigned to the region to which the pixel is more likely to belong statistically. The linear discrimination function (3.1), which is the logarithm of the ratio of the two probability functions, provides a criterion to this assignment problem.

$$\ln\left(\frac{p_1}{p_2}\right) = D_1 - D_2 \tag{3.1}$$

where $D_i = \left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right)^T \boldsymbol{y} - \frac{1}{2}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i\right)^T \boldsymbol{\mu}_i$. When y is a scalar, D_i is reduced to $(\mu_i / \sigma^2)y - (\mu_i^2 / 2\sigma^2)$.

If the random vector $\boldsymbol{y}$ in i^{th} region is m-dimensional and each of the m random variables has the same mean and variance, the discrimination function can be rewritten as:

$$D_i = \frac{\mu_i}{\sigma^2}\left(\boldsymbol{C}^{-1}\boldsymbol{1}\right)^T \boldsymbol{y} - \frac{\mu_i^2}{2\sigma^2}\left(\boldsymbol{1}^T \boldsymbol{C}^{-1}\boldsymbol{1}\right), \tag{3.2}$$

where $\boldsymbol{C}$ is a correlation coefficient matrix and $\boldsymbol{1}$ is an ($m \times 1$) column vector. After mathematical expansions, it can be simplified to:

$$D_i = \frac{\mu_i}{\sigma^2}\left(\sum_k \beta_k y_k\right) - \frac{\mu_i^2}{2\sigma^2}\left(\sum_k \beta_k\right), \tag{3.3}$$

where $\boldsymbol{y} = [y_0, \ldots, y_{m-1}]^T, \boldsymbol{C}^{-1}\boldsymbol{1} = \boldsymbol{\beta}$ and $\boldsymbol{\beta} = [\beta_0, \ldots, \beta_{m-1}]^T$.

Let us define another random variable w as a linear combination of the y_k's with coefficients β_k's. Assuming that the y_k's are independent and normally distributed, the mean and the variance of the linear combination of the random variables y_k's are given in (3.4) and (3.5).

$$E\{w\} = E\left\{\sum_k \beta_k y_k\right\} = \sum_k \beta_k E\{y_k\} = \mu_i \sum_k \beta_k = \mu_i^* \tag{3.4}$$

$$\begin{aligned}\text{Var}\{w\} &= \text{Var}\left\{\sum_k \beta_k y_k\right\} = \sum_i \sum_j \beta_i \beta_j \text{Cov}\{y_i, y_j\} \\ &= \sigma^2 \boldsymbol{\beta}^T \boldsymbol{C}\boldsymbol{\beta} = \sigma^2 \boldsymbol{1}^T \boldsymbol{C}^{-1}\boldsymbol{C}\boldsymbol{C}^{-1}\boldsymbol{1} = \sigma^2 \sum_i \beta_i = \sigma^{2*}\end{aligned} \tag{3.5}$$

With these parameters, the probability distribution of w becomes $N(\mu_i^*, \sigma^{2*})$, and the corresponding linear discrimination function becomes (3.6).

$$\begin{aligned}D_i(w) &= \left(\frac{\mu_i^*}{\sigma^{*2}}\right) w - \frac{\mu_i^{*2}}{2\sigma^{*2}} \\ &= \frac{\mu_i \sum_k \beta_k}{\sigma^2 \sum_k \beta_k} w - \frac{\mu_i \sum_k \beta_k}{2\sigma^2 \sum_k \beta_k}\left(\mu_i \sum_k \beta_k\right) = \frac{\mu_i}{\sigma^2} w - \frac{\mu_i^2}{2\sigma^2}\left(\sum_k \beta_k\right).\end{aligned} \tag{3.6}$$

Comparing equations (3.3) and (3.6), one can see that the linear discrimination function for multivariate $\boldsymbol{y}$ containing random variables with the same mean and variance is equivalent to the linear discrimination function for a new univariate w which is a linear combination of the components of the multivariate. This result is very helpful in finding a method to reduce the intrinsic errors of the likelihood method.

3.2. Domain Conversion

In the proposed method, an individual pixel value is converted to a random vector $\boldsymbol{y}$ whose components are the neighboring pixels in the specified neighborhood system. After calculating the correlation coefficient matrix of the random vector $\boldsymbol{y}$, another image $\boldsymbol{W}$ is created in the new domain (from the original gray level image $\boldsymbol{Y}$) by a linear combination of the neighboring pixels. The coefficients in the linear combination are computed from the inverse of the correlation coefficient matrix. Because of the normal distribution of y's in $\boldsymbol{Y}$, the distribution of w's in $\boldsymbol{W}$ is also normal.

Every site in the $\boldsymbol{W}$ image has a value which is different from the gray level in $\boldsymbol{Y}$. The local likelihood and local posterior probabilities are now calculated on each site in the **W** image, and then the sites are assigned to the region that gives a larger value of the local posterior probability.

4. ITERATIVE SEGMENTATION PROCESS

Given the image $\boldsymbol{W}$, and the provisional estimate $\boldsymbol{X}^{(i-1)}$ of the label image, the $\boldsymbol{X}^{(i-1)}$ is updated to $\boldsymbol{X}^{(i)}$ by maximizing the local posterior probability in the i^{th} iteration. The initial estimate $\boldsymbol{X}^{(0)}$ of the label image is generated by clustering feature vectors that consists of mean and standard deviation of the given local area using k-means algorithm. During the first several iterations, the local likelihood (2.3) is used and then is replaced by (2.4). The iterations are terminated if the condition of the process satisfies the convergence criterion, and the last $\boldsymbol{X}^{(i)}$ becomes the final segmented image.

5. EXPERIMENTAL RESULTS

We made different synthetic images. All of these images contain two different regions; one is background and the other contains the objects that are supposed to be separated from the background. Background has a mean value 120, the objects have 150 and both have the same standard deviation of 20. We use the first order neighborhood system and the local priors are estimated in the 8×8 windows.

We evaluated the performance of the proposed method with five (Art, Chess, Circle, Geometry, Target) test images. As you can see in the table 1, the method with domain conversion (L.P+D.C) gives better result in terms of the rate of convergence. Even though the final error rate is similar in both methods, the domain conversion makes the algorithm converge more rapidly. The method without domain conversion (L.P) seems to converge around the 12^{th} iteration, while the other method (L.P+D.C) around the 5^{th} iteration, which is twice as fast as in the other. Fig. 1 and 2 shows the variation of error rate with two test images visually.

6. CONCLUSION

In this paper, we proposed a novel local posterior approach empowered by a domain conversion method to obtain a better and faster image segmentation technique. Furthermore, the framework of the iterative segmentation process with local posteriors could be adapted to other applications

7. REFERENCES

[1] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 48, no. 3, pp. 259-302, 1986.

[2] T. Kurita, N. Otsu and N. Abdelmalek, "Maximum likelihood thresholding based on population mixture models," *Pattern Recognition*, vol. 25, no. 10, pp. 1231-1240, 1992.

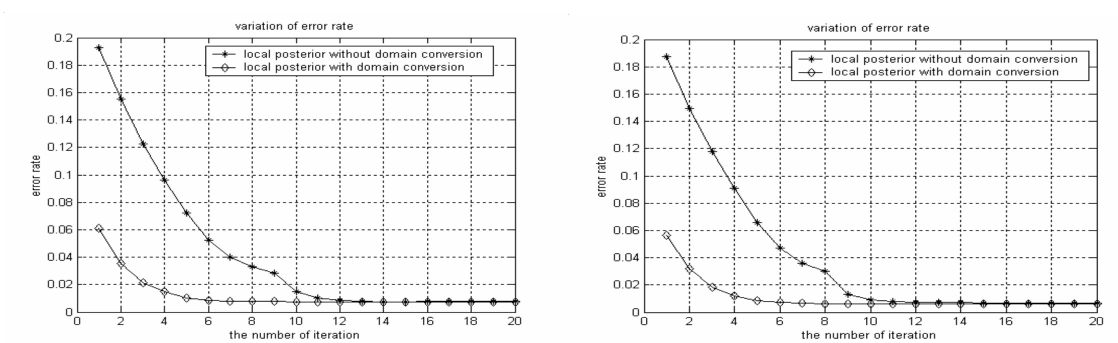
[3] C.K. Leung and F.K. Lam, "Maximum a posteriori spatial probability segmentation," *IEE Proceedings: Vision, Image and Signal Processing,* vol. 144, no. 3, pp. 161-167, June 1997.

[4] K.V. Mardia and T.J. Hainsworth, "A spatial thresholding method for image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, no. 6, pp.919-927, Nov. 1988.

Table 1. Variation of Error Rate on Iterations.

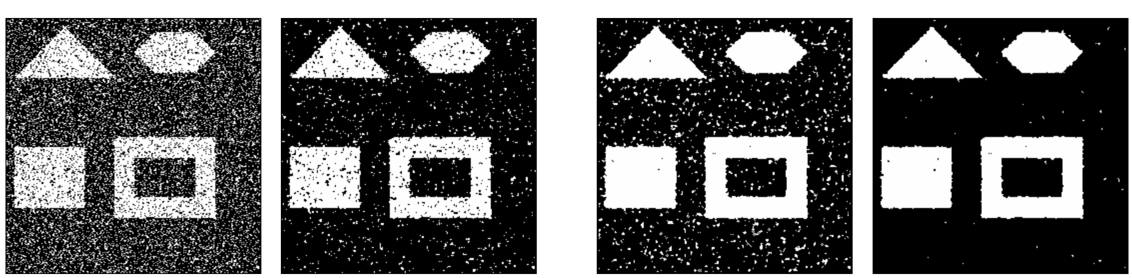
		1	2	3	4	5	10	15	20
Art	L.P	0.1730	0.1351	0.1055	0.0818	0.0631	0.0062	0.0054	0.0053
	L.P+D.C	0.0159	0.0102	0.0072	0.0065	0.0063	0.0062	0.0062	0.0062
Chess	L.P	0.1814	0.1431	0.1130	0.0876	0.0671	0.0295	0.0262	0.0262
	L.P+D.C	0.0385	0.0350	0.0361	0.0361	0.0359	0.0357	0.0357	0.0357
Circle	L.P	0.1849	0.1500	0.1182	0.0927	0.0696	0.0090	0.0059	0.0059
	L.P+D.C	0.0449	0.0212	0.0108	0.0072	0.0049	0.0037	0.0037	0.0037
Geometry	L.P	0.1926	0.1553	0.1227	0.0960	0.0721	0.0150	0.0076	0.0076
	L.P+D.C	0.0609	0.0351	0.0214	0.0148	0.0104	0.0076	0.0075	0.0075
Target	L.P	0.1875	0.1492	0.1180	0.0907	0.0658	0.0091	0.0070	0.0069
	L.P+D.C	0.0566	0.0322	0.0184	0.0118	0.0083	0.0063	0.0063	0.0063

(L.P: Local Posterior, D.C: Domain Conversion.)



a) Geometry (b) Target

Fig. 1. Variation of Error Rate.



(a) local posterior without domain conversion (b) local posterior with domain conversion

Fig. 2. Target (labeled images at iteration: 1^{st} ,5^{th}).