Int. J. Appl. Math. Comput. Sci., 2010, Vol. 20, No. 4, 711–726
DOI: 10.2478/v10006-010-0054-y

INTEGRATED REGION–BASED SEGMENTATION USING COLOR COMPONENTS AND TEXTURE FEATURES WITH PRIOR SHAPE KNOWLEDGE

MEHRYAR EMAMBAKHSH, HOSSEIN EBRAHIMNEZHAD, MOHAMMAD HOSSEIN SEDAAGHI

Faculty of Electrical Engineering
Sahand University of Technology, Sahand New City, Tabriz, Iran
e-mail: {m_emambakhsh,ebrahimnezhad,sedaaghi}@sut.ac.ir

Segmentation is the art of partitioning an image into different regions where each one has some degree of uniformity in its feature space. A number of methods have been proposed and blind segmentation is one of them. It uses intrinsic image features, such as pixel intensity, color components and texture. However, some virtues, like poor contrast, noise and occlusion, can weaken the procedure. To overcome them, prior knowledge of the object of interest has to be incorporated in a top-down procedure for segmentation. Consequently, in this work, a novel integrated algorithm is proposed combining bottom-up (blind) and top-down (including shape prior) techniques. First, a color space transformation is performed. Then, an energy function (based on nonlinear diffusion of color components and directional derivatives) is defined. Next, signed-distance functions are generated from different shapes of the object of interest. Finally, a variational framework (based on the level set) is employed to minimize the energy function. The experimental results demonstrate a good performance of the proposed method compared with others and show its robustness in the presence of noise and occlusion. The proposed algorithm is applicable in outdoor and medical image segmentation and also in optical character recognition (OCR).

Keywords: image segmentation, prior shape knowledge, level set, nonlinear diffusion, energy minimization.

1. Introduction

The advent of segmentation has introduced an efficient alternative to partitioning schemes for image processing. Its main goal is to separate image into disjoint, non-overlapping regions where each one has a uniformity in its predefined feature space. There exist many approaches to image segmentation in the literature, including watershed transformation (Kuo *et al*., 2008; Lu *et al*., 2006; Hrebień *et al*., 2008; Emambakhsh and Sedaaghi, 2009), graph-cuts (Greig *et al*., 1989), expectation maximization (EM) (Ramme *et al*., 2009), and region growing (Susomboon *et al*., 2006; Petera *et al*., 2008).

These methods generally have a high computational complexity making them unsuitable for multi-phase segmentation. Also, they are highly sensitive to local minima in their optimization steps. Moreover, watershed-based algorithms are greatly responsive to noise and often produce over-segmentation (Vincent and Soille, 1991).

Therefore, denoising and region merging algorithms are mainly employed as post- and pre-processing stages to reduce the number of regions. This will undoubtedly weaken the robustness of the algorithm while increasing the computational complexity.

On the other hand, two other approaches to segmentation are energy minimization (implemented by variational frameworks) and clustering methods. Medical image scission is one of the most popular fields of study where clustering algorithms have mainly been utilized. Some examples include: fuzzy C-means (FCM) (Wang *et al*., 2008), hierarchical clustering (Lai and Chang, 2009), Hopfield competitive learning networks (Kuo *et al*., 2008), K-means and self-organizing maps (SOMs) (Dokur *et al*., 2006; Ong *et al*., 2002).

Low computational complexity is one of the most important gains of clustering-based techniques. However, they have some serious disadvantages that must be taken into account. They are sensitive to noise and outliers in the feature space. They mainly acquire pixels' local information making them fragile in the presence of noise. On the other hand, energy minimization methods can measure image regional and statistical information resulting in a more meaningful and reasonable segmentation (Cremers *et al*., 2007). Moreover, regional-based methods are less

sensitive to noise. Algorithms based on clustering are generally more appropriate for blind segmentation where only intrinsic image features (e.g., gray level, color or texture) are considered. In other words, using prior knowledge is not usually straightforward in these approaches.

Conversely, a completely different procedure is employed in energy minimization approaches. An energy function, defined from image features, is targeted to be minimized. The minimization is implemented using a variational framework where active contours (level set (Osher and Sethian, 1988) and snakes (Kass *et al*., 1988)) are commonly applied. Splitting and merging a contour (or modifying its topology) have made level sets attractive for segmentation compared with snakes in recent studies (Cremers *et al*., 2007). They are less sensitive to initialization. Also, prior knowledge and probabilistic interpretations are easily engaged.

The major task in the level set is to minimize an energy function by evolving a contour implicitly. The nature of such a function is controversial in the literature. It is defined by the features extracted from the input image, where the features are simply the intensity of the pixels, color components, texture features, motion vectors computed from consecutive video frames by solving the optical flow equation (Mitiche and Sekkati, 2006; Hao *et al*., 2007) or even the integration of these features (Cremers *et al*., 2007). In the works of Wang and Vemuri (2004) as well as Feddern *et al*. (2006), structure tensors are employed for texture segmentation. Raja *et al*. (2010) used Gabor-wavelets for feature extraction from textured images.

A similar approach is reported by Andrysiak and Choras (2005), Sagiv *et al*. (2006) or Sandler and Lindenbaum (2006) to define an energy function for contour evolution based on the outputs of Gabor filters. However, feature extraction and defining the energy function based on structure tensors and Gabor filters have some disadvantages. Firstly, since image gradients need to be computed in structure tensors, a Gaussian kernel is utilized for denoising. Such smoothing degrades the edges leading to a weak segmentation. The same problem exists in Gabor filters. Moreover, a high dimensional feature space is required in both methods, which results in an increase in computational complexity. Perona and Malik (1990) and also Gerig *et al*. (1992) cured these problems (i.e., dimensionality and smoothing) using nonlinear diffusion. Rousson *et al*. (2003) and Cremers *et al*. (2007) put this method to work for texture segmentation. Nonlinear diffusion solely might not be sufficient for a successful segmentation. Lots of circumstances, such as texture inhomogeneity, poor image contrast, missing or diffused boundaries (Tsai *et al*., 2003), noise, intensity non-uniformity (Kuo *et al*., 2008) and occlusion can result in inappropriate segmentation. Therefore, prior knowledge of the object of interest has to be considered in segmentation.

In the work of Tsai *et al*. (2003), a method is proposed to use shape as prior knowledge for medical image segmentation. It is defined for both binary and grey scale images. To improve the efficiency of their approach, we have proposed a hybrid region-based method integrating top-down and bottom-up processes. Moreover, we employ a prior shape knowledge of the object of interest, texture features and color channels for segmentation. The proposed method has various applications in outdoor and medical image segmentation, as will be discussed in the simulation results. Our algorithm's robustness in the presence of noise and occlusion has also been evaluated. Preliminary results of this research appeared in the work of Emambakhsh *et al*. (2010).

This paper is organized as follows. Section 2 introduces the proposed method. The experimental results are provided in Section 3. Section 4 contains the conclusion.

2. Proposed method

In this section, we develop a novel scheme of image segmentation. It integrates the prior shape information of the object of interest with intensity information extracted from a textured color image using a nonlinear diffusion filter, in a level set framework. Figure 1 visualizes the block diagram of the proposed algorithm. A brief description of the proposed method is as follows. First of all, the input image is transformed from the RGB into the CIE L*a*b* color space. Then the directional derivatives of the L* channel are computed. The directional derivatives and L*a*b* color components are considered as inputs for nonlinear diffusion in the next step. The result of diffusion is a feature space and it is used in computing the energy function for level set evolution. In parallel with the

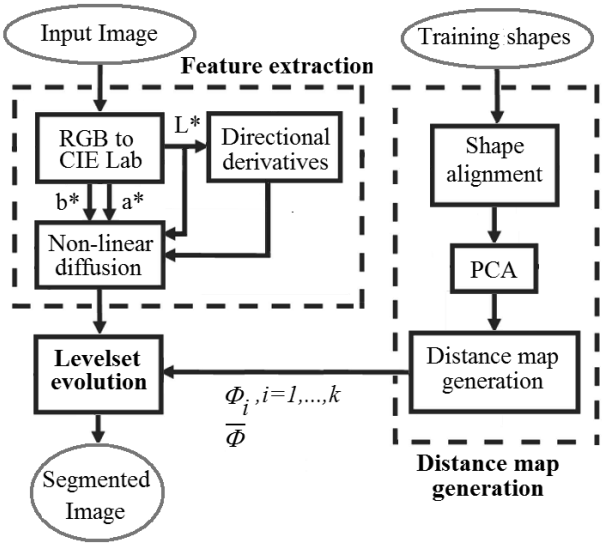

Fig. 1. Block diagram of the proposed algorithm.

above procedure, an algorithm similar to that presented by Tsai *et al*. (2003) is employed to create the distance maps for level set functions using training shapes. This process originates considering binary images including shape information. A shape alignment is next utilized on the training shapes. Then, superior shapes are captured by applying principal component analysis (PCA). Finally, distance maps are engaged to set up the level set functions from the major shapes. The produced distance maps are used in active contour evolution. After minimizing the energy function, the inner part of the contour will correspond to object location. Also, the outer part conceptualizes the background. The proposed method includes the following steps:

1. feature space construction,
2. distance map generation,
3. level set evolution.

2.1. Feature space construction. The feature space is constructed using the following stages:

- RGB to L*a*b* conversion,
- directional derivatives,
- nonlinear diffusion.

2.1.1. RGB to L*a*b* conversion. It can be understood from McAdams ellipses (Forsyth and Ponce, 2002) that the RGB color space is not a good choice for color image processing (Ong *et al*., 2002). Moreover, a feature space based on RGB components (and, generally speaking, linear color spaces) is highly correlated. Therefore, the background and foreground regions cannot be distinguished precisely. Consequently, the CIE L*a*b* color space with a nonlinear uniform domain is employed to produce a more detachable feature space.

2.1.2. Directional derivatives. The directional derivatives are computed on the L* color component. It helps to find the most salient edges in different directions. There exist reports on using only two directions for gradient computation (the gradients in the x and y directions) (Rousson *et al*., 2003; Cremers *et al*., 2007). We consider employing more directions because this significantly improves the result of the segmentation, especially for more complicate textures. Let I^L denote the L* color channel. Then, the output of the directional derivative will be

$$\begin{cases} u_{\theta_i} = \left[{I_x}^L \cos(\theta_i) + I_y^L \sin(\theta_i)\right]^2, \\ \theta_i = \theta_0 + (i)\Delta\theta, \\ \forall i = 0, 1, \ldots, n-1, \end{cases} \tag{1}$$

where I_x^L and I_y^L are the I^L derivatives along the x and y directions, respectively; u_{θ_i} is the directional derivative computed along θ_i; θ_0 represents the initial value measured (in degrees) in a counter-clockwise direction from the horizontal axis; θ_i is the i-th update used in computing the derivative; n is the number that specifies the total number of slices for dividing the unique circle. The values for L^*, a^*, b^* and also the vectors for the directional derivatives are considered as inputs for nonlinear diffusion. Therefore, the input matrix for the nonlinear diffusion is as follows:

$$u = \left(u_{\theta_0}, u_{\theta_1}, \ldots, u_{\theta_{n-1}}, I^L, I^a, I^b\right), \tag{2}$$

where I^L, I^a, and I^b represent the L*, a*, and b* color channels, respectively.

2.1.3. Nonlinear diffusion. Nonlinear diffusion is a method for image denoising and simplification. It was initially proposed by Perona and Malik (1990) for edge detection and its vector-valued version was presented by Gerig *et al*. (1992). We use it in setting up the feature space. The proposed feature space improves the robustness of segmentation in the vicinity of noise. The feature vectors (Eqn. (2)) is considered as an input for nonlinear diffusion. It is defined as in the work of Rousson *et al*. (2003):

$$\partial_t\, u_i' = \operatorname{div}\left[g\left(\sum_{k=0}^{N-1} |\nabla u_k|^2\right)\nabla u_i\right], \tag{3}$$

where ${u_i}'$ is the result of diffusion when the input vector is u_i; $g(\cdot)$ is a decreasing function of the gradient of the image; N denotes the dimension of the feature vector. We have used the following function in our research:

$$g_m(s) = \begin{cases} 1 - \exp\left(-\dfrac{C_m}{(s/\lambda)^m}\right), & s \geq 0, \\ 1, & s < 0, \end{cases} \tag{4}$$

where m is a constant used to define the damping rate; C_m is put in the equation to make the flux $S \cdot g(s)$ increase for $g < \lambda$ and decrease for $g \geq \lambda$. Also, λ is a parameter for controlling the contrast. Compared with the functions suggested by Rousson *et al*. (2003) and Cremers *et al*. (2007) for $g(\cdot)$, our function proposed in Eqn. (4) has more free parameters and flexibility. Solving Eqn. (3) results in a $(n+3)$-dimension feature vector, i.e.,

$$u' = \left(u_0', u_1', \ldots, u_{n+2}'\right). \tag{5}$$

The feature vector is used to compute the internal energy in level set evolution as explained in Section 2.3.

2.2. Distance map generation. The algorithm explained in the work of Tsai *et al*. (2003) is utilized to generate distance maps. It includes three stages. **Stage 1.** In the first step, binary images including the shape information of the object of interest are provided. These shapes have to be aligned. There are n training shapes. One of the images is considered as the reference while the others are translated, rotated and scaled in order to be aligned accordingly. In this paper, we have employed the alignment method of Tsai *et al*. (2003). The training set consists of n binary images $\{I^1, I^2, \ldots, I^n\}$ with pixel values equal to one for the inside, and zero for the outside of the object. The purpose is to find pose parameters, $\{\mathbf{P^1}, \mathbf{P^2}, \ldots, \mathbf{P^n}\}$. They are used to align all of the n training images according to the reference. $\mathbf{P}$ is a set of parameters related to scaling, translation and rotation.

Let $\tilde{I}$ denote the result of the transformation. Therefore, for each image, I, the transformed image will be (Tsai *et al*., 2003)

$$\begin{cases} \tilde{I}\left(\tilde{x}, \tilde{y}\right) = I(x, y), \\ \\ \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{bmatrix} = \mathrm{T}[\mathbf{P}] \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \end{cases} \tag{6}$$

where the transformation matrix $\mathrm{T}[\mathbf{P}]$ is determined by the following affine structure (Jain, 1989; Tsai *et al*., 2003):

$$\begin{aligned} \mathrm{T}[\mathbf{P}] = &\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & s_x & 0 \\ s_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &\cdot \begin{bmatrix} h_x & 0 & 0 \\ 0 & h_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{7}$$

The first matrix in Eqn. (7) is for x as well as y translations. The second one represents the shear transformation along the x and y directions. The third matrix is for scaling along the x and y directions. Finally, the fourth one is dedicated to rotation. Then, one of the training images is considered the reference ($\tilde{I}^j$) and the following energy is minimized to find the efficient pose parameters for image $\tilde{I}^i$ (Tsai *et al*., 2003):

$$E_{\text{align}} = \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \frac{\int\int_{\Omega} \left(\tilde{I}^i - \tilde{I}^j\right)^2 \mathrm{dA}}{\int\int_{\Omega} \left(\tilde{I}^i + \tilde{I}^j\right)^2 \mathrm{dA}}, \tag{8}$$

where Ω is the whole domain of the image. The gradient descent algorithm is employed to find the optimum values for the pose parameters, $\mathbf{P}$. Therefore, the gradient of Eqn. (8) with respect to $\mathbf{P^i}$ will be (Tsai *et al*., 2003)

$$\begin{aligned} &\nabla^i_{\mathbf{P}} E_{\text{align}} \\ &= \sum_{j=1, j\neq i}^{n} \left\{ \frac{2 \int\int_{\Omega} \left(\tilde{I}^i - \tilde{I}^j\right) \nabla^i_{\mathbf{P}} \tilde{I}^i \mathrm{dA}}{\int\int_{\Omega} \left(\tilde{I}^i + \tilde{I}^j\right)^2 \mathrm{dA}} \right. \\ &\left. - \frac{2 \int\int_{\Omega} \left(\tilde{I}^i - \tilde{I}^j\right)^2 \mathrm{dA} \int\int_{\Omega} \left(\tilde{I}^i + \tilde{I}^j\right) \nabla_{\mathbf{P}}{}^i \tilde{I}^i \mathrm{dA}}{\left(\int\int_{\Omega} \left(\tilde{I}^i + \tilde{I}^j\right)^2 \mathrm{dA}\right)^2} \right\} \end{aligned} \tag{9}$$

Stage 2. As the second stage, a distance map is computed for each training shape. In this study, the signed distance function (SDF) is employed to generate the distance maps, i.e., there are $\{\Psi_1, \Psi_2, \ldots, \Psi_n\}$ as SDFs for each aligned image (Osher and Sethian, 1988; Tsai *et al*., 2003). The zero levels of these SDFs are considered as the boundaries of the training shapes.

Stage 3. Finally, PCA is employed to extract the prominent shapes. First of all, the average of SDFs is computed and removed from the distance functions, $\{\Psi_1, \Psi_2, \ldots, \Psi_n\}$. The average is

$$\bar{\Phi} = \frac{1}{n} \sum_{i=1}^{n} \Psi_i. \tag{10}$$

Therefore, the average removed SDF is $\bar{\Psi}_i = \Psi_i - \bar{\Phi}$, where $i = 1, 2, \ldots, n$. $\{\bar{\Psi}_1, \bar{\Psi}_2, \ldots, \bar{\Psi}_n\}$ is the corresponding set for SDFs (with zero mean) for each input shape. Now a lexicographic ordering matrix, S, is generated from $\{\bar{\Psi}_1, \bar{\Psi}_2, \ldots, \bar{\Psi}_n\}$, i.e., $S = [\tilde{O}_1 \tilde{O}_2 \ldots \tilde{O}_n]$. Then, the covariance of S has to be computed. k eigenvectors corresponding to k greatest eigenvalues of the covariance matrix are selected. Finally, the (reduced dimension) shape matrix $\tilde{S}$ will be $\tilde{S} = S \times A = [\tilde{O}'_1 \tilde{O}'_2 \ldots \tilde{O}'_n]$. A is a matrix containing k eigenvectors corresponding to the most dominant shapes in the training set. The lexicographic order matrix $\tilde{S}$ is transformed to its initial distance maps set form, i.e., $\{\Phi_1, \Phi_2, \ldots, \Phi_k\}$.

2.3. Level set evolution. The extracted features from the input image, explained in Section 2.1, and the distance maps, $\{\Phi_1, \Phi_2, \ldots, \Phi_k\}$, discussed in Section 2.2, are used in a variational framework to evolve the contour implicitly using the level set. The level set update equation is as follows (Tsai *et al*., 2003):

$$\begin{cases} \Phi_{[\mathbf{w},\mathbf{P}]}(x, y) = \bar{\Phi}(\tilde{x}, \tilde{y}) + \sum_{i=1}^{k} w_i \Phi_i, \\ \mathbf{w} = [w_1, w_2, \ldots, w_k], \\ \sum_{i=1}^{k} w_i = 1, \end{cases} \tag{11}$$

where $\tilde{x}$ and $\tilde{y}$, and $\bar{\Phi}$ are computed using Eqns. (6) and (10). Hence $\mathbf{w}$ is a vector used for weighing each signed distance function. Moreover, Φ_i is the principal shape computed using PCA. Considering Bayesian inference, the energy function is minimized during the contour evolution as (Rousson *et al.*, 2003)

$$E(\Omega_1, \Omega_2) = -\int_{\Omega_1} \log\ p_1(u'(x))\,\mathrm{d}x - \int_{\Omega_2} \log\ p_2(u'(x))\,\mathrm{d}x, \tag{12}$$

where $u(x)$ is the feature vector captured from nonlinear diffusion (see Eqn. (5)) as u'. p_1 and p_2 are Gaussian probability density functions (PDFs) with mean vector and covariance matrix $\{\mu_1, \Sigma_1\}$ and $\{\mu_2, \Sigma_2\}$, respectively (Rousson *et al.*, 2003),

$$p_i(u'(x)) = \frac{1}{(2\pi)^2 |\Sigma_i|^{1/2}} \cdot \exp\left[\frac{1}{2}\left(u'(x) - \mu_i\right)^{\mathrm{T}}\ \Sigma_i^{-1}\ \left(u'(x) - \mu_i\right)\right], \tag{13}$$

where μ_i and Σ_i are a $[1 \times (n+3)]$ vector and a $[(n+3) \times (n+3)]$ matrix, respectively. They are calculated using the following equations (Rousson *et al.*, 2003):

$$\begin{cases} \mu_i(\Phi) &= \dfrac{\displaystyle\int_\Omega u'(x)\aleph_i\,\mathrm{d}x}{\displaystyle\int_\Omega \aleph_i\,\mathrm{d}x}, \\ \Sigma_i(\Phi) &= \dfrac{\displaystyle\int_\Omega (\mu_i - u'(x))\,(\mu_i - u'(x))^{\mathrm{T}}\,\aleph_i\,\mathrm{d}x}{\displaystyle\int_\Omega \aleph_i\,\mathrm{d}x}, \end{cases}$$
$$\aleph_1(z) = H_\epsilon(z), \quad \aleph_2(z) = 1 - H_\epsilon(z). \tag{14}$$

$H_\epsilon(z)$ and Φ are the Heaviside and level set functions, respectively. The zero level for Φ corresponds to the contour (Rousson *et al.*, 2003):

$$\Phi : \Omega \to \mathbb{R}, \qquad \Phi(x) = \begin{cases} D(x, \partial\Omega), & x \in \Omega_1, \\ -D(x, \partial\Omega), & x \in \Omega_2. \end{cases} \tag{15}$$

$D(x, \partial\Omega)$ introduces the Euclidean distance between x and $\partial\Omega$. $\partial\Omega$ is the approximate boundary for the contour. Finally, Ω_1 and Ω_2 represent the inside and outside of the contour, respectively. The aim is to find the optimized values for $\mathbf{w}$ and $\mathbf{P}$ that can be used in Eqn. (11) to implicitly minimize Eqn. (12). This problem can be solved using optimization algorithms such as gradient descent, Levenberg–Marqurdt (LM) (Levenberg, 1944), genetic algorithms (GA) (Goldberg, 1989), simulated annealing (SA) (Kirkpatrick *et al.*, 1983), and pattern search (Torczon, 1997; Lewis *et al.*, 1999). These global optimization approaches (compared with gradient descent, which optimizes locally) are less sensitive to initialization, but they have a very low convergence speed. The update equation for w and P in the gradient descent method will be (Tsai *et al.*, 2003)

$$\begin{cases} \mathbf{w}^{t+1} = \mathbf{w}^t - \alpha_{\mathbf{w}} \nabla_{\mathbf{w}} E, \\ \mathbf{P}^{t+1} = \mathbf{P}^t - \alpha_{\mathbf{P}} \nabla_{\mathbf{P}} E, \end{cases} \tag{16}$$

where α_w and α_p are coefficients that determine the convergence rate of the gradient descent algorithm.

Also, the parameters using the LM algorithm can be estimated in three steps (Levenberg, 1944):

1. With an initial estimate $\mathbf{w}$, compute the gradient vector and Hessian matrix $\mathbf{H}$ as follows:

$$\nabla_{\mathbf{w}} E = \left[\frac{\partial E}{\partial w_1} \quad \frac{\partial E}{\partial w_2} \quad \cdots \quad \frac{\partial E}{\partial w_k}\right],$$

$$\mathbf{H}_{\mathbf{w}} E = \nabla_{\mathbf{w}}^{\mathrm{T}} E\ \nabla_{\mathbf{w}} E = \begin{bmatrix} \left(\dfrac{\partial E}{\partial w_1}\right)^2 & \dfrac{\partial E}{\partial w_1}\dfrac{\partial E}{\partial w_2} & \cdots & \dfrac{\partial E}{\partial w_1}\dfrac{\partial E}{\partial w_k} \\ \dfrac{\partial E}{\partial w_2}\dfrac{\partial E}{\partial w_1} & \left(\dfrac{\partial E}{\partial w_1}\right)^2 & \cdots & \dfrac{\partial E}{\partial w_2}\dfrac{\partial E}{\partial w_k} \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{\partial E}{\partial w_k}\dfrac{\partial E}{\partial w_1} & \dfrac{\partial E}{\partial w_k}\dfrac{\partial E}{\partial w_2} & \cdots & \left(\dfrac{\partial E}{\partial w_k}\right)^2 \end{bmatrix}. \tag{17}$$

2. Then, update the parameter $\mathbf{w}$ by

$$\mathbf{w}^{t+1} = \mathbf{w}^t + \left(\mathbf{H}_{\mathbf{w}} E + \lambda I\right)^{-1} \nabla_{\mathbf{w}}^{\mathrm{T}} E, \tag{18}$$

where λ is a time-varying stabilization parameter.

3. Go back to Step 1 until the estimate of $\mathbf{w}$ converges.

The same algorithm is employed to find the optimum values of P in minimizing the energy function E:

$$\mathbf{P}^{t+1} = \mathbf{P}^t + \left(\mathbf{H}_{\mathbf{P}} E + \lambda I\right)^{-1} \nabla_{\mathbf{P}}^{\mathrm{T}} E \tag{19}$$

The LM algorithm is a standard nonlinear optimization approach which significantly surpasses the gradient descent method for medium sized problems (Roweis, 2010; Levenberg, 1944). Employing the knowledge about the curvature as well as the gradient of the error surface speeds up the minimization procedure. In fact, the LM algorithm constitutes a trade-off between the Gauss–Newton and the gradient descent. Using a small value of λ forces the LM algorithm towards the Gauss–Newton one.

Conversely, the LM algorithm with a large value of λ resembles the gradient descent. Generally, the LM algorithm is more suitable for multivariate function minimization compared with the classic gradient descent. However, it is of greater computational complexity compared with the gradient descent due to inversion of the Hessian. Specially, by increasing the number of variables, the complexity of computations grows drastically.

The update equations in the LM algorithm (Eqns. (18) and (19)) adjust λ according to the changes in the error function (Ranganathan, 2004). If the error increases as a result of the update (meaning that the quadratic approximation using the Hessian is not functioning well and the working point is not likely near a minimum), then restore the weights to their previous values and increase λ. This shows that the LM algorithm tends to a simple gradient descent. Conversely, if the error decreases (indicating that the quadratic approximation performs properly and the working point gets closer to a minimum), then preserve the current weights and decrease λ. This will thus result in allocating more weight on the Hessian.

The variational framework for updating the level set function (see Eqn. (11)), calculating the image statistics (Eqn. (14)), computing the energy function (Eqn. (12)) and updating the weights and pose parameters (Eqn. (16) or Eqns. (18) and (19)) is illustrated in Fig. 2.

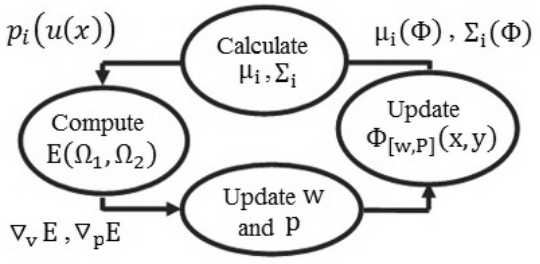

Fig. 2. Variational framework used for level set evolution.

3. Experimental results

In this section, the results of the proposed method are presented. The algorithm has been tested on three categories of images. The first one consists of outdoor images in which the object of interest has a specific shape. The second one includes cardiac magnetic resonance images (MRI) where tissue segmentation is performed. Finally, another application of our algorithm is investigated for OCR.

3.1. Outdoor image segmentation. There are problems associated with segmenting outdoor images, such as

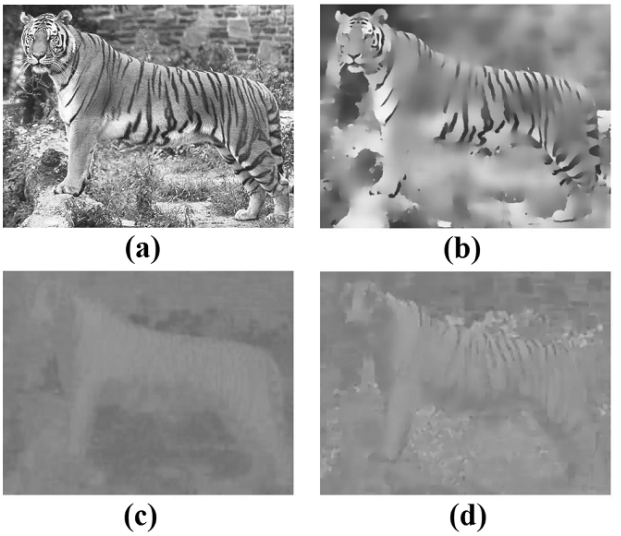

Fig. 3. Input image (a); results of nonlinear diffusion for L* (b), a* (c) and b* (d) color channels.

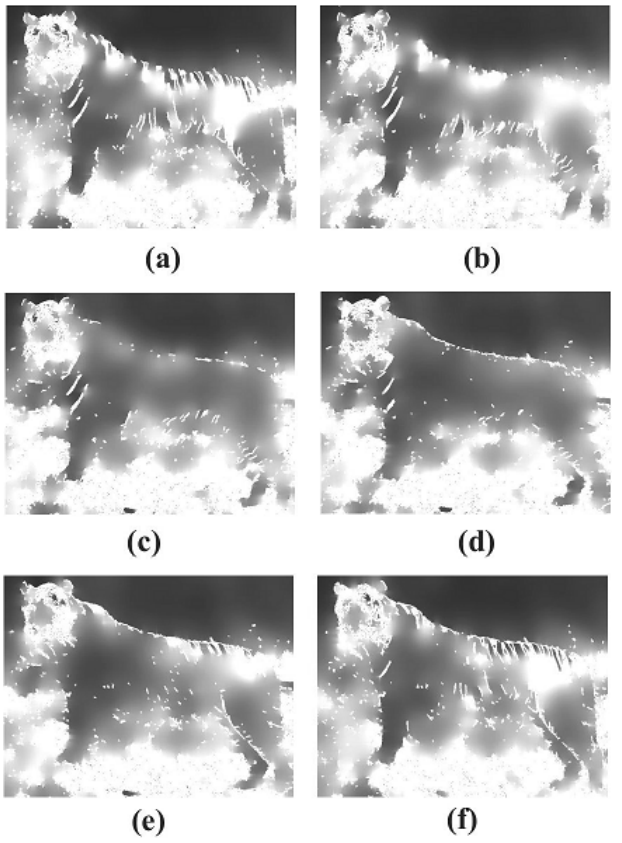

Fig. 4. Results for directional derivatives with nonlinear diffusion. The derivatives are computed for 0 (a), 30 (b), 60 (c), 90 (d), 120 (e), and 150 (f) degrees.

the shadow effect, an inhomogeneous texture, occlusion and camera imaging noise. Using color or texture features is not sufficient for a suitable segmentation. A prior knowledge of the object of interest should also be utilized in segmentation to overcome these problems. Figure 3(a) illustrates an outdoor image of a tiger with a specific texture in the foreground and a complicated background. Using only texture features to segment the tiger from the background results in failure due to a high foreground texture and background complexity. However, incorporating the tiger shape as prior knowledge for level set evolution as well as color and texture features produces desirable results.

Figures 3(b)–(d) illustrate the results of nonlinear diffusion for L*a*b* color components, respectively. Figure 4 demonstrates the results of directional derivatives. Directional derivatives are computed from Eqn. (1) with $\theta_0 = 0$ and $\Delta\theta = \pi/6$. The results are used as $u(x)$ in Eqn. (13) to compute the energy function in Eqn. (12). Due to background complexity, it can be simply verified from Fig. 4 that only using nonlinear diffusion outputs as inputs for a two-phase level set leads to over-segmentation. A two-phase level set is an active contour approach that considers only the background and the foreground (two classes) for the final segmentation. Therefore, the need for incorporating the prior shape knowledge of the object of interest seems to be undeniable.

To solve this problem, 16 side view binary images of a tiger with different poses are used as training shapes. They are aligned according to the procedure presented in Section 2.2. As an example, one of the training images before and after alignment is plotted in Fig. 5. In the alignment procedure, one of the images is considered the reference (Fig. 5(c)) and the others are aligned according to the reference. As an example, seven binary images are plotted in Fig. 6(a). The corresponding aligned shapes are visualized in Fig. 6(b). After alignment, PCA is utilized to capture the prominent shapes. We choose $k = 10$ in Eqn. (11), i.e., ten principal components are acquired from training shapes. Finally, SDFs are set up according to the principal shapes.

The distance functions are utilized as in Eqn. (11) to evolve the contour implicitly. $\bar{\Phi}$ (Eqn. (10)) is plotted in Fig. 7. The contour depicts the zero level set, i.e., $\bar{\Phi} \approx 0$. After computing SDFs from the training shapes, the contour is updated according to Eqn. (11) implicitly in a direction that minimizes the energy function (Eqn. (12)). The result of segmentation is visualized in Fig. 8(b). Figure 8(c) provides evidence for a comparison between our algorithm (a combination of bottom-up and top-down procedures) and those using texture features (a bottom-up or blind segmentation procedure). Due to object and background complexity, texture features cannot lead to an acceptable segmentation and over-segmentation will occur. Also, it is worth of noting that for updating the level set function in Fig. 8(c), where the prior shape has not been considered, the following level set equation is employed (Rousson *et al.*, 2003):

$$\Phi_t(x) = \delta_\epsilon\left(\Phi(x)\right)\left(v\ \mathrm{div}\left(\frac{\nabla\Phi}{|\nabla\Phi|}\right) + \log\frac{p_1(u'(x))}{p_2(u'(x))}\right), \tag{20}$$

where $\delta_\epsilon(z)$ is a regularized Dirac impulse function, and $p_1(u'(x))$ and $p_2(u'(x))$ are computed according to Eqn. (13). Here v determines the rate of active contour evolution.

Our region-based method also exhibits robustness in noisy image segmentation. Figures 9 (a)–(c) prove it. The image is degraded by additive Gaussian noise. The peak signal to noise ratio (PSNR) is 27.59 ($\mathbf{dB}$). Moreover, in the case of occlusion, our algorithm demonstrates desirable results. Figure 9(d) shows the same image occluded by black horizontal lines. The result for segmentation is plotted in Fig. 9(f). It is worth mentioning that in the initialization step (plotted in Figs. 9(b) 9(e)), two distance maps from the training shapes are used to speed up the segmentation. The initialization shown in Fig. 8(b) leads to the same result. However, it takes much more time for the contour to converge.

When computing the covariance matrix in Eqn. (14), it is deducted that if we diagonalize (whiten) Σ_i, the result will be better. In other words, the correlation reduction in the feature vectors (Eqn. (5)) produces a more reasonable result and fewer local minima will occur in the energy function (Eqn. (12)).

Another important issue that must be taken into account is the computational cost of the algorithm and the optimization method. We have compared five different optimization approaches in our work. They are the gradient descent algorithm (refer to Section 2.3), the Levenberg–Marquardt algorithm (Levenberg, 1944), genetic algorithms (Goldberg, 1989), simulated annealing (Kirkpatrick *et al.*, 1983) and pattern search (Torczon, 1997; Lewis *et al.*, 1999).

Among these algorithms, the gradient descent is the fastest one. However, it is highly sensitive to initialization and local minima in the energy function. If the initialization is not near the final point (i.e., a global minimum), it will never converge. This problem also exists for the LM algorithm. It is slower than the gradient descent because of its higher order partial derivative calculations (Levenberg, 1944). On the other hand, compared with the gradient descent algorithm, the convergence speed is much lower for pattern search and SA. Conversely, they are much less dependent on level set initialization and the initial contour can be located freely on the test image. Finally, despite the GA's ability to quickly scan a vast solution set, its major problem is that it might never stop evolution at the optimum point and the evolution may halt at local minima (or may not stop at all). Table 1 shows the

Table 1. Elapsed time for optimization algorithms.

Gradient Descent	LM	GA	SA	Pattern Search
12 (sec)	36 (sec)	960 (sec)	1210 (sec)	1405 (sec)

elapsed time for the optimization algorithms to be converged for segmenting Figs. 9(d)–(f). The algorithm was implemented on an Intel Core 2 Due CPU (T7250) with 2024 MB RAM.

As another example, we chose the tiger's image occluded by horizontal and vertical black lines as depicted in Fig. 10(a). It is completely obvious that the continuity of texture and color features was lost due to the occlusion. On the other hand, using the prior shape knowledge of the tiger can help to overcome this problem. The diffusion's result is an $n + 3$ feature vector. We have only demonstrated the result of the $L*$ channel. Figure 10(c) visualizes the result for the proposed method after solving the nonlinear diffusion in the image domain. Moreover, the same image is degraded by additive Gaussian noise with the PSNR of 5.34 $(\mathbf{dB})$ (Fig. 10(h)). It is clear that some edges have been completely blurred due to the noise leading to a hard distinction between background and foreground. However, employing the prior shape knowledge of the object of interest with intrinsic features simultaneously can resolve the problem.

The results for diffusion and segmentation are plotted in Figs. 10(i) and (j), respectively. To show the weakness of using bottom-up (Cremers *et al.*, 2007; Rousson *et al.*, 2003) and top-down (Tsai *et al.*, 2003) approaches solely for segmenting this image, the third and fourth columns in Fig. 10 are dedicated to these segmentation algorithms, respectively. It can be confirmed from Figs. 10(d) and (k) that only texture features and color components are not sufficient to segment such an inhomogeneous and complicated texture. Over-segmentation is inevitable and in some regions the occlusion is considered a foreground (Fig. 10(g)). Furthermore, using only shape prior knowledge (Tsai *et al.*, 2003) cannot provide the best result as can be verified from Figs. 10(e) and (l).

Morphological open-closing is also investigated as an alternative (of nonlinear diffusion) for feature extraction For this purpose, a rectangle structuring element ($1/18$ of the size of the input image) is utilized (Jain, 1989; Emambakhsh and Sedaaghi, 2009). The result is shown in Figs. 10(f) and (m). Similarly to nonlinear diffusion, texture uniformity is obtained and the occlusion has almost disappeared in the output image. But the problem in morphological filtering is its "blocking artifact" resulting from masking the structuring element over the input image. The masking generates unwanted rectangles in Figs. 10(f) and (m). Also, some parts of the foreground are merged into the background. This problem results in over-segmentation as shown in Figs. 10(g) and (n). On the other hand, nonlinear diffusion preserves image edges more precisely while the noise effect simultaneously decreases. Generally, choosing a suitable structuring element is very important when morphological operators are used for feature extraction from the texture. The type of structuring element highly depends on the texture. It will be even much more complicated for inhomogeneous textures, such as in Fig. 10(a).

To evaluate the performance of our algorithm numerically, the percentage of correct segmentation (PCS) is computed. It is performed by comparing the ground truth with the output of the proposed method. The noise power is increased (which reasonably results in a PSNR decrease) and the PCS is calculated for segmenting the tiger image shown in Fig. 10. The variation of the PCS vs. the PSNR is plotted in Fig. 11 for our algorithm when feature extraction is performed using three different methods: (i) nonlinear diffusion (as explained in our algorithm in Section 2.1.3), (ii) by computing morphological open-closing with a structuring element whose size is equal to one eighteenth of the input image's size (Jain, 1989; Emambakhsh and Sedaaghi, 2009), and (iii) by Gabor feature extraction (Andrysiak and Choras, 2005; Raja *et al.*, 2010; Sandler and Lindenbaum, 2006). Moreover, the results for the bottom-up algorithm (Cremers *et al.*, 2007; Rousson *et al.*, 2003) and the top-down approach (Tsai *et al.*, 2003) are plotted. As is clear from this figure, our approach (integrating intrinsic image features with prior shape knowledge) achieved the best PCS. Furthermore, compared with Gabor feature extraction and morphological open-closing, nonlinear diffusion has produced a higher PCS. Morphological open-closing preserves image edges more precisely compared with Gabor feature extraction. Conversely, Gabor filtering reduces noise more accurately. On the other hand, according to this figure, top-down methods show more robustness for noisy image segmentation compared with the bottom-up approach, but still generate a smaller PCS compared with our hybrid algorithm.

We also evaluated our algorithm on a Caltech image dataset (The-ViewCVS-Group, 2010) consisting of various images of objects and animals. Since different images from a single object were presented in this dataset, it is very useful to generate the training shape images and consequently segment the test images according to our algorithm. In our work, we selected rooster and stegosaurus images. Prominent shapes are captured from the dataset and SDFs are allocated to the resulting images.

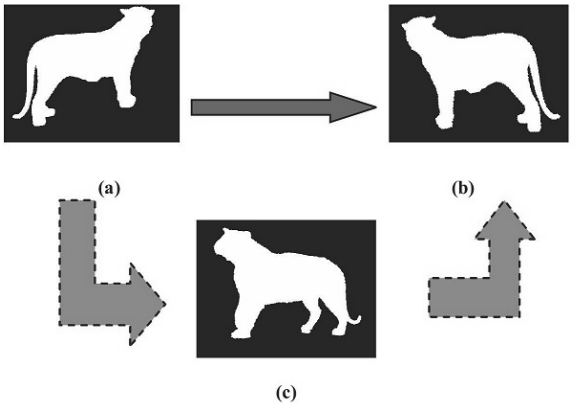

Fig. 5. Alignment procedure: before alignment (a), after alignment (b), reference shape (c).

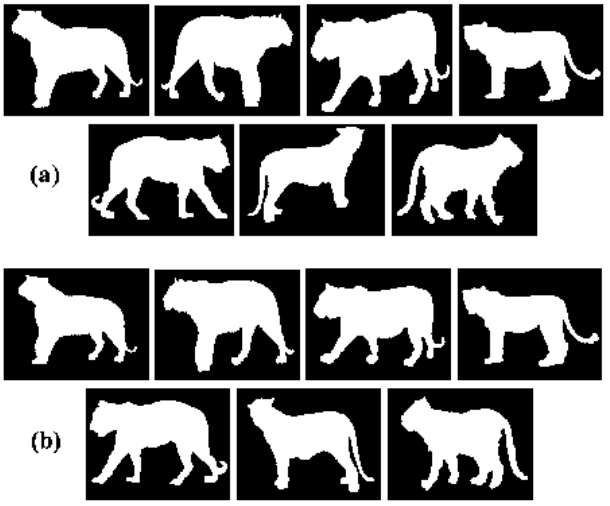

Fig. 6. Example of binary images (7 out of 16) used to generate the prior shapes: unaligned (a) and aligned binary images (b).

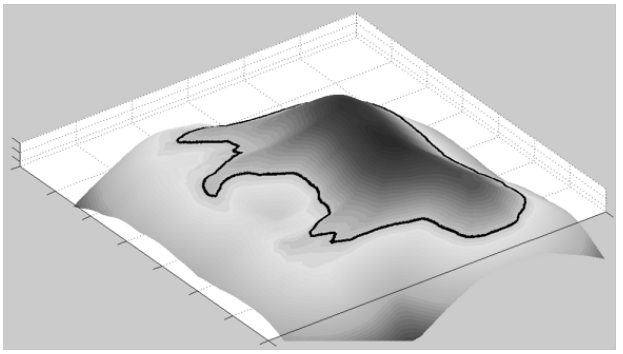

Fig. 7. Average of SDFs, $\bar{\Phi}$, computed using Eqn. (10).

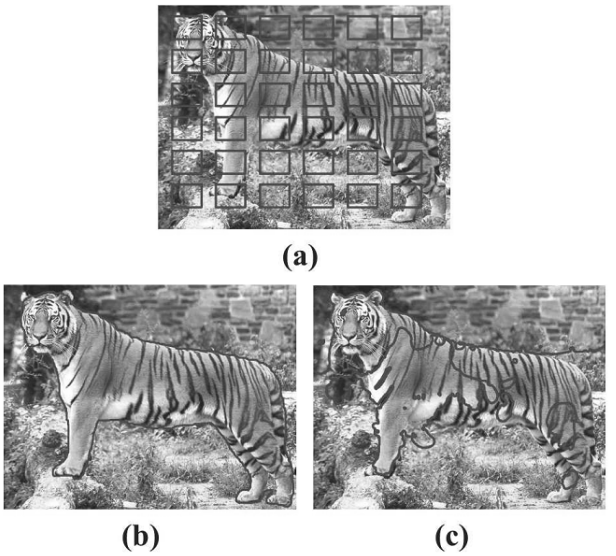

Fig. 8. Initialization for the level set function (a), results for our segmentation algorithm (b), algorithm without considering prior shape (c).

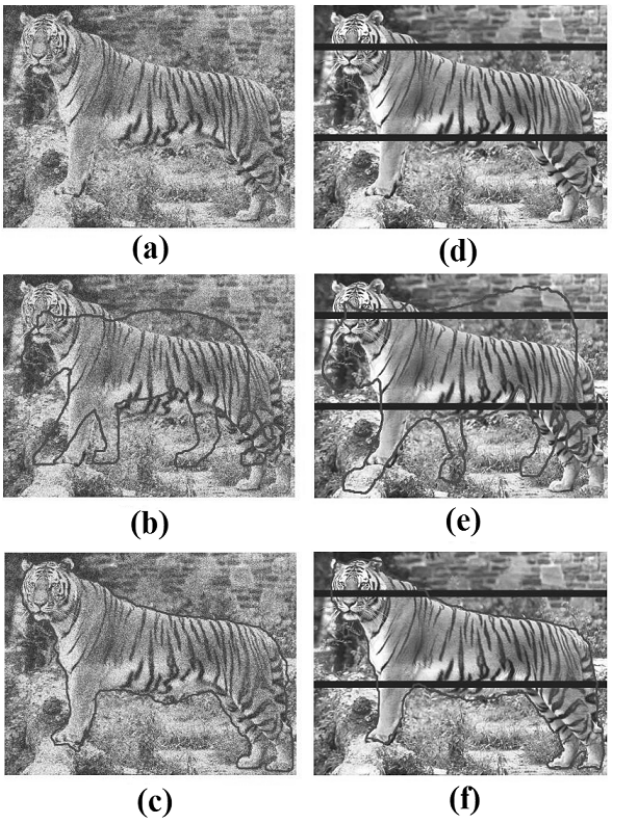

Fig. 9. Results of segmentation for the proposed method in two different cases. The left column indicates the case when the input image is corrupted by noise: noisy image (a), initialization (b), segmented image (c). The right column shows the results when the image is occluded: occluded image (d), initialization (e), result of the segmentation (f).

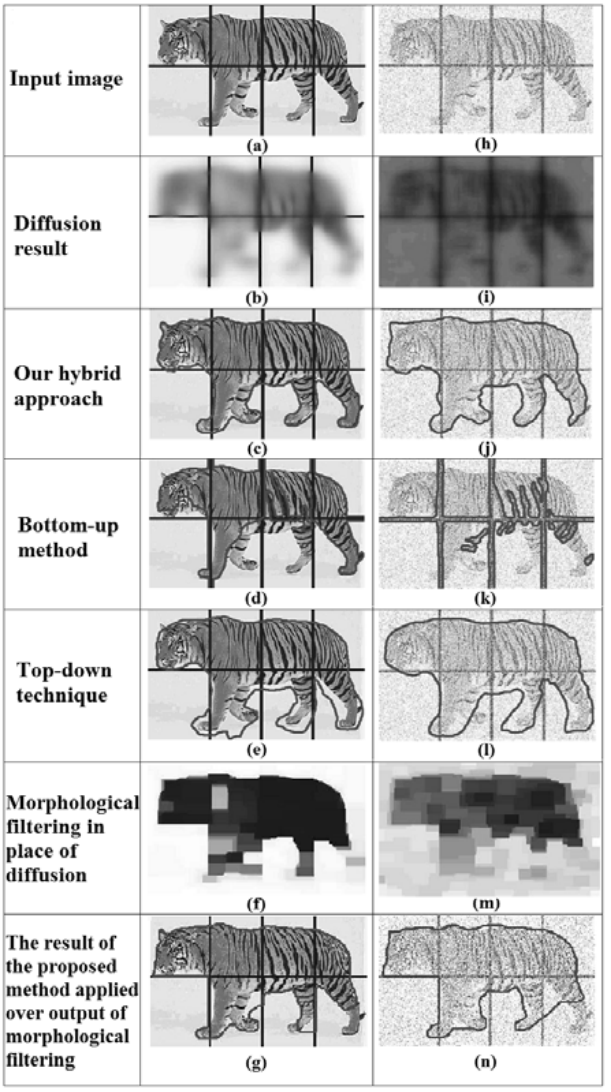

Fig. 10. Comparison of different methods for segmentation. The images in the left column of the figure correspond to the case when the input image is occluded. The images in the right column correspond to noisy image: The occluded and noisy input images (a) and (h), the results for diffusion (b) and (i), the result of our hybrid method where intrinsic image features and prior shape knowledge are simultaneously employed (c) and (j), bottom-up approach (d) and (k), top-down method (e) and (l), the results for open-closing, (f) and (m), the hybrid segmentation's result where morphological open-closing is used in place of diffusion (g) and (n).

Figures 12(a) and (c) show nine sample images used as training shapes as explained in Section 2.2. Also, similar to Fig. 7, $\bar{\Phi}$ (Eqn. (10)) is computed and plotted in Figs. 12(b) and (d), where the black enclosed contour represents the zero level.

Figure 13 is dedicated to four test images from this dataset. Figures 13(a)–(d) are the input images. In order

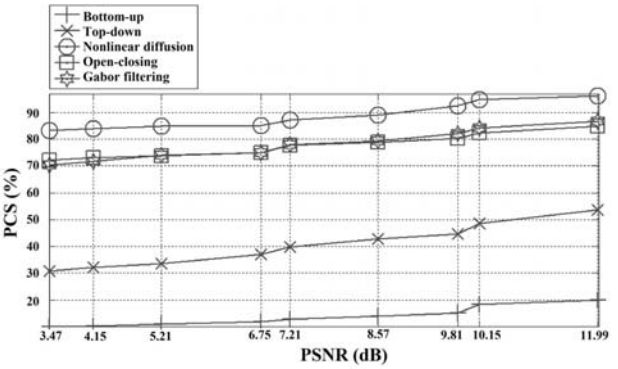

Fig. 11. PSNR vs. the PCS for the tiger image in Fig. 10.

to demonstrate the robustness of our algorithm against occlusion, rotation and noise, the images are degraded as in Figs. 13(e)–(g). Figure 13(d), containing a complicated texture with big variations in its pixel's color, is utilized without degradation. Figures 13(h)–(k) show contour initialization. Finally, the result for segmentation using the proposed method is shown in the last column.

Also, we have claimed that the CIE L*a*b* color space is much more suitable for setting up the feature space than RGB. An RGB color system demonstrates low detachability of the feature space. Also, there are a few local maxima in the histogram of RGB color components. This causes more complications in distinguishing the background from the foreground in an image. Consequently, segmentation based on the RGB color space produces a less precise result, as depicted in Fig. 14.

There are many different papers that analyzed the effects of the color space for segmentation purposes, (Skarbek and Koschan, 1994; Cheng *et al*., 2001). However, as completely discussed by Cheng *et al*. (2001), each color space has its own advantages and disadvantages and, still, there has not been a single color representation that can surpass others for segmenting all sorts of color images (Cheng *et al*., 2001). It is understood that for outdoor image segmentation, where real colors are taken into account, CIE L*a*b* (nonlinear color spaces) generates more acceptable results compared with RGB (linear color spaces). However, for artificial (synthetic) image segmentation, where colors are not necessarily natural, the RGB color representation may sometimes be preferred.

3.2. Medical image segmentation. Medical image analysis and segmentation is another field of study that our algorithm can be applied in. There are many situations such as patient unwanted movements during imaging, imaging noise, low contrast and occlusion resulted by other body tissues. They all lead to an inaccurate segmentation. Because of these problems, employing only intensity is not sufficient for a reasonable segmentation.

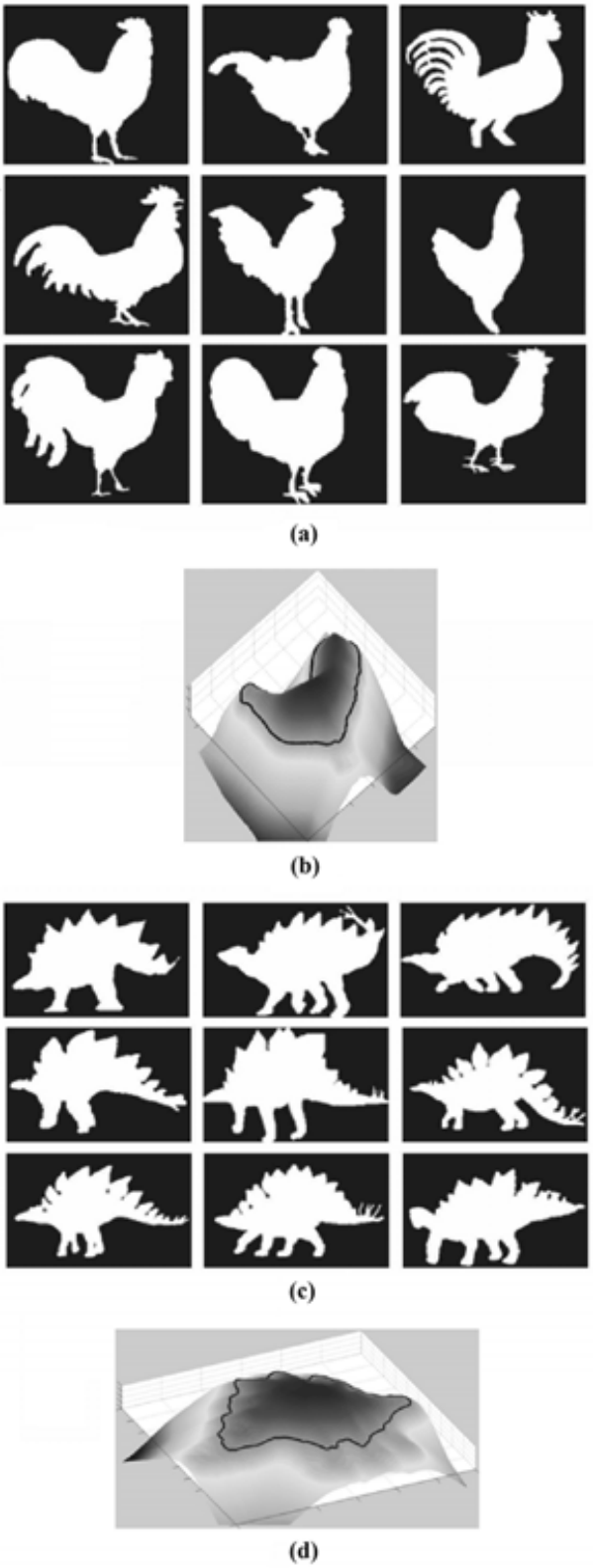

Fig. 12. Training shape images captured from rooster (a) and stegosaurus datasets (b); (c) and (d) represent the average of SDFs computed according to Eqn. (10).

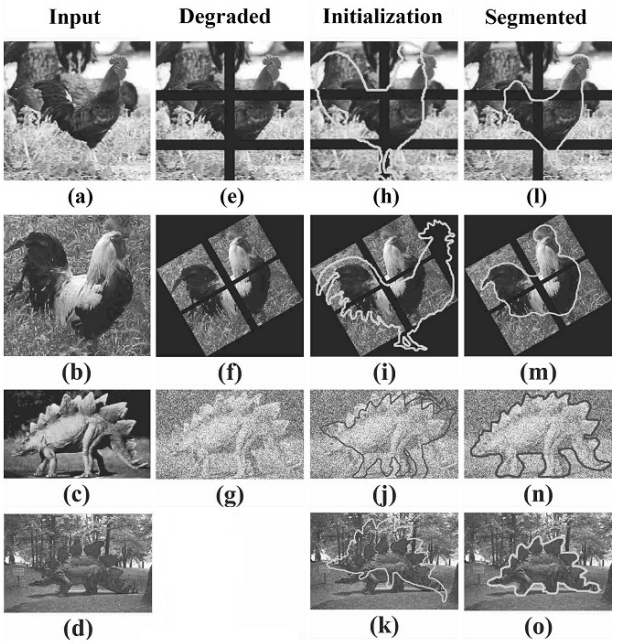

Fig. 13. Evaluation of the proposed method: input images (a)–(d), occluded, rotated and occluded, and noisy images, respectively (e)–(g), level set initialization (h)–(k), segmentation's result (l)–(o).

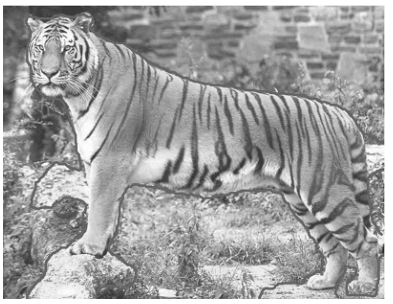

Fig. 14. Segmentation based on RGB color components.

Fig. 15. Ten samples of binary images used to generate the prior shapes.

As a result, using the tissue's prior shape knowledge can be beneficial to resolve the problem. In our work, a cardiac MRI sequence from York University (Andreopoulos and Tsotsos, 2010) is evaluated. To segment a predefined

Fig. 16. Result of segmentation for MRI sequences 120 to 135.

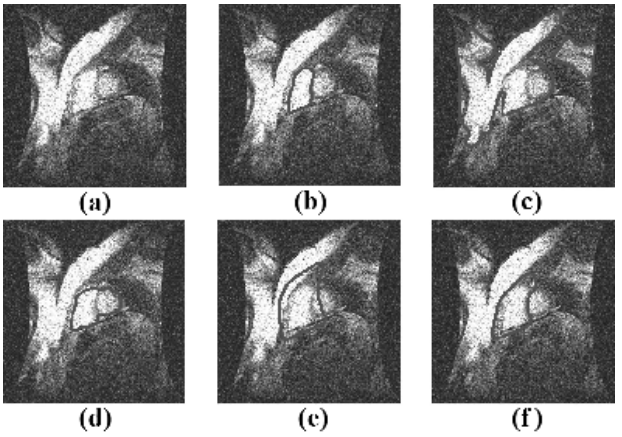

Fig. 17. Results of segmentation for noisy cardiac MRI: noisy image (a), segmentation using: proposed method (b), bottom-up algorithm (c), top-down technique (d), morphological filtering (e), Gabor-based feature spaces (f).

tissue in cardiac MRI, first, some training images containing the tissue's different shapes are considered. For this purpose, distance maps are generated using different MRI sequences. Figure 15 shows ten binary images generated from the MRI cardiac sequence (Andreopoulos and Tsotsos, 2010). Having used the binary images to make the distance maps in Eqn. (11), the segmentation for sequences 120–135 is performed and depicted in Fig. 16. Since the images are gray scale, the input of the nonlinear diffusion equation in (3) is n-dimensional (unlike color images having dimension $(n+3)$.

Our algorithm is more robust in the presence of noise compared with the top-down algorithm (Tsai *et al*., 2003) and the bottom-up approach (Cremers *et al*., 2007; Rousson *et al*., 2003) because of using both prior shape knowledge and intrinsic image features. Moreover, our feature space, which is based on non-linear diffusion, generates a more acceptable result compared Gabor and morphological-based feature spaces. To demonstrate it, Fig. 17(a) shows another MRI degraded by speckle and additive Gaussian noise. The PSNR is 25.81 (dB). The result for the proposed method is depicted in Fig. 17(b). The results for bottom-up and top-down approaches are plotted in Figs. 17(c) and (d). As is clear, bottom-up methods cannot localize the object and, due to noise, contour evolution in the top-down approach is trapped in a local minimum. The results of morphological and Gabor-based feature spaces are shown in Figs. 17(e) and (f), respectively. Because of the effects of blocking artifact in morphological filtering and edge smoothing in Gabor, the cardiac tissue is not determined precisely.

3.3. Special letter extraction. In Farsi (Persian) and Arabic character recognition, there are some letters known as special ones. These letters can be written in various forms, depending on their location in a word, and this makes their recognition very hard to attain. Pre-locating (or detecting) these letters will ease the character recognition procedure. However, it is a complicated task. In previous works (Biasdy *et al*., 2006; Einsele *et al*., 2008), morphological operators were used extensively to locate these letters.

Morphological matching is highly sensitive to scaling. To overcome the scaling obstacle, different scales of the target letter are used as the structuring element (Biasdy *et al*., 2006). However, this approach produces highly redundant computations. Moreover, morphological matching is deeply vulnerable to the rotation of the test image. A slight paper rotation during scanning leads to a rotation in characters and, consequently, the structuring element must be rotated in various degrees to find the correct matching.

Our approach is considered a proper candidate for such sophisticated application. For this purpose, first, some binary images of the special letter are provided to incorporate prior shape knowledge. These images are created from the special letters with various font styles. For example, in Fig. 18, one of these letters typed in various fonts is plotted. Although the dictation is different between the images, all of them represent a single letter. These training binary images are used as prior shape knowledge for segmentation. The result for special letter detection is shown in Fig. 19. The test images are plotted in Figs. 19(a)–(d). The fonts in test images are different

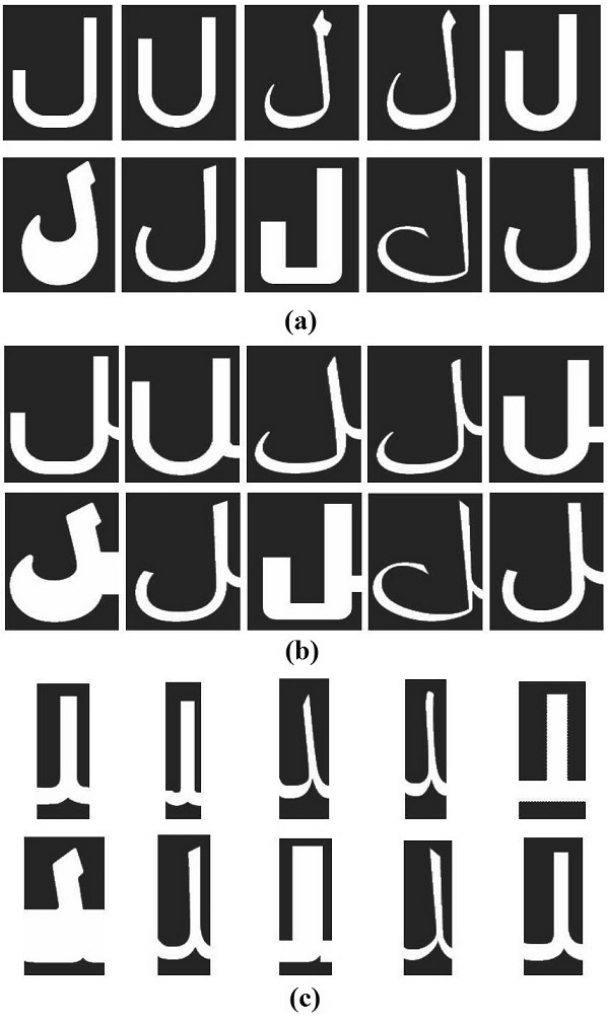

Fig. 18. Separated form of the letter (a), first and second attached form of the special letter, respectively (b), (c).

from those in training shapes.

Our algorithm was also compared with correlation-based matching. The results for the proposed method and the correlation-based algorithm are visualized in Figs. 19(e)–(h) and 19(i)–(p), respectively. In the correlation method, the correlation map is first computed according to the normalized cross correlation equation (Jain, 1989) (The arrows in Figs. 19(i)–(l) indicate the locations with maximum correlation.) Then, ellipses are plotted on the highest correlated points (Figs. 19(m)–(p)). The semi-minor and semi-major axes are allocated according to the template size. It is clear from the result that the exact boundary of the letter cannot be detected by correlation-based algorithms. Moreover, for rotated images, exact matching cannot be provided by these algorithms as has been depicted in Fig. 19(p). Finally, automatic detection of the threshold in correlation maps for locating matching points is a highly complicated task.

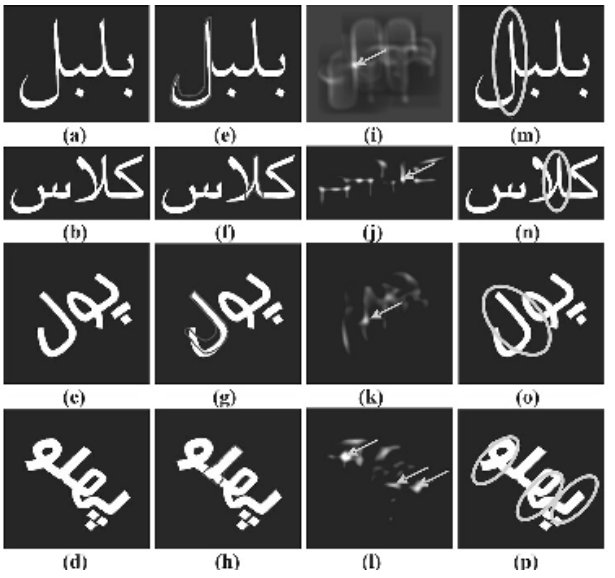

Fig. 19. Results for special letter detection: test images (a)–(d), the proposed method (e)–(h), the correlation maps (i)–(l), the correlation-based matching detection (m)–(p).

4. Conclusion

The main contribution of this research is the integration of color components and texture features with shape prior knowledge for region-based image segmentation. Our method consists of building a feature space based on nonlinear diffusion of CIE L*a*b* color components and directional derivatives, making an energy function, distance map generation, and level set evolution. It is applicable to outdoor and medical images and special letter (character) segmentation. Compared with previous top-down and bottom-up algorithms (Cremers *et al.*, 2007; Rousson *et al.*, 2003; Tsai *et al.*, 2003), the proposed method is much more robust for noisy and occluded images. This is evaluated by computing the PCS. Calculating the PCS for different noise powers show that our algorithm is superb for segmenting noisy images with a very low PSNR (even less than $3.5\ dB$). The advantages of our algorithm are as follows:

(i) Using shape prior knowledge and nonlinear diffusion makes our region-based algorithm robust for noisy images.

(ii) Compared with Gabor filters and structure tensors, the proposed feature space has lower dimensionality. Moreover, using nonlinear diffusion preserves image edges to a greater degree.

(iii) Incorporating prior shape knowledge improves the robustness of the proposed segmentation against occlusion. Moreover, prior knowledge can be used to overcome shadow effects in outdoor images.

(iv) Computing directional derivatives is very important for segmenting complicated and inhomogeneous textures more successfully.

(v) Compared with the RGB color space (and, generally, its linear transforms), employing nonlinear color spaces, especially CIE L*a*b* color components, produces better and more reasonable results since it generates a more detachable and uncorrelated feature space.

(vi) We have used the gradient descent algorithm, which locally minimizes the energy function and has a low computational complexity. However, genetic and pattern search algorithms (Torczon, 1997; Lewis *et al.*, 1999), which are well-known global optimization methods, can be used to minimize the energy function. Unlike gradient descent methods, they are less sensitive to initialization but have a high computational complexity and a low convergence speed.

Our work can also be employed for other applications such as object detection and recognition. In fact, having an appropriate segmentation eases feature extraction in object recognition and highly improves the result. Moreover, the proposed algorithm can be utilized in motion segmentation. The feature space may be enlarged to a combination of color, texture and motion vectors, captured from consecutive video frames where prior shape knowledge of the object of interest is also incorporated.

Finally, as indicated in the proposed method, the distribution of the feature space is estimated to follow a multivariate Gaussian model. In other words, a parametric model is employed and feature vectors are fit by a multivariate Gaussian function. However, there are some applications where other distributions might be suitable. Therefore, a more general approach is to estimate the distribution of the feature space with a non-parametric model. Of course, this will increase the computational complexity, but on the other hand, it can raise the generalization of segmentation.

References

Andreopoulos, A. and Tsotsos, J.K. (2010). Cardiac MRI dataset, `www.cse.yorku.ca/ mridataset/`.

Andrysiak, T. and Choras, M. (2005). Image retrieval based on hierarchical Gabor filters, *International Journal of Applied Mathematics and Computer Science* **15**(4): 471–480.

Biasdy, F., El-Sana, J. and Habash, N. (2006). Online arabic handwriting recognition using hidden Markov models, *Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition, La Baule, France*.

Cheng, H., Jing, X.H., Sun, Y. and Wang, J. (2001). Color image segmentation: Advances and prospects, *Pattern Recognition* **34**(12): 2259–2281.

Cremers, S.D., Rousson, M. and Deriche, R. (2007). A review of statistical approaches to level sets segmentation: Integrating colour, texture, motion and shape, *International Journal of Computer Vision* **72**(2): 195–215.

Dokur, Z., Iscan, Z. and Olmez, T. (2006). Segmentation of medical images by using wavelet transform and incremental self-organizing map, *in* A. Gelbukh and C.A. Reyes-Garcia (Eds.) *MICAI 2006: Advances ina Artificial Intelligence*, Lecture Notes in Computer Science, Vol. 4293, Springer, Berlin/Heidelberg, pp. 800–809.

Einsele, F., Ingold, R. and Hennebert, J. (2008). A language-independent, open-vocabulary system based on hmms for recognition of ultra low resolution words, *Journal of Universal Computer Science* **14**(18): 2982–2997.

Emambakhsh, M., Ebrahimnezhad, H. and Sedaaghi, M.H. (2010). A hybrid top-down/bottom up approach for image segmentation incorporating color and texture with prior shape knowledge, *18th Iranian Conference on Electrical Engineering ICEE2010, Isfahan, Iran*, pp. 270–275.

Emambakhsh, M. and Sedaaghi, M.H. (2009). Automatic MRI brain segmentation using local features, self-organizing maps, and watershed, *IEEE International Conference on Signal and Image Processing Applications, Kuala Lumpur, Malaysia*, pp. 123–128.

Feddern, C., Weickert, J. and Burgeth, B. (2006). Level-set methods for tensor-valued images, *Proceedings of the IEEE 2nd Workshop on VLSM, Nice, France*, pp. 65–72.

Forsyth, D.A. and Ponce, J. (2002). *Computer Vision: A Modern Approach*, Prentice-Hall of India, New Dehli.

Gerig, G., Kbler, O., Kikinis, R. and Jolesz, F.A. (1992). Nonlinear anisotropic filtering of MRI data, *IEEE Transactions on Medical Imaging* **11**(2): 221–232.

Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA.

Greig, D., Porteous, B. and Seheult, A. (1989). Exact maximum a posteriori estimation for binary images, *Journal of the Royal Statistical Society. Series B (Methodological)* **51**(2): 271–279.

Hao, J., Shen, Y. and Wang, Q. (2007). Segmentation for MRA image: An improved level-set approach, *IEEE Transactions on Instrumentation and Measurement* **56**(4): 1316–1321.

Hrebien, M., Stec, P., Nieczkowski, T. and Obuchowicz, A. (2008). Segmentation of breast cancer fine needle biopsy cytological images, *International Journal of Applied Mathematics and Computer Science* **18**(2): 159–170, DOI: 10.2478/v10006-008-0015-x.

Jain, A.K. (1989). *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ.

Kass, M., Witkin, A. and Terzopoulos, D. (1988). Snakes, active contour model, *International Journal of Computer Vision* **1**(4): 321–331.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983). Optimization by simulated annealing, *Science Magazine* **220**(4598): 671–680.

Kuo, W., Lin, C. and Sun, Y. (2008). Brain MR images segmentation using statistical ratio: Mapping between watershed and competitive Hopfield clustering network algorithms, *Computer Methods and Programs in Biomedicine* **91**(3): 191–198.

Lai, C. and Chang, C. (2009). A hierarchical evolutionary algorithm for automatic medical image segmentation, *Expert Systems with Applications* **36**(1): 248–259.

Levenberg, K. (1944). A method for the solution of certain problems in least-squares, *Quarterly Applied Mathematics* **2**(2): 164–168.

Lewis, Michael, R. and Torczon, V. (1999). Pattern search algorithms for bound constrained minimization, *SIAM Journal on Optimization* **9**(4): 1082–1099.

Lu, Y., Wang, J., Kong, J., Zhang, B. and Zhang, J. (2006). An integrated algorithm for MRI brain images segmentation, *Computer Vision Approaches to Medical Image Analysis* **4241**: 132–142.

Mitiche, A. and Sekkati, H. (2006). Optical flow 3d segmentation and interpretation: A variational method with active curve evolution and level sets, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(11): 1818–1829.

Ong, S.H., Yeo, N.C., Lee, K.H., Venkatesh, Y.V. and Kao, D.H. (2002). Segmentation of color images using a two-stage self-organizing network, *Image and Vision Computing* **20**(4): 279–289.

Osher, S. and Sethian, J. (1988). Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations, *Journal of Computationl Physics* **79**(1): 12–49.

Perona, P. and Malik, J. (1990). Scale space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(7): 629–639.

Petera, Z., Boussone, V., Bergote, C. and Peyrina, F. (2008). Aconstrained region growing approach based on watershed for the segmentation of low contrast structures in bone micro-CT images, *Pattern Recognition* **41**(7): 2358–2368.

Raja, K. B., Madheswaran, M. and Thyagarajah, K. (2010). Texture pattern analysis of kidney tissues for disorder identification and classification using dominant Gabor wavelet, *Machine Vision and Applications* **21**(3): 287–300.

Ramme, A.J., DeVries, N., Kallemyn, N.A., Magnotta, V.A. and Grosland, N. M. (2009). Semi-automated phalanx bone segmentation using the expectation maximization algorithm, *Journal of Digital Imaging* **22**(5): 483–491.

Ranganathan, A. (2004). The Levenberg-Marquardt algorithm, *Technical report* `http://www.scribd.com/doc/10093320/Levenberg-Marquardt-Algorithm`.

Rousson, M., Brox, T. and Deriche, R. (2003). Active unsupervised texture segmentation on a diffusion based feature space, *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, WI, USA*, pp. 699–704.

Roweis, S. (n.d.). Levenberg–Marquardt optimization, `http://www.cs.nyu.edu/roweis/notes/lm.pdf`.

Sandler, R. and Lindenbaum, M. (2006). Gabor filter analysis for texture segmentation, *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06), New York, NY, USA*, p. 178.

Skarbek, W. and Koschan, A. (1994). Color image segmentation—A survey, *Technical report*, University of Berlin, Berlin.

Sagiv, C., Sochen, N.A., and Zeevi, Y.Y. (2006). Integrated active contours for texture segmentation, *IEEE Transactions on Image Processing* **16**(6): 1633–1646.

Susomboon, R., Raicu, D. and Furst, J. (2006). Automatic single-organ segmentation in computed tomography images, *Sixth IEEE International Conference on Data Mining (ICDM'06), Hong Kong, China*, pp. 1081–1086.

The-ViewCVS-Group (2010). Caltech 101 dataset, `http://grey.colorado.edu`.

Torczon, V. (1997). On the convergence of pattern search algorithms, *SIAM Journal on Optimization* **7**(1): 1–25.

Tsai, A., Yezzi, A., Wells, J.W., Tempany, C., Tucker, D., Fan, A., Grimson, W.E. and Willsky, A. (2003). A shape-based approach to the segmentation of medical imagery using level sets, *IEEE Transactions on Medical Imaging* **22**(2): 137–154.

Vincent, L. and Soille, P. (1991). Watersheds in digital spaces: An efficient algorithm based on immersion simulations, *Pattern Analysis and Machine Intelligence* **13**(6): 583–598.

Wang, J., Kong, J., Lu, Y., Qi, M. and Zhang, B. (2008). A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints, *Computerized Medical Imaging and Graphics* **32**(8): 685–698.

Wang, Z. and Vemuri, B.C. (2004). An affine invariant tensor dissimilarity measure and its applications to tensor-valued image segmentation, *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA*, pp. 228–233.

Mehryar Emambakhsh received his B.Sc. and M.Sc. degrees in electrical engineering from Islamic Azad University, Shahre Rey Branch, Tehran, and the Sahand University of Technology, Tabriz, in 2007 and 2010, respectively. From 2009 to 2010, he worked as a computer programmer at Science Counter Inc. in Tehran over text processing, ontological information retrieval and web-page ranking algorithms. His research interests are image segmentation, object detection and recognition, as well as video content analysis and modeling.

Hossein Ebrahimnezhad received his B.Sc. and M.Sc. degrees in electrical engineering from Tabriz University and the K.N. Toosi University of Technology in 1994 and 1996, respectively. In 2007, he received the Ph.D. degree from Tarbiat Modares University. His research interests include image and video processing, computer vision and 3D reconstruction, soft computing, object detection and pattern analysis. Currently, he is an assistant professor at the Sahand University of Technology. He is a member of the IEEE.

Mohammad Hossein Sedaaghi was born in Tehran, Iran. He received the B.Sc. and M.Sc. degrees from the Sharif University of Technology, Tehran, in 1986 and 1987, respectively. In 1998, he received the Ph.D. degree from Liverpool University. He is now an associate professor at the Sahand University of Technology, Tabriz. His research interests include signal/image processing, pattern recognition, machine learning and biometrics.

Received: 24 February 2010
Revised: 5 July 2010
Re-revised: 28 July 2010