

Multivariate Image Segmentation Using Semantic Region Growing With Adaptive Edge Penalty

A. K. Qin, *Member, IEEE*, and David A. Clausi, *Senior Member, IEEE*

Abstract—Multivariate image segmentation is a challenging task, influenced by large intraclass variation that reduces class distinguishability as well as increased feature space sparseness and solution space complexity that impose computational cost and degrade algorithmic robustness. To deal with these problems, a Markov random field (MRF) based multivariate segmentation algorithm called “multivariate iterative region growing using semantics” (MIRGS) is presented. In MIRGS, the impact of intraclass variation and computational cost are reduced using the MRF spatial context model incorporated with adaptive edge penalty and applied to regions. Semantic region growing starting from watershed over-segmentation and performed alternatively with segmentation gradually reduces the solution space size, which improves segmentation effectiveness. As a multivariate iterative algorithm, MIRGS is highly sensitive to initial conditions. To suppress initialization sensitivity, it employs a region-level k -means (RKM) based initialization method, which consistently provides accurate initial conditions at low computational cost. Experiments show the superiority of RKM relative to two commonly used initialization methods. Segmentation tests on a variety of synthetic and natural multivariate images demonstrate that MIRGS consistently outperforms three other published algorithms.

Index Terms—Initialization sensitivity, Markov random field (MRF), multilevel logistic (MLL) model, multivariate segmentation, region adjacency graph (RAG), semantic region growing, vector-valued image, watershed.

I. INTRODUCTION

COMPUTER vision applications often require segmentation of digital imagery into semantically meaningful regions. The segmented regions can provide a basis for subsequent tasks such as object detection and recognition, scene understanding and content-based image retrieval. Therefore, ultimate performance depends upon segmentation accuracy.

Rapid advances in image technologies lead to various types of digital images. Multivariate (vector-valued) imagery (e.g., color images) depicts each site using a vector that characterizes the

same scene from distinct aspects, where the number of vector elements is called the feature space dimension. Univariate imagery (e.g., grayscale images) can be regarded as a special multivariate case in which each site is depicted by a scalar. This paper focuses on multivariate 2-D imagery which is defined on a discrete 2-D rectangular lattice.

Multivariate image segmentation has been widely applied in diverse fields [1]–[4]. Although theoretically feasible to extend many univariate segmentation techniques to their multivariate counterparts, practical performance is influenced by the multivariate nature of the image. Intraclass variation is typically present since semantically meaningful regions (classes) are often inhomogeneous due to scene characteristics, imaging environment, and image noise. Large intraclass variation usually reduces class distinguishability and, thus, degrades segmentation performance. Multivariate imagery is especially sensitive to large intraclass variation since every component image is a variation contributor. Moreover, computational cost of segmentation algorithms increases while algorithmic robustness tends to decrease with increasing feature space sparseness and solution space complexity.

Markov random field (MRF) based image segmentation [5] is advocated for its intrinsic capability of reducing the impact of intraclass variation using spatial context information. Yu and Clausi [4], [6] built an iterative region growing using semantics (IRGS) algorithm for univariate image segmentation, which incorporates edge information with the MRF model. IRGS uses watershed over-segmentation to build a region adjacency graph (RAG), which then undergoes vertex labeling and merging by alternating segmentation and semantic region growing procedures. Although IRGS demonstrates superior segmentation ability, it cannot take multivariate images as inputs. Moreover, IRGS is sensitive to initial conditions and may not consistently provide accurate segmentation results.

We present a region-level MRF-based segmentation algorithm named multivariate IRGS (MIRGS), which advances the merits of the univariate IRGS to deal with multivariate imagery. Similar to IRGS, MIRGS is also sensitive to initial conditions, which becomes more pronounced as the feature space dimension increases. Therefore, a region-level k -means (RKM) method is used to provide robust initialization for MIRGS at low computational cost.

The next section describes the mathematical background for multivariate image segmentation. Section III presents the region-level MLL model based segmentation, which leads to the full description of MIRGS that is tested in Section IV.

Manuscript received June 12, 2009; revised February 21, 2010. First published March 15, 2010; current version published July 16, 2010. This work was supported in part by a NSERC Discovery Grant, in part by the Canadian Federal International Polar Year (IPY) project, and in part by Geomatics for Informed Decisions (GEOIDE). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sharathchandra Pankanti.

A. K. Qin was with the Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada. He is now with the School of Automation, Southeast University, Nanjing, 210096, China (e-mail: qfred008@gmail.com).

D. A. Clausi is with the Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada (e-mail: dclausi@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2010.2045708

II. BACKGROUND

A. Image Segmentation Problem Statement

Let $\mathcal{S} = \{(i, j) | 1 \leq i \leq M, 1 \leq j \leq N\}$ denote a discrete 2-D rectangular lattice of size $M \times N$. $\mathbf{Y} = \{\mathbf{Y}_s | s \in \mathcal{S}\}$ represents the multivariate 2-D image defined on \mathcal{S} , where $\mathbf{Y}_s = \{Y_s^{(k)} | 1 \leq k \leq n\}$ is the n -dimensional feature vector defined on site s . Here, each image site denotes a pixel. $\mathbf{Y}^{(k)}$ represents the k^{th} univariate component image of \mathbf{Y} . Suppose an image is to be segmented into C classes. Let $\mathcal{L} = \{l | 1 \leq l \leq C\}$ denote the set of class labels. $\mathbf{X} = \{X_s | s \in \mathcal{S}\}$ is the random label field defined on \mathcal{S} , where X_s is the random variable taking a value in \mathcal{L} .

Given an observed image $\mathbf{y} = \{\mathbf{y}_s | s \in \mathcal{S}\}$ with $\mathbf{y}_s = \{y_s^{(k)} | 1 \leq k \leq n\}$ as an instance of \mathbf{Y} , image segmentation aims at finding a label field configuration $\mathbf{x} = \{x_s | s \in \mathcal{S}\}$ to optimize certain criteria defined on \mathbf{y} and \mathbf{x} . The obtained configuration partitions \mathcal{S} into C disjoint subsets $\Omega_1 \dots \Omega_C$ where Ω_i corresponds to a region or cluster of regions with defined similar characteristics.

B. Multivariate Image Segmentation Techniques

Multivariate imagery may be produced by capturing several images of a particular scene under varying circumstances (such as frequency) or by extracting multiple features from each site in a univariate image.

Among existing multivariate segmentation techniques, feature space thresholding and clustering methods [1], [7], [8] are simple and time-efficient. However, they produce noisy segmentations due to ignoring spatial context. Edge based methods [9]–[11] segment using region boundaries. Although less sensitive to regional inhomogeneity, edge-based performance is influenced by ill-defined noisy edges and often degrades when multiple region boundaries are to be found in a globally optimal sense.

Region based segmentation methods [2], [12], [13] merge or split regions using region statistics as descriptors and, as such, are less sensitive to noise. However, choosing suitable merging and splitting criteria and thresholds is difficult especially when region statistics are nonstationary.

Model based segmentation methods [9], [10], [14], [15] have solid mathematical foundations. They perform well given accurate model formulations and efficient optimization techniques but are challenged by nonstationary image properties. Hybridization [4], [6], [16], [17] provides a method to combine the strength of region, edge and model based approaches.

C. MRF-Based Image Segmentation

Given that \mathbf{X} is a MRF with respect to a certain neighborhood system on \mathcal{S} , the MRF-based image segmentation framework [5], [14] can be formulated in a general form as

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathbf{X}} p(\mathbf{y} | \mathbf{x}) P(\mathbf{x}) \quad (1)$$

where $p(\mathbf{y} | \mathbf{x})$ is the conditional probability density function of the observed image \mathbf{y} given a label field configuration \mathbf{x} . $P(\mathbf{x})$ follows the Gibbs distribution [5]

$$P(\mathbf{x}) = \frac{\exp(-E(\mathbf{x})/T)}{\sum_{\mathbf{x} \in \mathbf{X}} \exp(-E(\mathbf{x})/T)} \quad (2)$$

where T denotes temperature typically assumed to be 1 and $E(\mathbf{x})$ represents the Gibbs energy function of the configuration \mathbf{x} , which equals the sum of clique potentials over all possible cliques in \mathcal{S} where each clique is a subset of \mathcal{S} [5].

The MRF-based segmentation can be formulated at the region level [4], [6], [18]–[22] using the region adjacency graph (RAG) [7] based image representation. A RAG is denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where \mathcal{V} and \mathcal{E} represent the set of vertices and the set of arcs connecting neighboring vertices. Each vertex $v \in \mathcal{V}$ depicts an image region with \mathcal{S}_v denoting the set of image sites constituting that region. Each arc $e \in \mathcal{E}$ represents the boundary between two neighboring regions.

In the region-level MRF-based segmentation, an input image is first over-segmented into disjoint and relatively homogenous regions upon which the RAG is constructed. Suppose that a certain neighborhood system and cliques are defined on \mathcal{G} [18]. Let $\mathbf{X}^r = \{X_v^r | X_v^r \in \mathcal{L}, v \in \mathcal{V}\}$ denote the Markov random label field on \mathcal{G} [18] with $\mathbf{x}^r = \{x_v^r | x_v^r \in \mathcal{L}, v \in \mathcal{V}\}$ being its instance. Here, x_v^r denotes the label for all sites $s \in \mathcal{S}_v$. The region-level MRF-based segmentation is formulated as

$$\mathbf{x}^{r*} = \arg \max_{\mathbf{x}^r \in \mathbf{X}^r} p(\mathbf{y} | \mathbf{x}^r) P(\mathbf{x}^r). \quad (3)$$

Here, $P(\mathbf{x}^r)$ follows the Gibbs distribution with its energy function $E(\mathbf{x}^r)$ defined according to a certain MRF model on \mathcal{G} . Under the class-conditionally independent assumption, the $p(\mathbf{y} | \mathbf{x}^r)$ can be formulated as

$$p(\mathbf{y} | \mathbf{x}^r) = \prod_{i=1}^C \prod_{\mathcal{S}_v \subseteq \Omega_i} \prod_{s \in \mathcal{S}_v} p(\mathbf{y}_s | x_v^r = i). \quad (4)$$

Compared to the pixel-level formulation, the region-level approach can dramatically reduce the solution space to facilitate optimization. Moreover, taking image regions instead of pixels as processing units implicitly reduces the influence of intraregion variation. However, performance is influenced by the initial over-segmentation. Too many initial regions may still lead to poor suboptimal solutions. Too few initial regions may sacrifice some important structure information, which cannot be recovered. Therefore, a desirable initial over-segmentation scheme should produce a minimal number of regions with boundaries preserved. Existing over-segmentation schemes include watershed transform [21]–[23], partition mode test [24], tone-region based segmentation [25], pixels extraction [20], normalized cuts [26] and superpixel lattices [27].

The MRF-based image segmentation corresponds to a combinatorial optimization problem. Various optimization techniques [5] have been used to search for its optimal solution.

D. Initialization of MRF-Based Segmentation

Most optimization techniques solving the MRF-based segmentation problem are iterative and, thus, initialization-sensitive. However, typically such papers do not describe their initialization methods or use random initialization that can lead to poor suboptimal solutions.

Initialization for the MRF-based segmentation is concerned with both feature model class statistics and the label field configuration, which are interrelated. The initialized feature model class statistics can determine the initial label field configuration and vice-versa. Among existing initialization methods, k -means and finite mixture models [3], [28], [29] are simple and relatively time-efficient although they themselves are sensitive to initial conditions. A multiagent system [30] provides an evolutionary optimization framework using the population-based initialization, which solves segmentation problems at high computational cost. Tree-structure vector quantization, agglomerative hierarchical clustering and region growing or splitting schemes [31]–[33] explore proper initial conditions in a hierarchical manner, which can be effective but time-consuming. Existing initialization methods are typically performed in the feature space without considering spatial context information. Therefore, the obtained initial conditions may correspond to unfavorable solutions for the MRF-based segmentation and prevent convergence to preferred solutions.

III. MIRGS ALGORITHM

A. Multilevel Logistic (MLL) Model Based Segmentation

The MLL model [5] is a popular MRF spatial context model, which imposes the local label consistency constraint to suppress noisy label field configurations. Its typical pixel-level formulation on \mathcal{S} only considers pair-site cliques with the corresponding Gibbs energy function defined as

$$E(\mathbf{x}) = \sum_{\langle s, t \rangle \in \mathcal{C}} U(x_s, x_t) \quad (5)$$

$$U(x_s, x_t) = \begin{cases} \beta, & \text{if } x_s \neq x_t \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where $\langle s, t \rangle$ and $U(x_s, x_t)$ denote a pair-site clique consisting of two neighboring sites $s, t \in \mathcal{S}$ and its potential, \mathcal{C} represents all pair-site cliques on \mathcal{S} , and β is a positive number. Since $E(\mathbf{x})$ (5) is intrinsically related to the total class boundary length, the pixel-level MLL model actually penalizes the existence of class boundaries.

The MLL model formulated on \mathcal{G} leads to the region-level MLL model based segmentation. Depending upon the definition of region boundaries as RAG arcs, the MLL model on \mathcal{G} may have different forms. If each pixel belongs to one of the regions obtained from the initial over-segmentation, i.e., $\bigcup_{v \in \mathcal{V}} \mathcal{S}_v = \mathcal{S}$, region boundaries will be implicitly defined in-between image sites and characterized by sets of neighboring pixel pairs astride the boundary. In such a case, the region-level MLL model becomes the same as the pixel-level one. However, region boundaries, thus, defined cannot be efficiently stored and manipulated as RAG arcs.

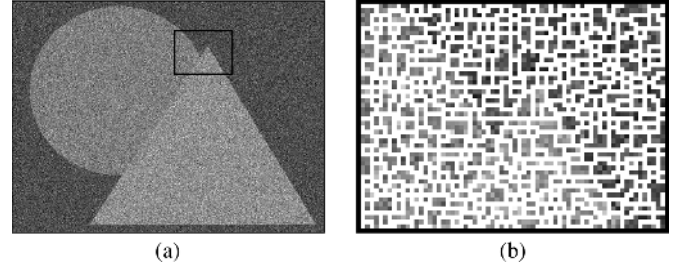


Fig. 1. Over-segmentation example using watershed transform. (a) Grayscale image $SG_{0.01}^A$ (378 \times 279) described in Section IV-A. (b) Over-segmentation within the region enclosed by the black box in (a) using watershed transform.

Here, we use an onsite region boundary definition [29] as follows:

- 1) $\bigcup_{v \in \mathcal{V}} \mathcal{S}_v \bigcup_{v \in \mathcal{V}} \partial \mathcal{S}_v = \mathcal{S}$;
- 2) $\forall v, w : \mathcal{S}_v \cap \partial \mathcal{S}_w = \emptyset$;
- 3) $\forall v \neq w : \mathcal{S}_v \cap \mathcal{S}_w = \emptyset$;
- 4) $\forall v : \partial \mathcal{S}_v \subseteq \bigcup_{w \in \mathcal{V}, w \neq v} \partial \mathcal{S}_w$

where $\partial \mathcal{S}_v$ denotes the one-pixel boundary outlining region \mathcal{S}_v . This definition indicates that each pixel belongs to either a region or a region boundary, where all regions are mutually separated by region boundaries. A region's boundary pixels are normally boundary pixels for at least one other neighboring region, e.g., $\partial \mathcal{S}_v \cap \partial \mathcal{S}_w$ represents the set of common boundary sites between neighboring regions \mathcal{S}_v and \mathcal{S}_w . Region boundaries, thus, defined can be easily represented using the popular chain code data structure [7]. Therefore, their storage and manipulation as RAG arcs are both space and time efficient.

According to this onsite boundary definition, the Gibbs energy function of the region-level MLL model is defined as

$$E(\mathbf{x}^r) = \sum_{\langle v, w \rangle \in \mathcal{C}^r} U(x_v^r, x_w^r) \quad (7)$$

$$U(x_v^r, x_w^r) = \begin{cases} \beta \sum_{s \in \partial \mathcal{S}_v \cap \partial \mathcal{S}_w} 0.5, & \text{if } x_v^r \neq x_w^r \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where $\langle v, w \rangle$ and $U(x_v^r, x_w^r)$ denote a pair-vertex clique involving two neighboring vertices v and w and its potential. \mathcal{C}^r represents all pair-vertex cliques on \mathcal{G} . Since $E(\mathbf{x}^r)$ (7) equals β times the total class boundary length, this region-level MLL model also penalizes the existence of class boundaries.

We use the well-established watershed transform [23] for the initial over-segmentation due to its high efficiency. An over-segmentation example is shown in Fig. 1. This transform partitions an input image into disjoint and relatively homogenous regions (catchment basins) as well as one-pixel region borders (watersheds) based upon the gradient magnitude image. The number of catchment basins equals the number of local minima detected in the gradient magnitude image.

Using (4), assume that $p(\mathbf{y}_s | x_v^r)$ defines a class-conditional multivariate Gaussian distribution, the region-level MLL model based segmentation can be formulated as

$$\mathbf{x}^{r*} = \arg \min_{\mathbf{x}^r \in \mathbf{X}^r} \left\{ \frac{1}{n} E_f(\mathbf{y}, \mathbf{x}^r) + \beta E_s(\mathbf{x}^r) \right\} \quad (9)$$

$$E_f(\mathbf{y}, \mathbf{x}^r) = \sum_{i=1}^C \sum_{\mathcal{S}_v \subseteq \Omega_i} \sum_{s \in \mathcal{S}_v} \left\{ \frac{1}{2} \log(|\Sigma_i|) + \frac{1}{2} (\mathbf{y}_s - \mathbf{u}_i)^T \times \Sigma_i^{-1} (\mathbf{y}_s - \mathbf{u}_i) \right\} \quad (10)$$

$$E_s(\mathbf{x}^r) = \sum_{i=1}^{C-1} \sum_{j=i+1}^C \sum_{s \in \partial\Omega_i \cap \partial\Omega_j} 1. \quad (11)$$

Here, $E_f(\mathbf{y}, \mathbf{x}^r)$ and $\beta E_s(\mathbf{x}^r)$ are energy functions related to the feature and spatial context models. \mathbf{u}_i and Σ_i denote mean vector and covariance matrix of class i . $\partial\Omega_i \cap \partial\Omega_j$ denotes the set of common boundary sites between classes Ω_i and Ω_j . Note that $\beta E_s(\mathbf{x}^r)$ equals $E(\mathbf{x}^r)$ (7), and β is regarded as a weighting factor. In multivariate segmentation, $E_f(\mathbf{y}, \mathbf{x}^r)$ increases proportionally with the feature space dimension n while $E_s(\mathbf{x}^r)$ does not. Therefore, β has to be dependent upon n to make a proper balance between $E_f(\mathbf{y}, \mathbf{x}^r)$ and $E_s(\mathbf{x}^r)$. To ensure the proper balance, we normalize $E_f(\mathbf{y}, \mathbf{x}^r)$ in (9) by n .

In the MRF-based segmentation, an inaccurate spatial context at early stages of optimization can mislead the estimation of feature model class statistics. The estimation error can be propagated in subsequent iterations, which generally results in an incorrect segmentation. To address this issue, Deng and Clausi [15] introduced a variable weighting $\alpha(\tau)$ into the feature model, which is hereby extended to the region level

$$\mathbf{x}^{r*} = \arg \min_{\mathbf{x}^r \in \mathbf{X}^r} \{ \alpha(\tau) E_f(\mathbf{y}, \mathbf{x}^r) + \beta E_s(\mathbf{x}^r) \}$$

$$\alpha(\tau) = a 0.9^{\tau-1} + \frac{1}{n} \quad (12)$$

where a is a constant parameter and τ is the current iteration. This adaptive weighting scheme gradually increases the influence of the spatial context model to prevent an inaccurate spatial context to mislead the estimation of feature model class statistics at early iterations. Note that the feature model normalization issue has already been taken into account in (12) via $\alpha(\tau)$.

After all iterations for solving (9) or (12) are finished, we need to deal with region boundary sites ($\bigcup_{v \in \mathcal{V}} \partial\mathcal{S}_v$). Among them, sites not belonging to class boundaries are directly assigned to the same class to which all its neighboring regions belong. Class boundary sites may be removed using maximum-likelihood classification [7] so that each site is assigned to one of the classes to which its neighbors belong, namely, $x_s = \arg \max_{i \in \{x_t \in \mathcal{L} | t \in \mathcal{N}_s\}} p(\mathbf{y}_s | i)$, $s \in \sum_{i=1}^{C-1} \sum_{j=i+1}^C \partial\Omega_i \cap \partial\Omega_j$ where \mathcal{N}_s denotes the neighborhood of site s and $p(\mathbf{y}_s | i)$ is obtained from the completed iterations of MIRGS.

Hereafter, the region-level MLL model based segmentation with and without the variable weighting factor $\alpha(\tau)$ are named as CMLL_T (constant MLL) and VMLL_T (variable MLL).

B. MIRGS Properties

The multivariate extension from IRGS to MIRGS involves both major and minor changes.

Major changes are related to the region boundary definition. IRGS defines region boundaries in-between image sites, which is not desirable for the following reasons:

- Watershed pixels obtained by watershed transform must be assigned to one of its neighboring regions. Common assignment criteria based upon region statistics are less effective due to feature space sparseness especially in multivariate cases. Consequently, watershed pixels located at edges may be wrongly assigned to undesirable regions and cannot be rectified subsequently. Such incorrect assignment may distort regional statistics, which successively influences segmentation performance. Moreover, this watershed removal step is time-consuming.
- Region boundaries are characterized by sets of neighboring pixel pairs astride the boundary and cannot be represented in a convenient data structure. Therefore, it is space and time consuming to store region boundaries as RAG arcs and to manipulate them.
- IRGS defines edge strength as intensity value differences between neighboring sites astride the boundary. However, a more reasonable measurement of edge strength is the gradient magnitudes calculated on image sites.

MIRGS uses the onsite region boundary definition (Section III-A) to address these undesirable aspects. Specifically, watersheds are directly used as region boundaries to avoid the problematic watershed removal process. Chain code [7] based boundary representation enables efficient storage and manipulation of region boundaries as RAG arcs. Edge strength is naturally defined as gradient magnitudes on boundary sites.

Minor changes include:

- IRGS calculates the gradient magnitude image using a traditional gradient operator [8], which is only applicable to univariate input images. MIRGS calculates gradient from multivariate input images using a vector gradient method [34], which extends watershed transform to multivariate imagery.
- MIRGS generalizes the univariate feature model of IRGS into its multivariate counterpart, which is further normalized by the feature space dimension.
- MIRGS generalizes the univariate Fisher criterion used by IRGS to adjust the spatial context model's parameter to its multivariate counterpart used by MIRGS.

MIRGS inherits many attractive properties from the univariate IRGS [6]:

- MIRGS uses a monotonically decreasing edge penalty function with the region-level MLL model.
- MIRGS changes edge penalty at each iteration to generate a sequence of spatial context models.
- MIRGS performs on a hierarchical RAG with the bottom-up nonincreasing vertex number, which gradually reduces the solution space size to improve segmentation effectiveness.

By incorporating an edge penalty function with $E_s(\mathbf{x}^r)$ (11) in the CMLL_T formulation (9)

$$E_s(\mathbf{x}^r) = \sum_{i=1}^{C-1} \sum_{j=i+1}^C \sum_{s \in \partial\Omega_i \cap \partial\Omega_j} g(\nabla_s) \quad (13)$$

$$g(\nabla_s) = \exp \left[- \left(\frac{\nabla_s}{K(\tau)} \right)^2 \right]. \quad (14)$$

MIRGS can be formulated as (15), shown at the bottom of the page, where $g(\nabla_s)$ is a monotonically decreasing edge penalty function with $\nabla_s \in [0, 1]$ representing the normalized gradient magnitude on site s as the edge strength measurement. Here, $g(\cdot)$ is defined as an exponential function with a varying parameter $K(\tau)$. $K(\tau)$ takes positive values and monotonically increases with the iteration number τ , which controls how quickly penalty decays with increasing edge strength. As $K(\tau)$ increases, the penalty difference between strong and weak edge strength decreases. When $K(\tau)$ approaches infinity, edge penalty equals one. Therefore, the formulation (15) can be regarded as making a smooth transition from the feature model based segmentation to the MLL model based segmentation as iteration proceeds, which implicitly solves the issue addressed by $\alpha(\tau)$ [15].

MIRGS applies watershed transform [23] on the gradient magnitude image calculated from the given multivariate image (Section III-C) to produce an over-segmentation upon which the initial RAG is built. Before the first iteration, feature model class statistics (\mathbf{u}_i and Σ_i) and the label field configuration \mathbf{x}^r (class labels of vertices in the RAG) are initialized. At the τ ($\tau \geq 1$) iteration, given the current feature model class statistics, Gibbs sampling [14] is applied on the current RAG to find a suboptimal solution for (15). Specifically, each RAG vertex is visited once in a random order. Suppose $v \in \mathcal{V}$ is being visited and the current label field configuration is $\mathbf{x}^r(\text{cur})$, probabilities $P_i = P(x_v^r = i, x_w^r = x_w^r(\text{cur}), \forall w \in \mathcal{V}, w \neq v | \mathbf{y})$, $i = 1 \dots C$ are first calculated using the Gibbs distribution with the energy term defined as per (15) and the temperature parameter being a fixed constant. Then, class label x_v^r is set to i with probability P_i . After Gibbs sampling, a greedy semantic region growing process is carried out to update the RAG. Specifically, consider each adjacent vertex pair that has the same class label in the current RAG, e.g., vertex pair v and w . The merging criterion δE is defined by analogy with Section 4.1 in [6] as¹

$$\begin{aligned} \delta E = & \frac{1}{2n} \sum_{s \in \mathcal{S}_{vw}} \log(|\Sigma_{vw}|) - \frac{1}{2n} \sum_{s \in \mathcal{S}_v} \log(|\Sigma_v|) \\ & - \frac{1}{2n} \sum_{s \in \mathcal{S}_w} \log(|\Sigma_w|) - \beta \sum_{s \in \partial \mathcal{S}_v \cap \partial \mathcal{S}_w} \exp \left[- \left(\frac{\nabla_s}{K(\tau)} \right)^2 \right] \end{aligned} \quad (16)$$

¹If the cardinality of \mathcal{S}_v , \mathcal{S}_w , or \mathcal{S}_{vw} is too small, the corresponding covariance matrices might become singular and, thus, have their determinant values equal to zero, which nullifies logarithm operations. As such, we assign a small positive value (10^{-12}) to the determinant of any singular covariance matrix so as to always make feasible the calculation of (16).

where $\mathcal{S}_{vw} = \mathcal{S}_v \cup \mathcal{S}_w$ and Σ_{vw} denotes the covariance matrix with respect to \mathcal{S}_{vw} . Among all of the adjacent vertex pairs under consideration, the neighboring regions corresponding to the vertex pair having the smallest δE are merged if this δE is negative, and subsequently update the RAG and the corresponding region statistics. This merging process is repeated until the smallest δE is non-negative. Feature model class statistics (\mathbf{u}_i and Σ_i) are then recomputed based upon the updated RAG as

$$w_{vi} = \begin{cases} 1, & \text{if } x_v^r = i \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

$$\mathbf{u}_i = \frac{\sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}_v} w_{vi} \mathbf{y}_s}{\sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}_v} w_{vi}} \quad (18)$$

$$\Sigma_i = \frac{\sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}_v} w_{vi} (\mathbf{y}_s - \mathbf{u}_i)(\mathbf{y}_s - \mathbf{u}_i)^T}{\sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}_v} w_{vi}}. \quad (19)$$

Here, singular Σ_i will be regularized [35]. Iterations continue until the prespecified maximum iteration number τ_{max} is reached.

C. Gradient Magnitude Computation on Multivariate Imagery

Both watershed transform [23] and edge strength measuring require a single gradient magnitude image derived from multivariate imagery. This is calculated using the vector gradient method [34]. Instead of combining gradient magnitudes derived independently from each univariate component image according to certain rules, the vector gradient method directly calculates the gradient in a vector field.

Given the observed multivariate image \mathbf{y} , the vector gradient approach computes the gradient magnitude and direction on site s as the square root of the largest eigenvalue of the matrix $\mathbf{D}_s^T \mathbf{D}_s$ and its corresponding eigenvector, where \mathbf{D}_s is an $n \times 2$ matrix defined as

$$\mathbf{D}_s = \begin{bmatrix} \partial_v \mathbf{y}^{(1)}(s) & \partial_h \mathbf{y}^{(1)}(s) \\ \vdots & \vdots \\ \partial_v \mathbf{y}^{(k)}(s) & \partial_h \mathbf{y}^{(k)}(s) \\ \vdots & \vdots \\ \partial_v \mathbf{y}^{(n)}(s) & \partial_h \mathbf{y}^{(n)}(s) \end{bmatrix} \quad (20)$$

where $\partial_v \mathbf{y}^{(k)}(s)$ and $\partial_h \mathbf{y}^{(k)}(s)$ denote the first partial derivatives of the k^{th} univariate component image of \mathbf{y} on site s with respect to vertical and horizontal directions, respectively. In MIRGS, partial derivatives are computed using the Gaussian derivative method [7]. The largest eigenvalue λ_s of the matrix

$$\begin{aligned} \mathbf{x}^{r*} = \arg \min_{\mathbf{x}^r \in \mathbf{X}^r} & \left\{ \frac{1}{n} \sum_{i=1}^C \sum_{\mathcal{S}_v \subseteq \Omega_i} \sum_{s \in \mathcal{S}_v} \left\{ \frac{1}{2} \log(|\Sigma_i|) + \frac{1}{2} (\mathbf{y}_s - \mathbf{u}_i)^T \Sigma_i^{-1} (\mathbf{y}_s - \mathbf{u}_i) \right\} \right. \\ & \left. + \beta \sum_{i=1}^{C-1} \sum_{j=i+1}^C \sum_{s \in \partial \Omega_i \cap \partial \Omega_j} \exp \left[- \left(\frac{\nabla_s}{K(\tau)} \right)^2 \right] \right\} \end{aligned} \quad (15)$$

$\mathbf{D}_s^T \mathbf{D}_s$ can be calculated analytically using Eqns. (4)–(8) in [34]. Accordingly, the gradient magnitude on site s is $\sqrt{\lambda_s}$ with its normalized value ∇_s calculated as

$$\nabla_s = \sqrt{\lambda_s} / \max_{s \in S} \sqrt{\lambda_s}. \quad (21)$$

D. Parameter Description

The parameter β in (15) can be either set empirically as a fixed value *a priori* [36] or estimated iteratively during segmentation [37], [38]. The former is ad hoc with respect to the images under examination while the latter is influenced by image noise and complexity. Following [6], at each iteration, MIRGS first uses a lookup table generated by the Monte Carlo method [38] to obtain an intermediate parameter β_0 according to the ratio of the current total class boundary length over the image size, and then determines β using β_0 and the two-class Fisher criterion [35]

$$\beta = c_1 \frac{J/c_2}{1 + J/c_2} \beta_0 \quad (22)$$

$$J = \min_{i,j} \text{trace} \left(S_{W_{ij}}^{-1} S_{B_{ij}} \right) \quad (23)$$

$$S_{W_{ij}} = \frac{N_i}{N_i + N_j} \Sigma_i + \frac{N_j}{N_i + N_j} \Sigma_j \quad (24)$$

$$S_{B_{ij}} = (\mathbf{u}_i - \mathbf{u}_j)(\mathbf{u}_i - \mathbf{u}_j)^T \quad (25)$$

where J is the minimum two-class Fisher criterion value among all class pairs, c_1 and c_2 are two parameters controlling the adjusting rule, $S_{W_{ij}}$ and $S_{B_{ij}}$ denote the pooled within-class scatter matrix and between-class scatter matrix regarding classes i and j , and N_i represents the cardinality of Ω_i . J measures the overall class separability in the feature space, so that the larger the J , the better the separability. Therefore, when the feature model dominates the energy function (15) at early iterations, J is relatively large and accordingly β is large to emphasize the spatial context model. When the spatial context model gradually plays a dominant role with increasing iterations, J decreases to reduce β so as to avoid the under-segmentation risk [6].

The parameter $K(\tau)$ in the edge penalty function is defined in a manner similar to [6]

$$K(\tau) = \begin{cases} f(2\tau), & \text{if } 1 \leq \tau \leq 50 \\ 1.1K(\tau - 1), & \text{if } \tau > 50 \end{cases} \quad (26)$$

where $f(i)$ takes a value in $[0, 1]$ to which i percent of the class boundary sites have the smaller ∇_s in comparison.

E. RKM-Based Initialization

Multivariate iterative image segmentation, like MIRGS, requires robust initialization to consistently achieve satisfying solutions. Initialization sensitivity in multivariate cases is more severe than that in univariate cases due to increased feature space sparseness and solution space complexity. Thus, a region-level k -means (RKM) based initialization method is incorporated into MIRGS to provide robust initialization. RKM takes

TABLE I
ALGORITHMIC DESCRIPTION OF MIRGS

1. Initially, set c_1 and c_2 in (22), τ_{int_1} , τ_{int_2} and τ_{int_3} in Section III-E, and specify the number of classes C and the maximum number of iterations τ_{max} .
2. Perform watershed transform on the gradient magnitude image calculated from the given image to obtain an over-segmentation upon which the initial RAG is constructed.
3. Perform the RKM-based initialization to obtain initial feature model class statistics and label field configuration. Let $\tau = 1$.
4. Update β and $K(\tau)$ using (22)–(26).
5. Perform Gibbs sampling on the current RAG to find a suboptimal solution for (15). Each RAG vertex is visited once in a random order. If $\tau = \tau_{max}$, go to step 9.
6. Calculate δE using (16) for each pair of adjacent vertices in the current RAG that have the same class label.
7. If the minimum δE is negative, merge the two neighboring regions depicted by the corresponding vertex pair, and then update the RAG. Go back to step 6 if merging happens.
8. Recompute feature model class statistics from the current RAG using (17)–(19). Let $\tau = \tau + 1$, and go back to step 4.
9. Assign any region boundary sites not belonging to class boundaries to the same class to which its neighboring regions belong.

image regions S_v produced by watershed over-segmentation as processing units instead of pixels.

RKM seeks a set of optimal class mean vectors \mathbf{u}_i , $i = 1 \dots C$ to minimize the following energy function:

$$E_{RKM} = \sum_{i=1}^C \sum_{S_v \subseteq \Omega_i} \sum_{s \in S_v} (\mathbf{y}_s - \mathbf{u}_i)^T (\mathbf{y}_s - \mathbf{u}_i). \quad (27)$$

To reduce initialization sensitivity, RKM employs a multistart scheme. τ_{int_1} sets of class mean vectors are randomly initialized. Each set performs τ_{int_2} iterations. At each iteration, the label field configuration \mathbf{x}^r is updated using the current class mean vectors according to the nearest center rule as

$$\forall v \in \mathcal{V}, \quad x_v^r = \arg \min_{i \in \mathcal{L}} \sum_{s \in S_v} (\mathbf{y}_s - \mathbf{u}_i)^T (\mathbf{y}_s - \mathbf{u}_i) \quad (28)$$

\mathbf{u}_i , $i = 1 \dots C$ is then recomputed using (17) and (18). From the derived τ_{int_1} sets of class mean vectors, the set with the smallest energy (27) is chosen as the starting point and refined further by τ_{int_3} iterations. Both multistart (τ_{int_2}) and further refining (τ_{int_3}) iterations can save time by terminating once the ratio of the absolute energy difference between consecutive iterations over the energy at the former iteration is below 10^{-6} .

Class mean vectors obtained by RKM can determine a label field configuration using (28) from which class covariance matrices can be calculated using (17) and (19). Regularization [35] will be applied in the case of singular covariance matrices. Such class mean vectors, class covariance matrices, and label field configuration provide the initial condition for further segmentation.

The RKM-based initialization method has several attractive features:

- low computational cost due to its simple formulation and region-level optimization;
- minimal initialization sensitivity;
- implicitly takes into account spatial context information via (28), i.e., class labels on any sites in one region depend

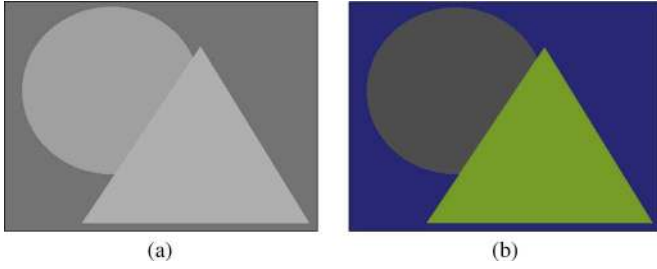


Fig. 2. Original synthetic images for generating noisy test images. (a) Grayscale image (378×279) of three gray levels: 96, 144, 160. (b) Color image (378×279) of three RGB colors: (39, 39, 117), (78, 78, 78), (117, 156, 39).

upon the feature properties on all sites in the same region and, thus, can provide accurate initial conditions for the MRF-based segmentation.

F. Overall Algorithm

The overall algorithm of MIRGS is described in Table I. In unusual cases, the final segmentation result obtained by MIRGS contains the number of classes fewer than C . This is triggered by step 5 and considered as segmentation failure.

IV. EXPERIMENTS

Segmentation tests are performed on various synthetic and natural multivariate images to evaluate two aspects:

- 1) *Initialization Analysis*: Determine which of the three initialization methods (random—RAND, expectation-maximization Gaussian mixture model—GMM, and RKM) provides accurate initial conditions for further segmentation.
- 2) *Segmentation Analysis*: Using the preferred initialization method from 1), compare the (i) accuracy and (ii) computation time of three MRF-based segmentation algorithms: CMLL_r, VMLL_r and MIRGS as well as GMM, where GMM acts as a base reference case.

A. Test Images

1) *Grayscale Images*: The synthetic grayscale image $SG_{0.01}^A$ shown in Fig. 3(a) is a univariate image. It is generated by adding Gaussian noise with mean (μ) 0 and variance (σ_n^2) 0.01 to a synthetic grayscale image shown in Fig. 2(a), which has three gray levels (96, 144, 160) corresponding to three classes (background, circle and triangle). Here, σ_n^2 denotes the normalized variance, i.e., $\sigma_n^2 = \sigma^2/255^2$ where σ^2 is the actual variance used. The histogram of $SG_{0.01}^A$ [Fig. 3(c)] shows that the image noise creates tremendous intraclass variation resulting in a uni-modal distribution.

2) *Color Images*: Segmentation tests on color images are performed in the CIE $L^*a^*b^*$ color space [39] due to CIE $L^*a^*b^*$'s desirable perceptually uniform property. Two types of color images with synthetic and natural intraclass variation are used.

The synthetic color image shown in Fig. 2(b) is a color version of Fig. 2(a), which consists of three RGB colors [(39, 39, 117), (78, 78, 78), (117, 156, 39)] corresponding to the three

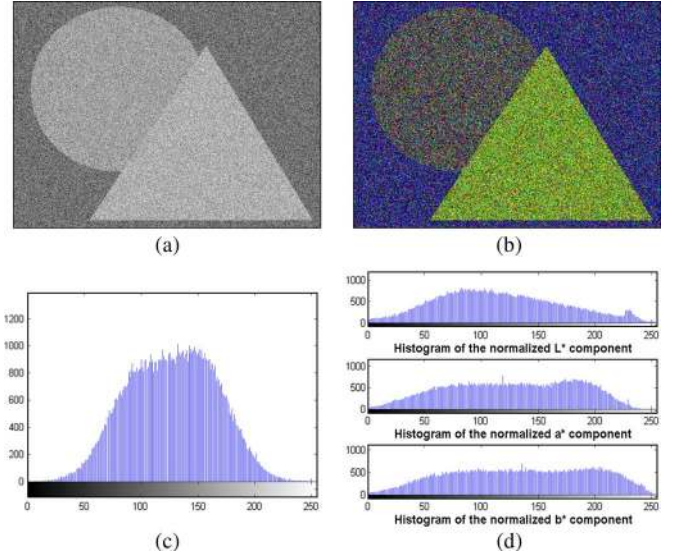


Fig. 3. Histograms of two noisy test image examples. (a) Grayscale image $SC_{0.01}^A$. (b) Color image $SC_{0.07}^A$. (c) Histogram of (a). (d) Histogram of (b) in the normalized CIE $L^*a^*b^*$ space.

classes. Additive Gaussian noise ($\mu = 0$ and $\sigma_n^2 = 0.01, 0.03, 0.05, 0.07, 0.09$) and multiplicative Gaussian noise ($\mu = 1$ and $\sigma_n^2 = 0.1, 0.3, 0.5, 0.7, 0.9$), are independently imposed on each RGB channel of Fig. 2(b), respectively to generate two groups of synthetic color images ($SC_{\sigma_n^2}^A$ and $SC_{\sigma_n^2}^M$). Fig. 3(d) illustrates the histogram of an example image $SC_{0.07}^A$ (Fig. 3(b)) in the normalized CIE $L^*a^*b^*$ color space where each of L^* , a^* and b^* component images is normalized to [0, 255]. As in Fig. 3(d), the individual class distributions are not distinct.

Natural color images are inherently characterized by considerable intraclass variation since classes in the scene are typically not homogeneous. Here, three natural color images (NC1, NC2, NC3) are used (Fig. 6). NC1 contains three classes (horses, green grasses and yellow flowers), NC2 contains three classes (grass field, car and raceway), and NC3 contains four classes (lake, mountain, blue sky, and clouds).

3) *Textured Images*: Three categories of nine Brodatz [40] mosaic images (Fig. 5) corresponding to different degrees of segmentation difficulty are used. The filter bank composed of 24 complex Gabor filters at four frequencies (22.63, 11.31, 5.66, and 2.83 pixels per cycle) and six orientations ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$, and 150°) is applied on each image [41]. The magnitudes of 24 filtered complex images constitute the multivariate features.

Category 1 includes three published mosaic images that have relatively small intraclass variation and, thus, easy to segment: T1 [15], [42] with four Brodatz textures² D68, D55, D84, and D77, T2 [43] with five Brodatz textures D77, D84, D55, D53, and D24, and T3 [41], [42] with five Brodatz textures D77, D55, D84, D17, and D24. As indicated in [42], the D17 (herring bone) should be regarded as containing two different textures when orientation-sensitive Gabor filters are used. Therefore, there are four, five, six classes in T1, T2, and T3, respectively. Textured images in this category are relatively easy to segment since their

²The D* denotes the numbering system used in the Brodatz album.

TABLE II

MAXIMUM VALUE (MAX), 3RD QUARTILE (Q3), 1ST QUARTILE (Q1), MINIMUM VALUE (MIN) AND MEAN VALUE (MEAN) OF acc (%) AND κ BY APPLYING MIRGS USING RANDOM, GMM AND RKM INITIALIZATION 10 TIMES ON GRAYSCALE IMAGE $SG_{0.01}^A$, COLOR IMAGE $SC_{0.07}^A$ AND TEXTURED IMAGE T2

		Grayscale Image			Color Image			Textured Image		
		RAND	GMM	RKM	RAND	GMM	RKM	RAND	GMM	RKM
acc	Max	99.30	99.30	99.30	99.10	72.20	99.20	95.50	92.10	97.30
	Q3	99.20	99.30	99.30	99.10	72.10	98.90	93.20	92.10	97.30
	Q1	72.20	99.30	99.30	72.20	72.10	98.90	69.80	92.10	97.30
	Min	72.20	99.30	99.30	72.20	72.10	98.90	56.70	92.10	97.30
	Mean	91.05	99.30	99.30	90.98	72.11	98.93	80.12	92.10	97.30
κ	Max	0.990	0.989	0.989	0.987	0.569	0.988	0.944	0.910	0.966
	Q3	0.987	0.989	0.989	0.986	0.568	0.984	0.915	0.910	0.966
	Q1	0.571	0.989	0.989	0.539	0.568	0.983	0.622	0.910	0.966
	Min	0.571	0.988	0.989	0.539	0.568	0.982	0.459	0.910	0.966
	Mean	0.862	0.989	0.989	0.851	0.568	0.984	0.752	0.910	0.966

TABLE III

MAXIMUM VALUE (MAX), 3RD QUARTILE (Q3), 1ST QUARTILE (Q1), MINIMUM VALUE (MIN) AND MEAN VALUE (MEAN) OF COMPUTATION TIME (MINUTES) OF THE INITIALIZATION PART AND THE WHOLE SEGMENTATION BY APPLYING MIRGS USING RANDOM, GMM AND RKM INITIALIZATION 10 TIMES ON GRAYSCALE IMAGE $SG_{0.01}^A$, COLOR IMAGE $SC_{0.07}^A$ AND TEXTURED IMAGE T2

		Grayscale Image			Color Image			Textured Image		
		RAND	GMM	RKM	RAND	GMM	RKM	RAND	GMM	RKM
Init Time	Max	0.00	2.12	0.07	0.00	7.37	0.12	0.02	30.42	0.42
	Q3	0.00	2.10	0.07	0.00	7.25	0.12	0.02	30.15	0.42
	Q1	0.00	2.07	0.05	0.00	7.17	0.10	0.02	28.58	0.38
	Min	0.00	2.05	0.05	0.00	7.07	0.10	0.00	27.02	0.35
	Mean	0.00	2.08	0.06	0.00	7.21	0.11	0.01	29.10	0.40
Seg Time	Max	0.88	2.98	0.92	1.15	8.48	1.27	13.48	47.82	18.57
	Q3	0.88	2.93	0.92	1.13	8.38	1.25	11.52	47.52	18.47
	Q1	0.78	2.92	0.90	1.10	8.30	1.23	8.10	45.98	17.48
	Min	0.77	2.92	0.90	1.03	8.20	1.22	7.57	44.35	17.42
	Mean	0.84	2.94	0.91	1.11	8.34	1.24	9.77	46.47	18.08

intraclass variation in the feature space is not too large to considerably reduce class separability.

Category 2 involves three mosaic images created using composite textures to artificially increase intraclass variation, which leads to higher segmentation difficulty. The composite texture is generated by linearly combining intensities of several Brodatz textures on each image site. T4 is composed of D5 (top-left), $0.5 * D8 + 0.5 * D84$ (top-right), D84 (bottom-left) and D92 (bottom-right). T5 is composed of $0.8 * D24 + 0.2 * D1$ (top-left), $0.8 * D29 + 0.2 * D1$ (top-right), $0.8 * D57 + 0.2 * D1$ (bottom-left) and $0.8 * D92 + 0.2 * D1$ (bottom-right). Due to the use of a common component texture D1, the class separability among all 4 composite textures in the feature space is reduced. T6 is created from $0.7 * D15 + 0.3 * D104$ (top-left), $0.7 * D54 + 0.3 * D104$ (top-right), $0.7 * D55 + 0.3 * D104$ (bottom-left) and $0.7 * D110 + 0.3 * D104$ (bottom-right) where the weight of the common component D104 in composite textures is increased to further reduce the class separability in the feature space.

Category 3 contains three mosaic images generated using Brodatz textures with similar visual perception. T7 consists of D84 (top-left), D54 (top-right), D112 (bottom-left) and D22 (bottom-right). T8 consists of D81 (top-left), D85 (top-right), D82 (bottom-left) and D80 (bottom-right). T9 consists of D6 (top-left), D19 (top-right), D55 (bottom-left) and D84 (bottom-right).

B. Experimental Setup—Initialization Analysis

Three initialization methods—random (RAND), expectation-maximization Gaussian mixture model (GMM) clustering [35],

and RKM—will be compared in terms of their capability to initialize MIRGS for segmenting gray, color and textured images ($SG_{0.01}^A$, $SC_{0.07}^A$ and T2) that have different feature space dimensions (1, 3, and 24).

RAND randomly labels each vertex in the initial RAG (i.e., randomly initializes \mathbf{x}^r) and then calculates initial feature model class statistics using (17)–(19). GMM is a pixel-level initialization method, which here uses the multistart pixel-level k -means for its own initialization. The final class mean vectors and covariance matrices estimated by GMM are used as initial feature model class statistics from which the initial \mathbf{x}^r is determined using the feature model. In both RAND and GMM, singular covariance matrices are regularized.

For GMM and RKM, setting τ_{int_1} to 25 and τ_{int_2} to 20 provides a good tradeoff between accurate initialization and reasonable computational speed. τ_{int_3} is set to 500, which is normally not exhausted since the early stopping criterion is met.

MIRGS's parameters are obtained by trial and error and remain the same for all tests: $c_1 = 5$, $c_2 = 0.2$, $\tau_{max} = 100$. In fact, c_1 is more dependent upon applications than other parameters [44]. We fixed the value of c_1 throughout our experiment since we found that $c_1 = 5$ can produce satisfactory segmentation results for the current test suite. In MIRGS, Gibbs sampling is performed using a small constant temperature 0.01. The number of classes C is specified *a priori*.

With ground truth, segmentation accuracy (acc) (namely the percentage of correctly labeled pixels) and kappa coefficient (κ) [45] (that measures the degree of agreement between the segmentation result and ground truth) can be determined. They are used to quantitatively evaluate segmentation performance.

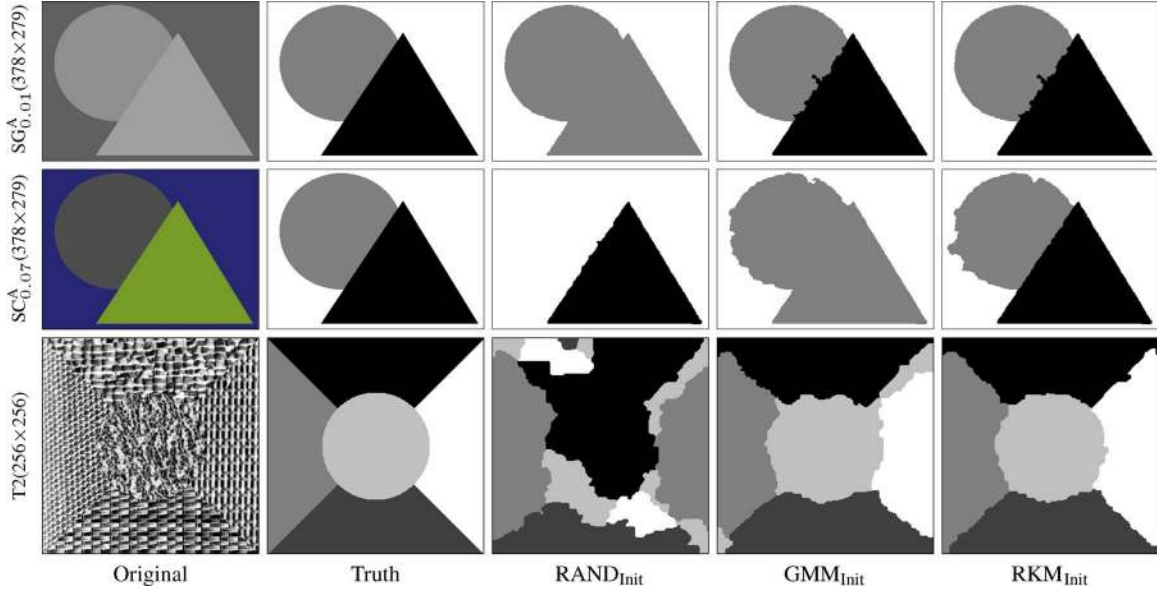


Fig. 4. Poorest segmentation results corresponding to the minimum acc and $kappa$ among 10 runs by applying MIRGS with RAND, GMM and RKM initialization on grayscale image $SG_{0.01}^A$, color image $SC_{0.07}^A$ and textured image T2, respectively. Results with regard to the RKM initialization are most visually agreeable to segmentation ground truth.

TABLE IV

MEAN VALUES OF acc (%) AND κ OVER 10 RUNS BY APPLYING GMM, CMLL_r, VMLL_r AND MIRGS ON TWO GROUPS OF 10 SYNTHETIC COLOR IMAGES WITH ADDITIVE AND MULTIPLICATIVE GAUSSIAN NOISE AT VARYING NOISE LEVELS (AN EXAMPLE IS SHOWN IN FIG. 4). BOLD SHOWS OPTIMAL MEAN VALUES OF acc AND κ AS WELL AS THOSE INDISCERNIBLE FROM THE OPTIMAL BASED UPON WILCOXON SIGNED-RANK TEST WITH SIGNIFICANCE LEVEL 0.05

Images			GMM	CMLL _r	VMLL _r	MIRGS
	σ_n^2					
SC ^A _{σ_n^2}	0.01	acc	88.90	99.60	99.60	99.70
		κ	0.831	0.994	0.995	0.995
	0.03	acc	68.59	99.06	99.40	99.50
		κ	0.533	0.985	0.990	0.993
	0.05	acc	55.10	98.18	99.15	99.32
		κ	0.317	0.972	0.987	0.990
	0.07	acc	51.30	96.96	98.78	98.93
		κ	0.267	0.954	0.981	0.984
	0.09	acc	37.20	94.24	98.35	98.74
		κ	-0.009	0.912	0.975	0.981
SC ^M _{σ_n^2}	0.1	acc	91.80	99.60	99.60	99.70
		κ	0.874	0.994	0.994	0.995
	0.3	acc	76.80	99.35	99.45	99.60
		κ	0.649	0.990	0.992	0.993
	0.5	acc	68.20	99.25	99.45	99.40
		κ	0.516	0.989	0.992	0.991
	0.7	acc	59.40	98.85	99.27	99.40
		κ	0.376	0.983	0.989	0.991
	0.9	acc	59.80	98.89	99.28	99.30
		κ	0.388	0.983	0.989	0.990

All algorithms are implemented in C++, encapsulated in the MAGIC system [46], and performed on a Windows XP PC with an Intel P4 3.0 GHz CPU using 1 GB memory. Computation time is recorded for each case.

For each test image, MIRGS uses each initialization method 10 times starting from different random seeds while all three initialization methods use the same random seed with respect to any individual run.

TABLE V

MEAN VALUES OF acc (%) AND κ OVER 10 RUNS BY APPLYING GMM, CMLL_r, VMLL_r AND MIRGS ON THREE CATEGORIES OF NINE TEXTURED IMAGES AS SHOWN IN FIG. 5. BOLD SHOWS OPTIMAL MEAN VALUES OF acc AND κ AS WELL AS THOSE INDISCERNIBLE FROM THE OPTIMAL BASED UPON WILCOXON SIGNED-RANK TEST WITH SIGNIFICANCE LEVEL 0.05. “Cat” STANDS FOR “CATEGORY”

Images			GMM	CMLL _r	VMLL _r	MIRGS
Cat 1	T1	acc	90.20	98.43	97.96	98.40
		κ	0.870	0.979	0.973	0.979
	T2	acc	87.40	96.54	95.73	97.30
		κ	0.843	0.957	0.947	0.966
	T3	acc	82.42	95.13	95.04	96.32
		κ	0.787	0.941	0.939	0.955
Cat 2	T4	acc	75.80	92.14	93.28	98.59
		κ	0.677	0.895	0.910	0.981
	T5	acc	79.70	94.43	94.08	98.93
		κ	0.730	0.926	0.921	0.986
	T6	acc	78.80	91.56	93.93	97.23
		κ	0.718	0.888	0.919	0.963
Cat 3	T7	acc	61.80	92.39	91.28	98.91
		κ	0.490	0.899	0.884	0.985
	T8	acc	67.30	83.90	83.64	98.11
		κ	0.564	0.785	0.782	0.975
	T9	acc	89.10	73.63	92.47	95.14
		κ	0.854	0.648	0.900	0.935

C. Tests—Initialization Analysis

For each test image, acc , κ , and computation time of the initialization part and the whole MIRGS segmentation over 10 runs are summarized in Tables II and III. Minimum and maximum values (Min and Max), first and third quartiles (Q1 and Q3), and the mean value (Mean) are shown. Fig. 4 illustrates the poorest segmentation results corresponding to the Min of acc and $kappa$ in Table II.

RAND is the fastest initialization method (Table III) but shows inconsistent segmentation performance (Table II). For all three test images, interquartile ranges between Q1 and

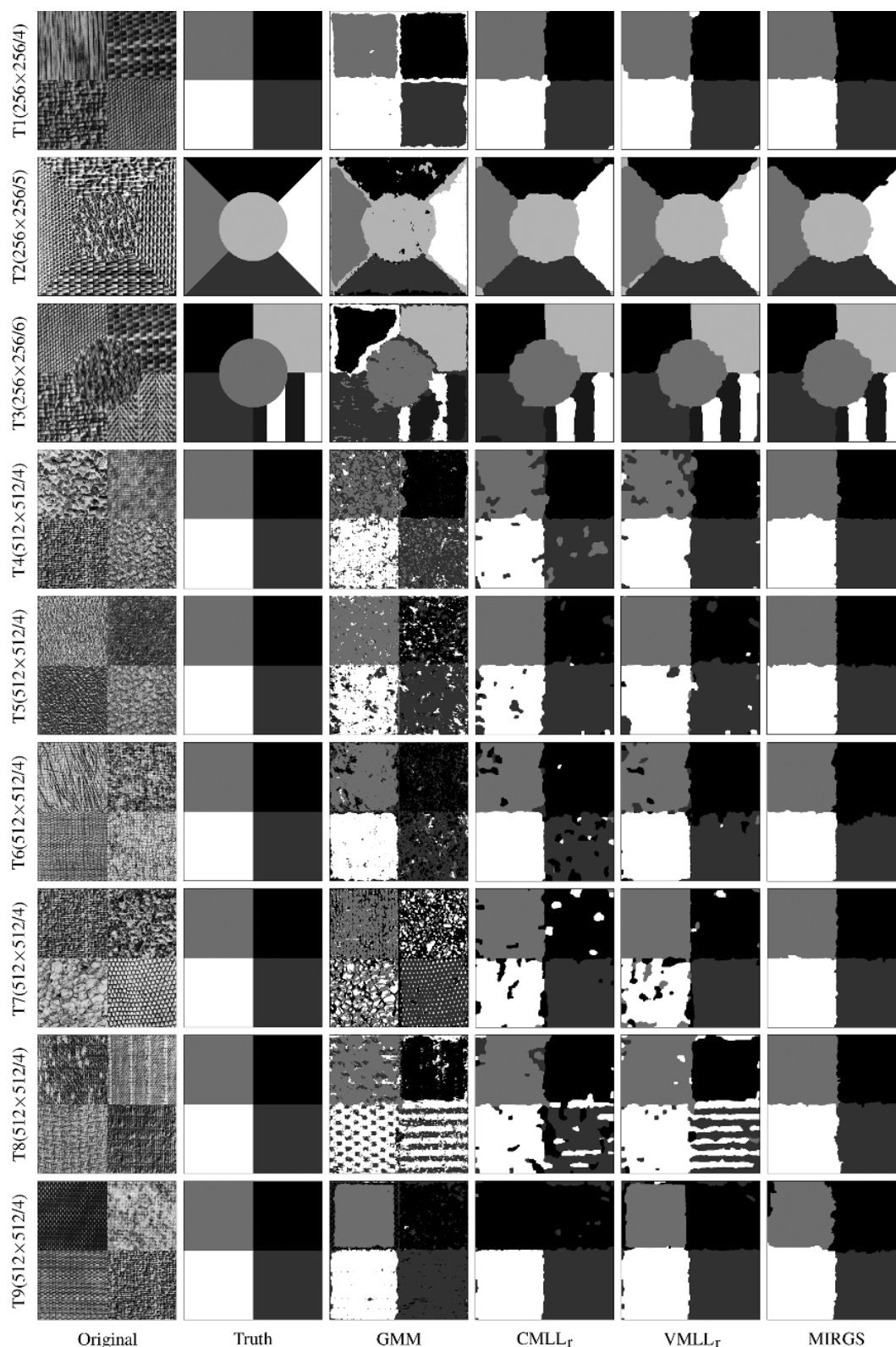


Fig. 5. Segmentation results on three groups of nine textured images using GMM, CMLL_r, VMML_r and MIRGS. Each row refers to one textured image (T1 to T9 from top to bottom with image size/number of classes in bracket). Each column from left to right shows the original image, segmentation ground truth and segmentation results obtained by GMM, CMLL_r, VMML_r and MIRGS. MIRGS consistently gives the most visually agreeable results.

Q3 with respect to both acc and κ are large. Discrepancies between average and best segmentation results, as indicated by differences between Mean and Max regarding both acc and κ , are also large. In fact, although very promising segmentation

results corresponding to the high Max of acc and κ can be obtained, extremely poor results for RAND indicated by the low Min of acc and κ reveal the occurrence of segmentation failure, as shown in Fig. 4.

The GMM-based initialization method is less sensitive to initial random seeds. Accordingly, MIRGS initialized by GMM yields consistent segmentation performance, which is demonstrated in Table II by small interquartile ranges as well as small differences between Mean and Max with respect to both acc and κ for all three test images. However, GMM has high computational cost. As indicated in Table III, the GMM-based initialization is very time-consuming. For all three test images, its computation time is longer than the subsequent segmentation part.

Since no spatial context information is taken into account, GMM may provide unfavorable initial conditions for MIRGS and, thus, lead to unsatisfactory segmentation results. For the univariate image, the GMM-based initialization facilitates MIRGS to achieve very good performance in all of 10 runs (Table II). For the multivariate color and texture feature images, the GMM-based initialization consistently generates unsatisfactory segmentation results. GMM fails in all of 10 runs for the color image, with one such example shown in Fig. 4. In fact, one of three class mean vectors and covariance matrices derived by GMM corresponds to a noise-induced class with relatively small size. According to the initial \mathbf{x}^r determined using the feature model, only a few vertices in the initial RAG are assigned with this noise-induced class. Since most neighbors of these vertices belong to other classes, MIRGS inevitably yields the segmentation result with one class missed, causing failure.

The RKM-based initialization method consistently provides accurate initial conditions for MIRGS. Table II indicates MIRGS incorporating RKM yields strong performance in all of 10 runs for all three test images. Regarding both acc and κ , high Max and Min, small interquartile ranges, and small differences between Mean and Max are noted. Even the poorest result corresponding to the Min of acc and κ as shown in Fig. 4 is visually close to ground truth. RKM has low computational cost. As depicted in Table III, for all three test images, computation time of RKM is significantly shorter than that of GMM, and is trivial compared to that of the subsequent MIRGS.

As such, RKM is the most effective method to suppress initialization sensitivity among the three schemes.

D. Experimental Setup—Segmentation Analysis

Four segmentation methods—GMM, CMLL_T, VMLL_T and MIRGS—will be compared using a variety of multivariate images (two groups of 10 synthetic color images $SC_{\sigma_n^2}^A$ and $SC_{\sigma_n^2}^M$, three categories of nine textured images T1–T9, and three natural color images NC1–NC3) to investigate the influence of intraclass variation on segmentation performance.

GMM uses the same formulation and parameter setting as in initialization analysis, which is a base case to refer the success of the other methods. The other three methods all use the RKM-based initialization and the β adjustment rule (22)–(25) with the same parameter settings for a fair performance comparison. For VMLL_T, a is set to 80 as in [15]. Gibbs sampler with a simulated annealing schedule $T(\tau) = 2 \times 0.98^{\tau-1}$ is applied using 300 iterations to solve CMLL_T and VMLL_T. RKM and MIRGS use the same parameter settings as in initialization analysis. For all segmentation methods, the number of classes C is specified *a priori*.

Segmentation performance is measured using acc and κ if ground truth accompanies test images. Computation time is recorded to evaluate computational speed.

For each test image, each segmentation method is applied 10 times starting from different random seeds while all four methods share the same random seed with respect to any individual run.

E. Tests—Segmentation Accuracy

For synthetic color and textured images accompanied by ground truth, mean values of acc and κ over 10 runs are reported in Tables IV and V. The Wilcoxon signed-rank test [47] with significance level 0.05 is applied on values of acc or κ in 10 runs regarding any two segmentation algorithms to determine whether the performance difference between these two algorithms is statistically significant. For each test image, the segmentation method with the highest mean acc and κ values is reported in bold. Any method that generates mean acc and κ values not statistically significantly different from the highest is also shown in bold. In Tables IV and V, MIRGS performs best for all test images.

For synthetic color images (Table IV), performance of all segmentation algorithms reduces as the noise level increases. This is most noticeable for GMM, moderate for CMLL_T, and slight for VMLL_T and MIRGS. For textured image (Table V), GMM has the poorest performance. CMLL_T and VMLL_T perform well on Category 1, but unsatisfactorily on the more difficult categories 2 and 3. Fig. 5 illustrates one segmentation example with respect to each segmentation algorithm applied on every textured image to provide direct visual perception. All segmentation algorithms produce consistent segmentation results over 10 runs. Obviously, all segmentation algorithms except MIRGS yield spotty segmentation results especially along class and image borders where higher intraclass variation may occur. MIRGS generates visually agreeable segmentation results on all of nine textured images.

For natural color images without ground truth, segmentation results have to be evaluated via visual perception. Fig. 6 shows one example of segmentation results in 10 runs with respect to each segmentation algorithm applied on every natural color image. All segmentation algorithms produce consistent segmentation results over 10 runs. Different classes are separated by closed boundary lines with distinct colors. For image NC1, MIRGS successfully partitions the scene into three classes (horses, green grass and yellow flowers). CMLL_T cannot effectively differentiate grass and flowers and incorrectly assigns some body parts of horses (white spot on the head of the large horse and hooves in the forelegs of both small and large horses) to the class of grass. VMLL_T also mistakenly splits some body parts of horses (white spot on the head and one hoof in the foreleg of the large horse). For image NC2, three classes (grass field, car and raceway) are well demarcated by MIRGS. CMLL_T splits the car into several pieces and incorrectly assign some pieces to the class of grass field. VMLL_T roughly separates three classes with large errors occurring around class borders. For image NC3, MIRGS generally separates the four classes (lake including a small sailboat inside, mountain, blue sky and clouds) except for a small region of dark cloud that is

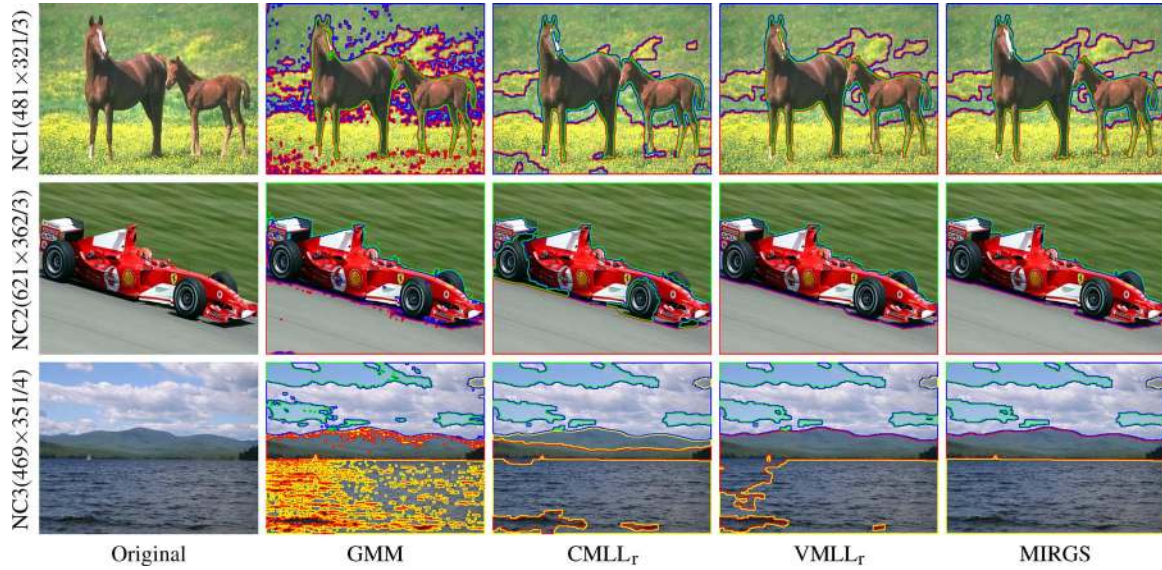


Fig. 6. Segmentation results on three natural color images using GMM, CMLL_r, VMML_r and MIRGS with respect to one example run. Each row refers to one natural color image (NC1 to NC3 from top to bottom with image size/number of classes to be segmented in bracket). Each column from left to right denotes the original image and segmentation results obtained by GMLL, CMLL_r, VMML_r and MIRGS. Class boundaries are outlined in distinct colors. MIRGS consistently gives the most visually agreeable results.

TABLE VI
MEAN (STANDARD DEVIATION) OF COMPUTATION TIME (MINUTES) AVERAGED OVER 10 RUNS BY APPLYING GMM, CMLL_r, VMML_r AND MIRGS ON TWO GROUPS OF 10 SYNTHETIC COLOR IMAGES WITH ADDITIVE AND MULTIPLICATIVE GAUSSIAN NOISE AT VARYING NOISE LEVELS

Images		GMM	CMLL _r	VMML _r	MIRGS
$SC_{\sigma_n^2}^A$	σ_n^2				
	0.01	2.86(0.07)	0.68(0.02)	0.70(0.02)	0.99(0.01)
	0.03	4.58(0.08)	0.75(0.01)	0.76(0.02)	1.10(0.02)
	0.05	7.98(0.13)	0.77(0.01)	0.78(0.01)	1.18(0.01)
	0.07	7.51(0.09)	0.80(0.02)	0.81(0.01)	1.24(0.02)
	0.09	8.27(0.08)	0.80(0.01)	0.82(0.01)	1.21(0.01)
$SC_{\sigma_n^2}^M$	0.1	2.72(0.10)	0.75(0.01)	0.76(0.01)	0.94(0.01)
	0.3	3.99(0.12)	0.80(0.01)	0.82(0.01)	1.08(0.01)
	0.5	3.39(0.07)	0.87(0.01)	0.88(0.01)	1.22(0.02)
	0.7	7.03(0.08)	0.83(0.01)	0.85(0.01)	1.29(0.01)
	0.9	4.89(0.07)	0.87(0.02)	0.88(0.01)	1.30(0.02)

incorrectly assigned to the lake class due to the strong feature similarity. This error similarly occurs in the results obtained by the other three methods. CMLL_r and VMML_r cannot properly differentiate lake and mountain. For all three test images, the GMM always yields spotty segmentation results due to intraclass variation.

Overall, MIRGS has the strongest capability of handling intraclass variation among four algorithms.

F. Tests—Segmentation Computation Time

MIRGS adopts semantic region growing to yield a hierarchical RAG based image representation with bottom-up nonincreasing vertex numbers. Although such a hierarchical structure reduces the size of the solution space and, thus, tends to speed up segmentation, its construction may be time-consuming. Computational complexity regarding region merging and the subsequent updating of the RAG depends upon the complexity of the image under segmentation.

TABLE VII
MEAN (STANDARD DEVIATION) OF COMPUTATION TIME (MINUTES) AVERAGED OVER 10 RUNS BY APPLYING GMM, CMLL_r, VMML_r AND MIRGS ON THREE CATEGORIES OF NINE TEXTURED IMAGES. “CAT” STANDS FOR “CATEGORY”

Images		GMM	CMLL _r	VMML _r	MIRGS
Cat 1	T1	21.63(0.50)	5.39(0.71)	5.24(0.25)	32.94(0.96)
	T2	28.13(0.93)	7.37(0.43)	7.48(0.64)	18.08(0.51)
	T3	53.82(2.16)	7.88(0.02)	7.91(0.02)	29.76(0.20)
Cat 2	T4	121.77(2.91)	18.46(0.06)	18.48(0.06)	84.28(4.26)
	T5	206.18(1.58)	20.68(0.06)	20.70(0.07)	103.37(2.44)
	T6	105.93(0.98)	19.45(0.11)	19.43(0.08)	138.33(7.20)
Cat 3	T7	170.52(2.05)	16.74(0.05)	16.77(0.05)	38.91(1.53)
	T8	120.03(2.50)	22.49(0.21)	22.45(0.06)	109.75(4.62)
	T9	103.36(2.24)	19.54(0.72)	19.37(0.55)	120.15(13.13)

TABLE VIII
MEAN (STANDARD DEVIATION) OF COMPUTATION TIME (MINUTES) AVERAGED OVER 10 RUNS BY APPLYING GMM, CMLL_r, VMML_r AND MIRGS ON THREE NATURAL COLOR IMAGES

Images	GMM	CMLL _r	VMML _r	MIRGS
NC1	5.11(0.07)	0.56(0.01)	0.57(0.01)	0.26(0.01)
NC2	3.81(0.19)	0.69(0.01)	0.70(0.01)	0.99(0.01)
NC3	8.85(0.20)	0.77(0.03)	0.79(0.02)	0.36(0.01)

Tables VI, VII and VIII report computation time (minutes) with respect to the segmentation experiments in Section IV-E. Generally speaking, MIRGS has reasonable computation time for all test images, which is comparable to that of GMM. In most cases except for natural color images NC1 and NC3, MIRGS takes longer computation time than CMLL_r and VMML_r due to the computational cost associated with the region growing procedure, however this extra computing time is warranted to capture accurate segmentation.

V. CONCLUSION

We present a MRF-based multivariate segmentation algorithm named MIRGS, which extends the applicability of IRGS

to multivariate images while inheriting the merits of IRGS. To suppress initialization sensitivity, MIRGS uses a RKM-based initialization method, which consistently provides accurate initial conditions at low computational cost. The superiority of RKM relative to two commonly used initialization schemes has been demonstrated on images with different feature space dimensions. For a variety of synthetic and natural multivariate images, MIRGS consistently achieves the highest segmentation accuracy compared to three other published algorithms.

Computation time of MIRGS is closely related to the construction time of the hierarchical RAG. Therefore, fast methods to establish the hierarchical RAG are important for future investigation. Moreover, to automatically determine the number of classes instead of prespecifying it as an algorithmic parameter is desirable. Some previous attempts [32] made in the context of the MRF-based segmentation deserve being investigated under the MIRGS framework. In addition, MIRGS mentioned in this paper is performed in an unsupervised manner, which does not take into account any domain knowledge. However, the domain knowledge can be easily incorporated into (15) and used in step 5 in Table I to produce the supervised MIRGS, which deserves future investigation.

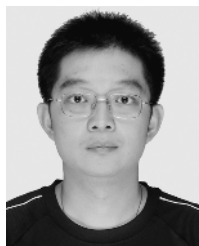
ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their contributing comments.

REFERENCES

- [1] M. C. Clark *et al.*, "Automatic tumor segmentation using knowledge-based techniques," *IEEE Trans. Med. Imag.*, vol. 17, no. 2, pp. 187–201, Apr. 1998.
- [2] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2071–2082, Dec. 2000.
- [3] Y. Wang, K. F. Loe, T. Tan, and J. K. Wu, "Spatiotemporal video segmentation based on graphical models," *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 937–947, Jul. 2005.
- [4] Q. Yu and D. A. Clausi, "SAR sea-ice image analysis based on iterative region growing using semantics," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3919–3931, Dec. 2007.
- [5] S. Z. Li, *Markov Random Field Modeling in Image Analysis*, 2nd ed. New York: Springer-Verlag, 2001.
- [6] Q. Yu and D. A. Clausi, "IRGS: Image segmentation using edge penalties and region growing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2126–2139, Dec. 2008.
- [7] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, 3rd ed. Andover, U.K.: Thomson Learning, 2008.
- [8] J. C. Noordam, W. H. A. M. van den Broek, and L. M. C. Buydens, "Multivariate image segmentation with cluster size insensitive fuzzy c-means," *Chemomet. Intell. Lab. Syst.*, vol. 64, no. 1, pp. 65–78, Oct. 2002.
- [9] J. Shah, "Curve evolution and segmentation functionals: Application to color images," in *Proc. Int. Conf. Image Processing (ICIP'96)*, Lausanne, Switzerland, Sep. 1996, pp. 461–464.
- [10] G. Sapiro, "Color snakes," *Comput. Vis. Imag. Understand.*, vol. 68, no. 2, pp. 247–253, Nov. 1997.
- [11] W. Y. Ma and B. S. Manjunath, "EdgeFlow: A technique for boundary detection and image segmentation," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1375–1388, Aug. 2000.
- [12] R. Nock and F. Nielsen, "Statistical region merging," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 11, pp. 1452–1458, Nov. 2004.
- [13] Q. M. Luo and T. M. Khoshgoftaar, "Unsupervised multiscale color image segmentation based on MDL principle," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2755–2761, Sep. 2006.
- [14] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [15] H. Deng and D. A. Clausi, "Unsupervised image segmentation using a simple MRF model with a new implementation scheme," *Pattern Recognit.*, vol. 37, no. 12, pp. 2323–2335, 2004.
- [16] S. C. Zhu and A. L. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 9, pp. 884–900, Sep. 1996.
- [17] J. P. Fan, D. K. Y. Yau, A. K. Elmagarmid, and W. G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1454–1466, Oct. 2001.
- [18] R. Kindermann and J. L. Snell, *Markov Random Fields and Their Applications*. Providence, RI: American Mathematical Society, 1980.
- [19] I.-Y. Kim and H.-S. Yang, "A systematic way for region-based image segmentation based on Markov random field model," *Pattern Recognit. Lett.*, vol. 15, no. 10, pp. 969–976, Oct. 1994.
- [20] F. G. Yang and T. Z. Jiang, "Pixon-based image segmentation with Markov random fields," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1552–1559, Dec. 2003.
- [21] G. S. Xia, C. He, and H. Sun, "Integration SAR image segmentation method using Markov random field on region adjacency graph," *IET Radar, Sonar Nav.*, vol. 1, no. 5, pp. 348–354, 2007.
- [22] S.-S. Huang, L.-C. Fu, and P.-Y. Hsiao, "Region-level motion-based background modeling and subtraction using MRFs," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1446–1456, May 2007.
- [23] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.
- [24] M. Suk and S. Chung, "A new segmentation technique based on partition mode test," *Pattern Recognit.*, vol. 16, no. 5, pp. 469–480, 1983.
- [25] B. Kartikeyan and A. Sarkar, "A unified approach for image segmentation using exact statistics," *Comput. Vis. Graph. Image Process.*, vol. 48, no. 2, pp. 217–229, Nov. 1989.
- [26] X. Ren and J. Malik, "Learning a classification model for segmentation," in *Proc. IEEE Int. Conf. Computer Vision (ICCV'03)*, Nice, France, Oct. 2003, pp. 10–17.
- [27] A. P. Moore, S. J. D. Prince, J. Warrell, U. Mohammed, and G. Jones, "Superpixel lattices," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'08)*, Anchorage, AK, Jun. 2008, pp. 1–8.
- [28] R. Fjøltoft *et al.*, "Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 675–686, Mar. 2003.
- [29] Q. Yu, "Automated SAR Sea Ice Interpretation" Ph.D. dissertation, Univ. Waterloo, Waterloo, ON, Canada, May 2006 [Online]. Available: <http://www.eng.uwaterloo.ca/~dclausi/students.html>
- [30] K. E. Melkemia, M. Batoucheb, and S. Fofouf, "A multiagent system approach for image segmentation using genetic algorithms and extremal optimization heuristics," *Pattern Recognit. Lett.*, vol. 27, no. 11, pp. 1230–1238, Aug. 2006.
- [31] J. K. Fwu and P. M. Djurić, "EM algorithm for image segmentation initialized by a tree structure scheme," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 349–352, Feb. 1997.
- [32] T. N. Tran, R. Wehrens, D. H. Hoekman, and L. M. C. Buydens, "Initialization of Markov random field clustering of large remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1912–1919, Aug. 2005.
- [33] D. X. Xu and J. N. Hwang, "A topology independent active contour tracking," in *Proc. IEEE Int. Workshop on Neural Networks for Signal Processing*, Madison, WI, Aug. 1999, pp. 429–438.
- [34] H. C. Lee and D. R. Cok, "Detecting boundaries in a vector field," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1181–1194, May 1991.
- [35] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Hoboken, NJ: Wiley, 2000.
- [36] E. Rignot and R. Chellappa, "Segmentation of polarimetric synthetic aperture radar data," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 281–300, Jul. 1992.
- [37] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random field using simulated annealing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 8, pp. 799–813, Aug. 1989.
- [38] X. Descombes, R. D. Morris, J. Zerubia, and M. Berthod, "Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood," *IEEE Trans. Image Process.*, vol. 8, no. 7, pp. 954–963, Jul. 1999.
- [39] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*, 1st ed. Berlin, Germany: Springer-Verlag, 2000.

- [40] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.
- [41] D. A. Clausi and M. E. Jernigan, "Designing Gabor filters for optimal texture separability," *Pattern Recognit.*, vol. 33, no. 11, pp. 1835–1849, 2000.
- [42] D. A. Clausi, "K-means Iterative Fisher (KIF) unsupervised clustering algorithm applied to image texture segmentation," *Pattern Recognit.*, vol. 35, no. 9, pp. 1959–1972, 2002.
- [43] T. Randen and J. H. Husoy, "Filtering for texture classification: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 4, pp. 291–310, Jul. 1999.
- [44] P. Yu, "Segmentation of Radarsat-2 Dual Polarization Sea Ice Imagery," Master's thesis, Univ. Waterloo, Waterloo, ON, Canada, Aug. 2009.
- [45] W. D. Hudson and C. W. Ramm, "Correct formulation of the kappa coefficient of agreement," *Photogram. Eng. Remote Sens.*, vol. 53, no. 4, pp. 421–422, 1987.
- [46] D. A. Clausi, A. K. Qin, M. S. Chowdhury, P. Yu, and P. Maillard, "MAGIC: MAP-Guided Ice Classification," *Can. J. Remote Sens.*, to be published.
- [47] S. Siegel, *Non-Parametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.



A. K. Qin (M'05) received the B.E. degree from Southeast University, Nanjing, China, in 2001, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2007.

During the Ph.D. study, he was with the Illinois Genetic Algorithms Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL, as a visiting scholar for three months in 2005, and worked as a project officer at Nanyang Technological University, Singapore, from 2006 to 2007. After obtaining the Ph.D. degree, he joined the Department of Sys-

tems Design Engineering at the University of Waterloo, Waterloo, ON, Canada, as a postdoctoral fellow from 2007 to 2009. He is currently a faculty member with the School of Automation at Southeast University, Nanjing, China. His research interests include pattern recognition, image processing, computer vision, machine learning, evolutionary computation, remote sensing data analysis, biometrics, and bioinformatics.



David A. Clausi (S'93–M'96–SM'03) received the B.A.Sc., M.A.Sc., and Ph.D. degrees in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1990, 1992, and 1996, respectively.

After completing his doctorate, he worked in the medical imaging field at Agfa (Waterloo). He started his academic career in 1997 as an Assistant Professor in Geomatics Engineering at the University of Calgary, Alberta, Canada. In 1999, he returned to his alma mater and is now a Professor specializing in the fields of Intelligent and Environmental Systems. He is an active interdisciplinary and multidisciplinary researcher. He has an extensive publication record, publishing refereed journal and conference papers in the diverse fields of remote sensing, computer vision, algorithm design, and biomechanics. His primary research interest is the automated interpretation of synthetic aperture radar (SAR) sea ice imagery, in support of operational activities. His research efforts have led to successful commercial implementations.

Dr. Clausi was the co-chair of IAPR Technical Committee 7—Remote Sensing during 2004–2006. He has received numerous scholarships, paper awards, and two Teaching Excellence Awards.