

# Edge preserving spatially varying mixtures for image segmentation

Giorgos Sfikas

Christophoros Nikou \*

Nikolaos Galatsanos

University of Ioannina  
Department of Computer Science  
45110, Ioannina, Greece  
{sfikas,cnikou,galatsanos}@cs.uoi.gr

## Abstract

*A new hierarchical Bayesian model is proposed for image segmentation based on Gaussian mixture models (GMM) with a prior enforcing spatial smoothness. According to this prior, the local differences of the contextual mixing proportions (i.e. the probabilities of class labels) are Student's t-distributed. The generative properties of the Student's t-pdf allow this prior to impose smoothness and simultaneously model the edges between the segments of the image. A maximum a posteriori (MAP) expectation-maximization (EM) based algorithm is used for Bayesian inference. An important feature of this algorithm is that all the parameters are automatically estimated from the data in closed form. Numerical experiments are presented that demonstrate the superiority of the proposed model for image segmentation as compared to standard GMM-based approaches and to GMM segmentation techniques with "standard" spatial smoothness constraints.*

## 1. Introduction

Clustering-based image segmentation methods rely on arranging data into groups having common characteristics [19]. During the last decade, the main research directions in the relevant literature are focused on mixture model [16, 14], graph theoretic approaches [17, 8, 21], methods based on the mean shift algorithm [6, 4] and rate distortion theory techniques [20].

Modeling the probability density function (pdf) of pixel attributes (e.g. intensity, texture) with finite mixture models (FMM) [1] is also a natural way to cluster data be-

cause it automatically provides a grouping based on the components of the mixture that generated them. Furthermore, the likelihood of a FMM is a rigorous metric for clustering performance [1]. The parameters of the FMM model with Gaussian components can be estimated very efficiently through maximum likelihood (ML) estimation using the Expectation-Maximization (EM) algorithm [7]. Furthermore, it can be shown that Gaussian components allow efficient representation of a large variety of pdf. Thus, Gaussian mixture models (GMM), are commonly employed in image segmentation tasks.

A drawback of the standard ML approach for GMM image segmentation is that commonality of location is not taken into account when grouping the data. In other words, the prior knowledge that adjacent pixels most likely belong to the same cluster is not used. To overcome this shortcoming, spatial smoothness constraints have been imposed. A common approach is the use of an MRF. Many MRF variants have been proposed, see for example [11]. However, determination of the amount of the imposed smoothness automatically requires knowledge of the normalization constant of the MRF. Since this is not known analytically, learning strategies were proposed [22, 9].

Research efforts in imposing spatial smoothness for image segmentation can be grouped into two categories. In the methods of the first category, spatial smoothness is imposed on the discrete *hidden* variables of the FMM that represent class labels [12, 21]. These approaches may be categorized in a more general area involving simultaneous image recovery and segmentation which is better known as *image modeling* [15]. More specifically, spatial regularization is achieved by imposing a discrete MRF on the classification labels of neighboring pixels that penalizes solutions where neighboring pixels belong to different classes. Inference in this category of models is non trivial and generally performed through EM-like alternating optimization, Markov Chain Monte Carlo (MCMC) or inexact variational inference.

---

\*This work was co-funded by the European Union in the framework of the project Support of Computer Science Studies in the University of Ioannina of the Operational Program for Education and Initial Vocational Training of the 3rd Community Support Framework of the Hellenic Ministry of Education, funded by 20% from national sources and by 80% from the European Social Fund (ESF).

In the second category of methods, the MRF-based smoothness constraint is not imposed on the labels but on the contextual mixing proportions, that is on the probabilities of the pixel labels. This model is called spatially variant finite mixture model (SVFMM) [16] and avoids the inference problems of DMRFs. For instance, in [14], a new family of smoothness priors was proposed for the contextual mixing proportions based on the Gauss-Markov random fields that takes into account cluster statistics, thus enforcing different smoothness strength for each cluster. The model was also refined to capture information in different spatial directions. In these models, maximum *a posteriori* (MAP) estimation of the contextual mixing proportions via the MAP-EM algorithm is possible. However, the main disadvantage is that smoothness is imposed in the neighbourhood of each pixel without taking into account that the respective pixel may be an edge pixel or its neighbourhood consists of edge pixels.

In this paper, we present a new hierarchical Bayesian model for mixture model-based image segmentation with spatial constraints. This model assumes that the local differences of the *contextual mixing proportions* follow a Student's *t*-distribution. The generative model of Student's *t*-distribution contains two levels. The lower level is a Gaussian pdf with precision (inverse variance) varying with each pixel and the higher level a Gamma pdf. The varying precisions with each pixel of the Gaussians of this model capture the *local image variations* and thus allow the smoothness constraints to incorporate the image edge structure.

A MAP-EM algorithm was used for Bayesian inference with this model. An important feature of this algorithm is that *all the necessary parameters* are estimated from the data. Thus, the proposed segmentation algorithm is *automatic* in the sense that it does not require empirical selection of parameters like other state-of-the-art methods (n-cuts, mean-shift).

The model was extensively evaluated on the 300 images of the Berkeley image data base and was compared with other GMM based methods that do not require parameter selection. More specifically, it compared favorably to standard GMM and to GMM with "standard" spatial smoothness constraints [14].

## 2. The edge preserving spatially varying GMM

The K-kernel spatially varying GMM [16, 14] differs from the standard GMM [1] in the definition of the mixing proportions. More precisely, in the SVGMM, each pixel  $x^n$ ,  $n = 1, \dots, N$  has a distinct vector of mixing proportions denoted by  $\pi_j^n$ ,  $j = 1, \dots, K$ , with  $K$  being the number of Gaussian kernels. We call these parameters *contextual mixing proportions* to distinguish them from the mixing proportions of a standard GMM. Hence, the probability of a

distinct pixel is expressed by:

$$f(x^n; \pi, \mu, \Sigma) = \sum_{j=1}^K \pi_j^n \mathcal{N}(x^n; \mu_j, \Sigma_j), \quad (1)$$

where  $0 \leq \pi_j^n \leq 1$ ,  $\sum_{j=1}^K \pi_j^n = 1$  for  $j = 1, 2, \dots, K$  and  $n = 1, 2, \dots, N$ ,  $\mu_j$  are the Gaussian kernel means and  $\Sigma_j$  are the Gaussian kernel covariance matrices.

Generally, in image processing and computer vision, we assume that, conditioned on a hidden variable  $Z$ , pixels  $X = \{x^1, x^2, \dots, x^N\}$  are independent and Gaussian-distributed:

$$p(X|Z) = \prod_{j=1}^K \prod_{n=1}^N \mathcal{N}(x^n | \mu_j, \Sigma_j)^{z_j^n}, \quad (2)$$

where the set of  $N \times K$  latent variables  $Z = \{z_j^n\}_{n=1..N, k=1..K}$  is introduced to make inference tractable for the model. The  $Z$  variables are distributed multinomially:

$$p(Z|\Pi) = \prod_{j=1}^K \prod_{n=1}^N (\pi_j^n)^{z_j^n}, \quad (3)$$

where each  $z_n$  is a binary vector, with  $z_j^n = 1$  if datum  $n$  is generated by the  $j$ -th kernel and  $z_j^n = 0$  otherwise.

Considering the set of *contextual mixing proportions*  $\Pi$  as random variables and assuming a proper prior, we can incorporate the intuitive fact that neighbouring pixels are more likely to share the same class label. We assume a Markov random field on  $\Pi$ , which equivalently means that  $\Pi$  is governed by a Gibbs distribution [11] generally expressed by:

$$p(\Pi) \propto \prod_C e^{-\psi_c(\Pi)},$$

where  $\psi_c$  is a function on clique  $c$ , called *clique potential* function in the literature, and the product is over all minimal cliques of the Markov random field.

In the herein proposed model, we consider clique potential functions imposing local differences of *contextual mixing proportions* to follow a univariate Student's *t*-distribution.

A  $d$ -dimensional random variable  $X$  follows a multivariate *t*-distribution,  $X \sim St(\mu, \Sigma, \nu)$ , with mean  $\mu$ , positive definite, symmetric and real  $d \times d$  covariance matrix  $\Sigma$  and has  $\nu \in [0, \infty)$  degrees of freedom when [1], given the weight  $u$ , the variable  $X$  has the multivariate normal distribution with mean  $\mu$  and covariance  $\Sigma/u$ :

$$X|\mu, \Sigma, u \sim \mathcal{N}(\mu, \Sigma/u), \quad (4)$$

and the weight  $u$  follows a Gamma distribution parameterized by  $\nu$ :

$$u \sim \mathcal{G}(\nu/2, \nu/2). \quad (5)$$

Integrating out the weights from the joint density leads to the density function of the marginal distribution:

$$p(x; \mu, \Sigma, \nu) = \frac{\Gamma(\frac{\nu+d}{2}) |\Sigma|^{-\frac{1}{2}}}{(\pi\nu)^{\frac{d}{2}} \Gamma(\frac{\nu}{2}) [1 + \nu^{-1} \delta(x, \mu; \Sigma)]^{\frac{\nu+d}{2}}} \quad (6)$$

where  $\delta(x, \mu; \Sigma) = (x - \mu)^T \Sigma^{-1} (x - \mu)$  is the Mahalanobis squared distance and  $\Gamma$  is the Gamma function [1]. It can be shown that for  $\nu \rightarrow \infty$  the Student's  $t$ -distribution tends to a Gaussian distribution with covariance  $\Sigma$ . Also, if  $\nu > 1$ ,  $\mu$  is the mean of  $X$  and if  $\nu > 2$ ,  $\nu(\nu - 2)^{-1} \Sigma$  is the covariance matrix of  $X$ . Therefore, the family of  $t$ -distributions provides a heavy-tailed alternative to the normal family with mean  $\mu$  and covariance matrix that is equal to a scalar multiple of  $\Sigma$ , if  $\nu > 2$  (fig. 1)[1].

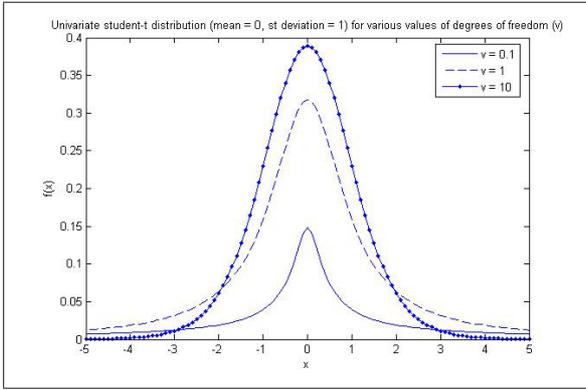


Figure 1. The Student's  $t$ -distribution for various degrees of freedom. As  $\nu \rightarrow \infty$  the distribution tends to a Gaussian. For small values of  $\nu$  the distribution has heavier tails than a Gaussian.

Therefore, the clique potential functions are properly defined in order to impose:

$$\pi_j^n - \pi_j^k \sim \mathcal{St}(0, \beta_{jd}^2, \nu_{jd}), \quad \forall n, j, d, \forall k \in \gamma_d(n). \quad (7)$$

As it can be observed in eq. (7), we introduce  $K \times D$  different  $t$ -distributions, amounting to an equal number of parameter sets,  $\{\beta_{jd}^2, \nu_{jd}\}_{j=1..K, d=1..D}$ . In eq. (7),  $D$  stands for the number of a pixel's neighbourhood adjacency types, and  $\gamma_d(n)$  is the set of neighbours of pixel indexed  $n$ , with respect to the  $d^{\text{th}}$  adjacency type. In our model, we assume 4 neighbours for each pixel, and partition the corresponding adjacency types into horizontal and vertical, thus, setting  $D = 2$ . This variability of parameter sets aims to capture the fact that smoothness statistics may vary along clusters and spatial directions [14]. Therefore, the joint distribution on  $\Pi$  is given by:

$$p(\Pi; \beta, \nu) = \prod_{d=1}^D \prod_{j=1}^K \prod_{n=1}^N \prod_{k \in \gamma_d(n)} \mathcal{St}(\pi_j^n | \pi_j^k; \beta_{jd}^2, \nu_{jd}). \quad (8)$$

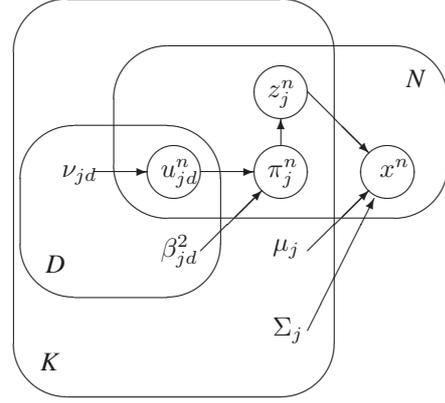


Figure 2. Graphical model for the edge preserving spatially variant Gaussian mixture model. Superscript  $n \in [1, N]$  denotes pixel index, subscript  $j \in [1, K]$  denotes kernel (segment) index and  $d \in [1, D]$  describes the neighbourhood direction type.

Following the definition of the  $t$ -distribution in eq. (4) and eq. (5) we introduce the latent variables  $U = \{u_{jd}^n\}_{n=1..N, j=1..K, d=1..D}$  and the distribution of the differences of local *contextual mixing proportions* becomes:

$$\pi_j^n - \pi_j^k \sim \mathcal{N}(0, \beta_{jd}^2 / u_j^{nk}),$$

$$u_j^{nk} \sim \mathcal{G}(\nu_{jd}/2, \nu_{jd}/2), \quad \forall n, j, d, \quad \forall k \in \gamma_d(n). \quad (9)$$

This generative model (fig. 2), apart from being tractable using the EM algorithm, as it will be demonstrated in the next section, allows better insight on our assumption of Student- $t$  cliques. Since  $u_j^{nk}$  depends on datum indexed by  $n$ , each weight difference in the MRF can be described by a different instance of a Gaussian distribution. Therefore, as  $u_j^{nk} \rightarrow +\infty$  the distribution tightens around zero, and forces neighboring *contextual mixing proportions* to be smooth. On the other hand, when  $u_j^{nk} \rightarrow 0$  the distribution tends to be uninformative, and forces no smoothness. This is a desirable property when there exists an edge between neighboring pixels. Thus, the  $U$ -variable maps provide a very detailed description of the image edge structure. Furthermore, they may be considered as a continuous generalization of the binary line-process variable idea in [2, 11].

### 3. Bayesian inference using MAP-EM

As shown in the graphical model in figure 2, the unknowns  $\Psi = \{\mu, \Sigma, \beta, \nu\}$  are considered as *parameters* and will be estimated in the M-step of the EM algorithm that follows. The  $Z, U$  are *hidden* random variables and will be inferred in the E-step of the same algorithm. The unknown quantities  $\Pi$ , although being random variables, they are treated as parameters and are estimated in the M-step. This is the reason we refer to this algorithm as MAP-EM [1].

To perform model inference, the evidence  $\ln p(X, \Pi; \Psi)$  with respect to the model parameters  $\Psi = \{\mu, \Sigma, \beta, \nu\}$  and the *contextual mixing proportions*  $\Pi$  has to be optimized. In EM terminology [7] this is the incomplete data log-likelihood while the complete log-likelihood is expressed by  $\ln p(X, \Pi, Z, U; \Psi)$ . The conditional expectation of the complete likelihood is an important quantity in EM - it is defined as

$$\mathcal{E}_{Z^{(t)}, U^{(t)}|X, \Pi^{(t)}} \left\{ \ln p(X, \Pi, Z, U; \Psi) \right\}. \quad (10)$$

The E-step consists in computing the joint expectation of the hidden variables  $Z$  and  $U$ , with respect to current iteration parameters  $\Psi^{(t)}$  where  $t$  denotes the number of current iteration. Observing the graphical model in fig. 2, we can see that given  $X$  and  $\Pi$ ,  $Z$  and  $U$  are conditionally independent; therefore  $\mathcal{E}_{Z, U|X, \Pi}(\cdot) = \mathcal{E}_{Z|X, \Pi}\{\mathcal{E}_{U|X, \Pi}(\cdot)\}$  and we can compute these expectations separately. So we have the updates  $\forall n, j, d, \forall k \in \gamma_d(n)$ :

$$\begin{aligned} \langle z_j^n \rangle^{(t)} &= \frac{\pi_j^{n(t)} \mathcal{N}(x^n; \mu_j^{(t)}, \Sigma_j^{(t)})}{\sum_{l=1}^K \pi_l^{n(t)} \mathcal{N}(x^n; \mu_l^{(t)}, \Sigma_l^{(t)})}, \\ \langle u_j^{nk} \rangle^{(t)} &= \zeta_j^{nk(t)} / \eta_j^{nk(t)}, \\ \langle \ln u_j^{nk} \rangle^{(t)} &= \psi(\zeta_j^{nk(t)}) - \ln \eta_j^{nk(t)}, \end{aligned}$$

where  $\psi(\cdot)$  stands for the digamma function, and parameters  $\zeta, \eta$  being:

$$\begin{aligned} \zeta_j^{nk(t)} &= \frac{1}{2} \left( \nu_{jd}^{(t)} + 1 \right), \\ \eta_j^{nk(t)} &= \frac{1}{2} \left( \nu_{jd}^{(t)} + \frac{(\pi_j^{n(t)} - \pi_j^{k(t)})^2}{\beta_{jd}^{2(t)}} \right). \end{aligned}$$

Maximization of the current complete likelihood (10) must be driven with respect to the model parameters  $\Psi$  and  $\Pi$ . With some manipulation, we can rewrite it equivalently as

$$\begin{aligned} \mathcal{E}_{Z|X, \Pi} \{ \ln p(X|Z; \mu, \Lambda) \} + \mathcal{E}_{Z|X, \Pi} \{ \ln p(Z|\Pi) \} + \\ + \mathcal{E}_{U|\Pi} \{ \ln p(\Pi|U; \beta) \} + \mathcal{E}_{U|\Pi} \{ \ln p(U; \nu) \}. \end{aligned}$$

In this form, parameter optimization is straightforward. The resulting update equations make up the M-step:

$$\begin{aligned} \mu_j^{(t+1)} &= \frac{\sum_{n=1}^N \langle z_j^n \rangle^{(t)} x^n}{\sum_{n=1}^N \langle z_j^n \rangle^{(t)}}, \\ \Sigma_j^{(t+1)} &= \frac{\sum_{n=1}^N \langle z_j^n \rangle^{(t)} (x^n - \mu_j^{(t+1)})(x^n - \mu_j^{(t+1)})^T}{\sum_{n=1}^N \langle z_j^n \rangle^{(t)}}, \end{aligned}$$

$$\beta_{jd}^{2(t+1)} = \frac{\sum_{n=1}^N \sum_{k \in \gamma_d(n)} \langle u_j^{nk} \rangle^{(t)} (\pi_j^{n(t)} - \pi_j^{k(t)})^2}{\sum_{n=1}^N |\gamma_d(n)|}.$$

Moreover, the *contextual mixing proportions*  $\pi_j^n$  are computed as the roots of a quadratic equation:

$$a_j^n \left( \pi_j^{n(t+1)} \right)^2 + b_j^n \left( \pi_j^{n(t+1)} \right) + c_j^{n(t+1)} = 0 \quad (11)$$

with coefficients:

$$\begin{aligned} a_j^n &= - \sum_{d=1}^D \left\{ \beta_{jd}^{-2(t)} \sum_{k \in \gamma_d(n)} \langle u_j^{nk} \rangle^{(t)} \right\}, \\ b_j^n &= \sum_{d=1}^D \left\{ \beta_{jd}^{-2(t)} \sum_{k \in \gamma_d(n)} \langle u_j^{nk} \rangle^{(t)} \pi_j^{k(t)} \right\}, \\ c_j^n &= \frac{1}{2} \langle z_j^n \rangle^{(t)}. \end{aligned}$$

The form of the coefficients guarantees that there is always a real non negative solution. However, the solutions of eq. (11) for a given pixel, indexed by  $n$ , will not in general satisfy the constraint  $\sum_{j=1}^K \pi_j^n = 1$ . In order to get proper mixing weight vectors we perform a projection step onto the constraints subspace using the quadratic programming algorithm described in [3].

Finally, setting the derivative of (10) with respect to the degrees of freedom equal to zero we obtain  $\nu_{jd}^{(t+1)}$  as the solutions of the equation:

$$\begin{aligned} \ln(\nu_{jd}^{(t+1)}/2) - \psi(\nu_{jd}^{(t+1)}/2) + \\ + \left[ \frac{\sum_{n=1}^N \sum_{k \in \gamma_d(n)} (\langle \ln u_j^{nk} \rangle^{(t)} - \langle u_j^{nk} \rangle^{(t)})}{\sum_{n=1}^N |\gamma_d(n)|} \right] + 1 = 0 \end{aligned}$$

with  $\psi(\cdot)$  being again the digamma function.

## 4. Experimental results

In our model, parameters  $U$  play a very important role in the preservation of the boundaries between image regions. The  $U$ -variable maps for the  $j^{th}$  kernel can be considered as the edges that separate the  $j^{th}$  segment of the image from the remaining segments. To demonstrate this point we show an example in figure 3. In this example, an image is segmented into  $K = 3$  segments thus 6  $U$ -variable maps are shown. The first row of this figure shows the original and the segmented images. Then, moving from top to bottom, the  $U$ -variable maps for the three image segments, namely *sky*, *roof and shadows*, *building* are shown, respectively. The left column highlights vertical edges and the right column underpins horizontal edges. Notice that in the second row of figure 3, where the  $U$ -variable maps for segment *sky* are shown, the edges between the segment *sky* and the rest

(*roof and shadows, building*) are mainly highlighted. The edges between the other segments, (*roof and shadows and building*) are mainly highlighted in the remaining two maps. Similarly, the edges between the segments *sky* and *building* are not highlighted in the third row of images as the  $U$ -variable maps for *roof and shadows* are underpinned.



Figure 3.  $U$ -variable maps: The first row shows the original image and the segmentation for  $K = 3$  clusters; the rows below show  $U$ -variable maps (expected values of  $u_j^{n_k}$  variables). Brighter values represent lower values of  $u$ . In each row, the  $U$ -variable maps for kernel indexed by  $j = 1$  (*sky*),  $j = 2$  (*roof and shadows*) and  $j = 3$  (*building*), are shown respectively. The left column corresponds to  $u$  values computed for horizontal adjacencies, and the right column for vertical adjacencies.

In our implementation, we have used a 4-dimensional feature vector to describe the image data. It is comprised by the *Lab* color space features and the *Blobworld* contrast texture descriptor as described in [5]. Prior to segmentation, each variate has been separately normalized in order not to have dominating features. We have evaluated the proposed Student’s  $t$ -based SVGMM (St-SVGMM) segmentation scheme on the 300 images of the Berkeley image database [13]. We have applied our algorithm for different values of the number of segments ( $K = \{3, 5, 7, 10, 15, 20\}$ ). For comparison purposes, we have also experimented with the standard GMM [1] and the GMM based segmentation with “standard” smoothness constraints [14] with the same

number of components.

The obtained segmentations were quantitatively evaluated with two performance measures: the Rand index (RI) [18] and the boundary displacement error (BDE) [10]. The RI measures the consistency between the ground truth and the computed segmentation map while the BDE measures error in terms of boundary displacement with respect to the ground truth. The statistics for these measures are presented in tables 1 and 2.

Based on the theoretical properties of the Student’s  $t$ -model one might have expected that the St-SVGMM introduced erroneous boundaries that did not agree with human segmentation. Therefore it would provide a worse RI as compared to the “classical” non preserving algorithm (SVGMM) [14]. However, as observed in the statistics of the RI (table 1), the St-SVGMM outperforms the standard GMM in all cases and the SVGMM in the overwhelming majority of the different number of components.

Also, in terms of correct region boundary estimation, expressed by the BDE (table 2), the St-SVGMM outperforms the SVGMM, as it is theoretically expected. However, it also outperforms standard GMM and the difference in performance increases with the number of segments. The explanation for this behavior is that the standard GMM since it does not integrate a smoothing step it generally computes correctly the boundaries between segments (it also outperforms the SVGMM in the same median values). However, as the number of segments increases, the complexity of the image cannot be captured by a simple GMM and smoothness constraints that model the image edge structure become increasingly beneficial.

Overall, the St-SVGMM not only preserves region boundaries but also improves the correct classification rates with respect to the standard methods. Some representative segmentation examples are shown in figure 4.

## 5. Conclusion

In this paper a segmentation algorithm based on clustering with GMM is proposed. The main novelty of this work is a smoothness prior which apart from constraining adjacent pixel to belong in the same cluster captures the image edge structure. Thus, it does not enforce smoothness across segment boundaries. Another important feature of the herein proposed segmentation algorithm is that all required parameters are estimated from the data. Thus, this algorithm is *automatic* and does not require empirical parameter selection like many recent state-of-the-art segmentation algorithms. An important perspective of this study is to automatically estimate the number of components  $K$ . To this end, criteria appropriate to constrained mixtures could be conceived.

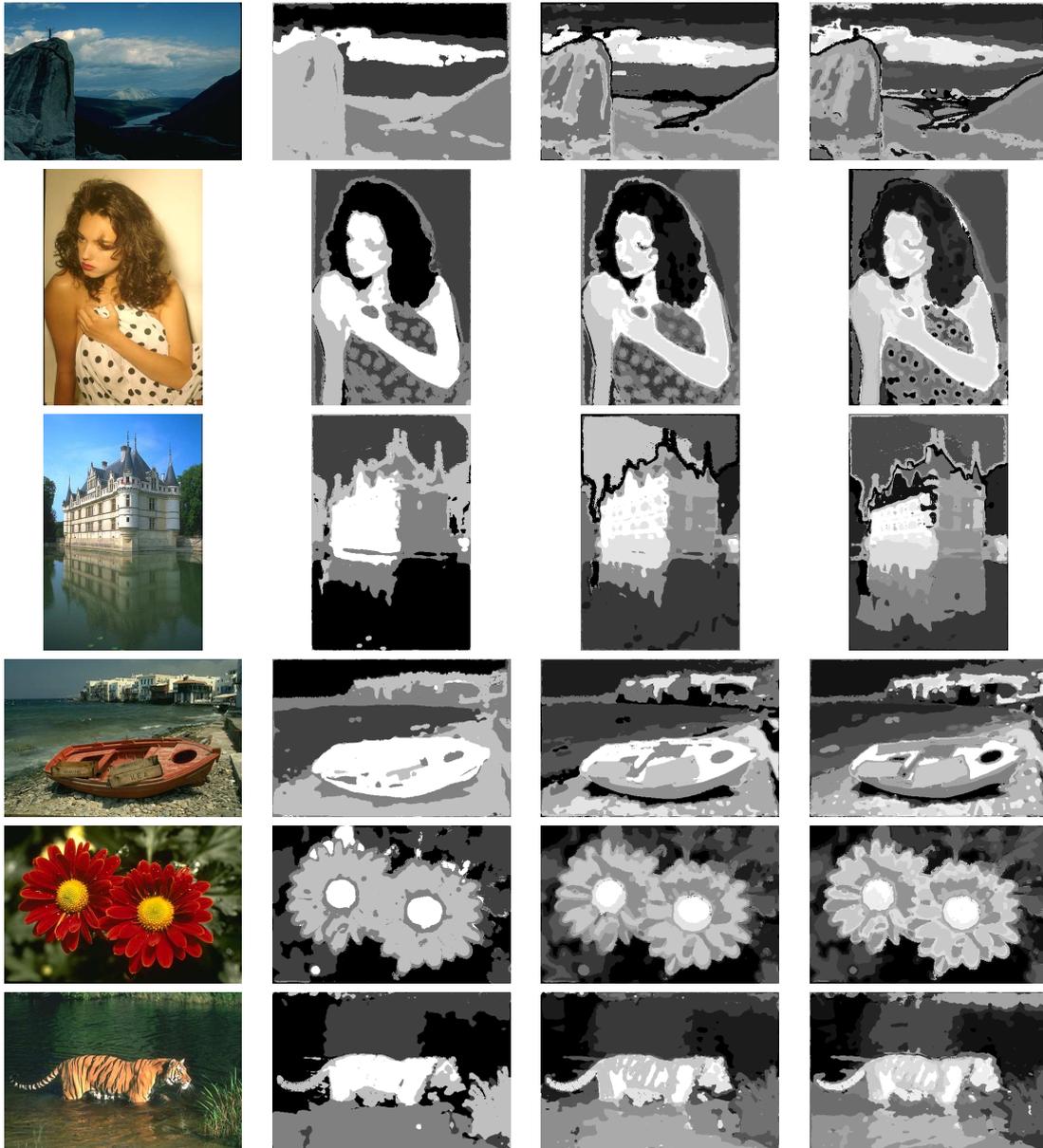


Figure 4. Segmentation examples using the proposed edge preserving spatially variant mixture. From left to right, the columns show: the original image, segmentation with  $K = 5$ ,  $K = 10$  and  $K = 15$  kernels.

Table 1. Statistics on the Rand Index (RI) over the 300 images of the Berkeley image data base for the compared methods. Higher values represent better segmentations.

GMM				SVGMM				St-SVGMM			
$K$	Mean	Median	St. dev.	$K$	Mean	Median	St. dev.	$K$	Mean	Median	St. dev.
3	0.6754	0.6796	0.0853	3	0.6860	0.6902	0.0851	3	<b>0.6871</b>	<b>0.6911</b>	0.0857
5	0.7101	0.7346	0.1025	5	0.7175	0.7451	0.1071	5	<b>0.7183</b>	<b>0.7460</b>	0.1071
7	0.7171	0.7528	0.1191	7	0.7228	<b>0.7589</b>	0.1214	7	<b>0.7233</b>	0.7586	0.1212
10	0.7166	0.7590	0.1329	10	0.7209	<b>0.7598</b>	0.1355	10	<b>0.7209</b>	0.7592	0.1357
15	0.7125	0.7545	0.1431	15	<b>0.7159</b>	0.7579	0.1465	15	0.7157	<b>0.7589</b>	0.1469
20	0.7092	0.7489	0.1473	20	0.7062	0.7452	0.1529	20	<b>0.7110</b>	<b>0.7525</b>	0.1524

Table 2. Statistics on boundary displacement error (BDE) over the 300 images of the Berkeley image data base for the compared methods. Lower values represent better segmentations.

GMM				SVGMM				St-SVGMM			
$K$	Mean	Median	St. dev.	$K$	Mean	Median	St. dev.	$K$	Mean	Median	St. dev.
3	4.7893	<b>4.1636</b>	2.3856	3	4.7870	4.2057	2.3967	3	<b>4.7633</b>	4.2109	2.3820
5	4.3862	<b>3.7571</b>	2.1734	5	4.3938	3.8141	2.1745	5	<b>4.3666</b>	3.7939	2.1713
7	4.2438	3.7077	2.0950	7	4.2118	3.6831	2.0554	7	<b>4.1902</b>	<b>3.6451</b>	2.0556
10	4.1370	3.6025	2.0088	10	4.0963	3.5038	1.9863	10	<b>4.0929</b>	<b>3.4969</b>	2.0033
15	4.0996	3.6353	1.9760	15	4.0341	3.5036	1.9401	15	<b>4.0119</b>	<b>3.4486</b>	1.9345
20	4.1283	3.6784	2.0110	20	4.1913	3.6549	1.9084	20	<b>4.0049</b>	<b>3.4735</b>	1.9286

## References

- [1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006. [1](#), [2](#), [3](#), [5](#)
- [2] A. Blake and A. Zisserman. *Visual Reconstruction*. The MIT Press, 1987. [3](#)
- [3] K. Blekas, A. Likas, N. Galatsanos, and I. Lagaris. A spatially constrained mixture model for image segmentation. *IEEE Transactions on Neural Networks*, 16(2):494–498, 2005. [4](#)
- [4] M. A. Carreira-Perpinan. Fast nonparametric clustering with Gaussian blurring mean-shift. In *Proceedings of the International Conference on Machine Learning (ICML'06)*, pages 153–160, 2006. [1](#)
- [5] C. Carson, S. Belongie, H. Greenspan, and J. Malik. Blobworld: image segmentation using expectation-maximization and its application to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1026–1038, 2002. [5](#)
- [6] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002. [1](#)
- [7] P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977. [1](#), [4](#)
- [8] P. F. Felzenswalb and D. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004. [1](#)
- [9] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. *International Journal of Computer Vision*, 40(1):25–47, 2000. [1](#)
- [10] J. Freixenet, X. Munoz, D. Raba, J. Marti, and X. Cuff. Yet another survey on image segmentation: Region and boundary information integration. In *Lecture Notes in Computer Science. Proceedings of the European Conference on Computer Vision (ECCV'02)*, pages 408–422, 2002. [5](#)
- [11] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6):721–741, 1984. [1](#), [2](#), [3](#)
- [12] J. Marroquin, E. Arce, and S. Botello. Hidden Markov measure field models for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1380–1387, 2003. [1](#)
- [13] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th International Conference on Computer Vision*, volume 2, pages 416–423, July 2001. [5](#)
- [14] C. Nikou, N. Galatsanos, and A. Likas. A class-adaptive spatially variant mixture model for image segmentation. *IEEE Transactions on Image Processing*, 16(4):1121–1130, 2007. [1](#), [2](#), [3](#), [5](#)
- [15] A. N. Rajagopalan and S. Chaudhuri. An MRF model-based approach to simultaneous recovery of depth and restoration from defocused images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(17):577–589, 1999. [1](#)
- [16] S. Sanjay-Gopal and T. Hebert. Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm. *IEEE Transactions on Image Processing*, 7(7):1014–1028, 1998. [1](#), [2](#)
- [17] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000. [1](#)
- [18] R. Unnikrishnan, C. Pantofaru, and M. Hebert. Toward objective evaluation of image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):929–944, 2007. [5](#)
- [19] R. Xu and D. Wunsch II. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678, 2005. [1](#)
- [20] A. Yang, J. Wright, S. Sastry, and Y. Ma. Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding*, Accepted for publication:In Press, 2007. [1](#)
- [21] R. Zabih and V. Kolmogorov. Spatially coherent clustering using graph cuts. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'04)*, volume 2, pages 437–444, 2004. [1](#)
- [22] C. Zhu and D. Mumford. Prior learning and Gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1236–1250, 1997. [1](#)