



GC-ASM: Synergistic integration of graph-cut and active shape model strategies for medical image segmentation [☆]

Xinjian Chen ^a, Jayaram K. Udupa ^{b,*}, Abass Alavi ^c, Drew A. Torigian ^c

^a School of Electronics and Information Engineering, Soochow University, Suzhou 215006, China

^b Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, United States

^c Hospital of the University of Pennsylvania, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104-6021, United States

ARTICLE INFO

Article history:

Received 3 June 2010

Accepted 1 December 2012

Available online 20 December 2012

Keywords:

Object recognition

Image segmentation

Statistical shape models

Graph cut

ABSTRACT

Image segmentation methods may be classified into two categories: purely image based and model based. Each of these two classes has its own advantages and disadvantages. In this paper, we propose a novel synergistic combination of the image based graph-cut (GC) method with the model based ASM method to arrive at the GC-ASM method for medical image segmentation. A multi-object GC cost function is proposed which effectively integrates the ASM shape information into the GC framework. The proposed method consists of two phases: model building and segmentation. In the model building phase, the ASM model is built and the parameters of the GC are estimated. The segmentation phase consists of two main steps: initialization (recognition) and delineation. For initialization, an automatic method is proposed which estimates the pose (translation, orientation, and scale) of the model, and obtains a rough segmentation result which also provides the shape information for the GC method. For delineation, an iterative GC-ASM algorithm is proposed which performs finer delineation based on the initialization results. The proposed methods are implemented to operate on 2D images and evaluated on clinical chest CT, abdominal CT, and foot MRI data sets. The results show the following: (a) An overall delineation accuracy of TPVF > 96%, FPVF < 0.6% can be achieved via GC-ASM for different objects, modalities, and body regions. (b) GC-ASM improves over ASM in its accuracy and precision to search region. (c) GC-ASM requires far fewer landmarks (about 1/3 of ASM) than ASM. (d) GC-ASM achieves full automation in the segmentation step compared to GC which requires seed specification and improves on the accuracy of GC. (e) One disadvantage of GC-ASM is its increased computational expense owing to the iterative nature of the algorithm.

Crown Copyright © 2012 Published by Elsevier Inc. All rights reserved.

1. Introduction

1.1. Background

Automatic image segmentation is a fundamental and challenging problem in computer vision and medical image analysis. In spite of several decades of research and many key advances, several challenges still remain in this area. The whole segmentation operation can be thought of as consisting of two related processes: recognition and delineation. *Recognition* is the high-level process of determining roughly the whereabouts of an object of interest and distinguishing it from other object-like entities in the image. *Delineation* is the low-level process of determining the precise spatial

extent of the object in the image. The efficient incorporation of high-level recognition help with accurate low-level delineation has remained a challenge in medical image segmentation.

From the consideration of the recognition process and its implications, it is important to distinguish between two types of activities in medical image processing – *Computer Aided Diagnosis* (CAD), and another, which, for the lack of a widely accepted name, we will refer to as *Computer Aided Visualization and Analysis* (CAVA). In CAD, the focus is on disease diagnosis via images. CAVA, however, deals with the science underlying computerized methods of image processing, analysis, and 3D visualization to facilitate new therapeutic strategies, basic clinical research, education, and training. In CAD, the goal of recognition (or detection) is to identify/detect an abnormality such as a mass or lesion in the images of a body region. In CAVA, on the other hand, the goal of recognition is to identify anatomic organs, not for the purpose of determining their presence/absence, but for subsequently delineating, quantifying, and visualizing them and their pathological deviations. In CAD, recognition/detection may be itself a goal, or it may be followed by

[☆] This paper has been recommended for acceptance by Nikos Paragios.

* Corresponding author. Address: Department of Radiology, University of Pennsylvania, 423 Guardian Drive, Philadelphia, PA 19104-6021, United States. Fax: +1 (215) 898 9145.

E-mail address: jay@mail.med.upenn.edu (J.K. Udupa).

delineation to study the morphology/architecture of the lesion for its classification as benign or malignant. In CAVA, thus far, far more emphasis has been given to delineation than for automatic organ recognition.

Different schemes have been employed in books and review papers [1,2] for classifying segmentation methods. The scheme we have used is motivated by the challenges that currently exist in segmentation. In this paper, the focus is on delineation, whether it is for CAVA or CAD, and on recognition from the perspective of CAVA and not CAD. Recognition for CAVA is usually implemented with the user providing information interactively, for example, in the form of seeds, initial boundary, or an initial model placed near the object. In view of these considerations, our review will focus mainly on delineation.

1.1.1. Classification of delineation methods

Delineation methods may be globally classified into three groups: Those that rely entirely on information available only in the given scene, those that use prior *shape models* (SMs) for the object of interest, and a third emerging hybrid group that combines the strengths of different methods. For brevity, we shall refer to these groups as *purely image based approaches* or *pl-approaches*, *SM-approaches*, and *hybrid*. Starting from the earliest publications in delineation [3], a majority of the methods investigated during the 4–5 decades of segmentation research are *pl-approaches*. *SM-* and *hybrid approaches* are recent and are rightfully attracting a great deal of attention.

1.1.2. *pl-approaches*

These methods maybe further classified into two groups – *boundary-based* (*Bpl*) and *region-based* (*Rpl*). Early attempts toward automating boundary tracing took *optimum boundary detection* approaches [4,5]. Their inadequacies led to the *active contour* (AC) methods [6–8]. In an attempt to avoid the post-delineation correction often required by these methods, *live wire* (LW) user-steered delineation methods have emerged [9–11]. In these techniques, recognition by a human operator and delineation by the computer take place cooperatively, synergistically, and with a certain degree of continuity in a tightly coupled manner. Another class of boundary-based delineation techniques called *level-set* (LS) methods [12–14] have emerged also for overcoming the inadequacies of the deformable boundary methods. They can handle changing topology, and can deal with local voxel level deformations. As for *region-based* approaches, from simple *intensity thresholding* [15] to methods of automatically finding the thresholds in an optimum fashion have been devised [16,17]. *Region growing* methods [18,19] evolved to overcome the drawbacks of thresholding. *Clustering* or *feature space partitioning* methods [20] are popular particularly in brain MR image analysis [21,22]. The commonly used clustering methods are *k-nearest neighbor*, *c-means*, and *fuzzy c-means* techniques. *Graph-based* approaches pose delineation as a graph problem and present solutions via graph search algorithms. Two actively pursued classes of methods in this group are *graph-cut* (GC) [23–25] and *fuzzy connectedness* (FC) [26–28]. In *Markov Random Field* (MRF), the spatial information in an image can be encoded statistically through contextual constraints of neighboring voxels [29,30]. In the *watershed* (WS) –based methods [31–33,57,58], the region occupied by an object is considered to be the set of all those voxels which get flooded under certain conditions. In the so-called *Mumford-Shah model* [34,35], the idea is to partition the domain of the given scene optimally into different object regions by minimizing a functional.

1.1.3. *SM approaches*

The basic premise of these methods is to first construct a spatial/geographic statistical model that contains salient shape

information and its expected variations among all manifestations of the object of interest. Subsequently this model is utilized to search for a shape in the given scene that agrees with the model. Methods in this group differ in how the model is represented and built, if and how scene intensity texture information is incorporated statistically into the model, how an initial position of the model is specified in the scene, how the particular shape is found in the scene, and how the model is used as a spatial prior in tissue classification. Prominent in this group are *Active Shape and Appearance Models* (ASM, AAM) [36–38], *m-Reps* [39,40], and *atlas-based methods* [41–43]. ASM/AAM methods use “landmarks” to represent shape and principal component analysis to capture the major modes of variation in shape observed in the training data sets. *m-reps* use a medial axis representation to capture shape information. *Atlas-based methods* rely on a previously created atlas that captures in it information about the object label, geometry, shape, scene intensity properties, and object assembly architecture defining the relationship among objects. These methods have proven useful especially in classifying tissues of the brain. The given image is first elastically matched to the standard template of the atlas. The prior likelihood from the atlas is then combined with image intensity to derive a posteriori tissue likelihood.

1.2. Motivation

1.2.1. Hybrid-approaches

The *pl-approaches* and the *SM-approaches* each have their own strengths and weaknesses. The premise of hybrid approaches is to combine the complementary strengths of the individual methods to arrive at a more powerful hybrid strategy that can overcome the weakness of the component methods. The *hybrid strategies* investigated in the literature combine methods as follows: (hs1): two or more *Bpl-methods* or two or more *Rpl-methods*; (hs2): *Bpl-methods* and *Rpl-methods*; (hs3): *Bpl-methods* with an *SM-method*; (hs4): *Rpl-methods* with an *SM-method*. Combination of two *SM-approaches* developed as a hybrid method is not known. Compared to *pl-* and *SM-approaches*, research done on hybrid approaches is far less advanced and is mainly under strategies (hs1) [44,45] and (hs2) [46,47]. In (hs1) and (hs2) strategies, we can advance segmentation to the point where we can, at best, harness all or most of the information present in the given image for influencing segmentation. In contrast, strategies (hs3) [48,49] and (hs4) [50,51] allow us, in addition to such information, to bring in information about the geography, shape, and relationships of objects. This latter area, however, is in its early stages, and, we believe, offers the greatest potential for major advances in segmentation algorithms. The proposed method falls under this latter category.

In this paper, continuing in the spirit of hybrid strategy (hs4), we propose a novel synergistic way of combining the *pl* graph-cut approach with the model based ASM method to arrive at the GC-ASM strategy. Some related research works [52–56,59,60,69–71] in this area are summarized here. Besbes et al. [52] proposed a discrete MRF based segmentation method which combines shape priors and regional statistics. However, this method does not perform segmentation at the pixel level. Freedman and Zhang [53] incorporated the shape template into the graph-cut formulation as a distance function. However, this method relies crucially on having user input. Based on the latter, Ayvaci and Freedman [54] proposed a joint-registration segmentation method which removes the user interaction requirement and resolves the problem of template registration. However, this method requires proper registration of the shape template for an accurate segmentation. Kumar et al. [55] used a Markov Random Field representation where the latent shape model variables are integrated via expectation maximization. While shape information is utilized in a principled Bayesian manner, this approach is computationally intensive

requiring separate energy minimization. Malcolm et al. [56] impose the shape prior model on the terminal edges and perform graph cuts iteratively starting with an initial contour. Their method constructs a statistical shape space using kernel principal component analysis. This method also relies on user input. Kohli et al. [59] present an algorithm for performing simultaneous segmentation and 3D pose estimation of a human body from multiple views. This method optimizes a cost function based on a Conditional Random Field that combines shape prior information and pose of the subject via dynamic graph cuts. Most closely related to our work, Vu and Manjunath [60] propose a shape prior segmentation method using graph cuts capable of segmenting multiple objects. The shape prior energy is based on a shape distance popular with level set approaches. However, the shape used is a simple fixed shape.

In contrast to these methods, the strategy proposed in this paper is an automatic multi-object segmentation method, although model building requires user help. And most importantly, different from all the above mentioned shape prior-integrated GC methods, this technique does not need to do shape registration because of the proposed automatic initialization step. The proposed GC-ASM process effectively combines the rich statistical shape information embodied in ASM with the energy minimization capabilities of the GC method for delineation. A multi-object GC cost function is proposed which effectively integrates the ASM shape information into the GC paradigm. The proposed method consists of two phases: model building and segmentation. In the model building phase, the ASM model is built and the parameters of GC are trained. The segmentation phase consists of two main steps: initialization (or recognition) and delineation. In the initialization step, the model pose (translation, orientation, and scale) is estimated automatically yielding also a rough segmentation result. In the delineation step, an iterative algorithm performs finer delineation based on the initialization result, alternating between image segmentation via GC and the update of the ASM shape model. The proposed method was tested on three 2D medical image data sets: abdominal CT, chest CT, and foot MRI. In these data sets, the acquired images are naturally roughly aligned.

To summarize, the main contributions of this paper are as follows: (1) A multi-object initialization/recognition method which effectively combines the GC and ASM methods. (2) A novel shape integrated iterative GC-ASM delineation method, which alternates between object delineation via GC and updating the ASM shape model. (3) A novel strategy for assessing the effectiveness of object recognition and its influence on object delineation.

The paper is organized as follows. In Section II, the complete methodology of GC-ASM is described. In Section III, a detailed evaluation of this method in terms of its accuracy and efficiency in both recognition and delineation by utilizing three different data sets is presented. In Section IV, we summarize our conclusions. A preliminary version of this paper appeared in the conference proceedings of the SPIE 2009 Medical Imaging Symposium [61].

2. Graph cut with active shape models

2.1. Overview of approach

In this section, an overview of the GC-ASM approach is presented. To test the feasibility and the effectiveness of the ideas underlying GC-ASM, this paper focuses on the segmentation problem in 2D images. Some sample results on 3D images are presented at the end.

Let $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_m$ be the physical objects of interest, in a given body region \mathbf{B} , such as the human liver, lungs, and heart, in the thoracic region. For segmenting the boundaries of these objects in an image of a particular manifestation of them, ASM captures

the statistical variations in the boundaries of these objects within the objects' family via a statistical shape model M associated with each object. In our case, M constitutes the model of an assembly of $m \geq 1$ objects. GC-ASM determines a cost structure K associated with M (and, hence, the set of objects represented by M) via the principles of optimal cut underlying the graph cut method. As per this cost structure, every shape instance x of the multiple objects in \mathbf{B} generated by M is assigned a total GC cost $K(x)$ in a given image. This cost is determined from the graph cut method. GC-ASM seeks that collection of object boundaries in the given image of \mathbf{B} , which satisfies the shape constraints of M and for which the cost $K(x)$ in the given image is the smallest possible. The main steps involved in GC-ASM are listed below, and a flowchart of the proposed method is given in Fig. 1. Each step is described in detail in each subsection of Section 2.

2.2. Procedure GC-ASM

2.2.1. Model building phase

- T1. Specify landmarks on boundaries of objects $\mathbf{O}_1, \dots, \mathbf{O}_m$ in the training images provided for body region \mathbf{B} .
- T2. Construct a shape model M for the objects in \mathbf{B} from the landmarks and training images.
- T3. Estimate parameters for GC delineation.

2.2.2. Segmentation phase

- S1. Initialization/recognition: Determine, in the given image I of \mathbf{B} , the pose at which M should be set in I so that the model boundaries are close to the real object boundaries in I . Let the shape instance of the multiple object assembly corresponding to the recognized site be x .
- S2. Delineation: For the shape instance x of the multiple object assembly, determine the best oriented boundaries in I as per the GC method.
- S3. If the convergence criterion is satisfied, output the best oriented boundaries found in S2 and stop. Otherwise, subject x to the constraints of model M and go to Step S2.

In this procedure, T1–T3 constitute training or model creation steps, and S1–S3 represent initialization/recognition and delineation steps. These steps are described in detail in Sub-sections

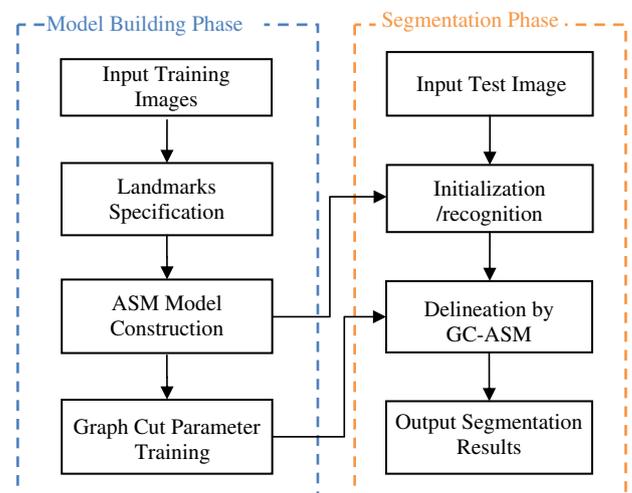


Fig. 1. A flowchart of the proposed method.

2.3.1–Sections 2.4.3. But first, the relevant GC principles are described in Section 2.2.

2.3. Multi-object shape integrated graph cut

GC segmentation can be formulated as an energy minimization problem such that for a set of pixels P and a set of labels L , the goal is to find a labeling $f: P \rightarrow L$ that minimizes the energy function $E(f)$.

$$E(f) = \sum_{p \in P} R_p(f_p) + \sum_{p \in P, q \in N_p} B_{p,q}(f_p, f_q) \quad (1)$$

where N_p is the set of pixels in the neighborhood of p , $R_p(f_p)$ is the cost of assigning label $f_p \in L$ to p , and $B_{p,q}(f_p, f_q)$ is the cost of assigning labels $f_p, f_q \in L$ to p and q . In two-class labeling, $L = \{0, 1\}$, the problem can be solved efficiently with graph cuts in polynomial time when $B_{p,q}$ is a submodular function, i.e., $B_{p,q}(0, 0) + B_{p,q}(1, 1) \leq B_{p,q}(0, 1) + B_{p,q}(1, 0)$ [62]. However, for the general multi-object segmentation problem, $L = \{0, 1, \dots, k\}$, $k \geq 2$, it is known to be NP-hard [63,75].

2.3.1. Shape integrated energy function

In our framework, the unary cost $R_p(f_p)$ is the sum of a data penalty $D_p(f_p)$ and a shape prior $S_p(f_p)$ term. The data term is defined based on image intensity and can be considered as a log likelihood of the image intensity for object [25,77,78]. The shape prior term is independent of image information, and the boundary term is based on the gradient of the image intensity [25].

The proposed shape integrated energy function is defined as follows:

$$E = \sum_{p \in P} [\alpha \cdot D_p(f_p) + \beta \cdot S_p(x_{\mathbf{O}})] + \sum_{p \in P, q \in N_p} \gamma \cdot B_{p,q}(f_p, f_q) \quad (2)$$

where α, β, γ are the weights for the data term, shape term S_p , and boundary term, respectively, satisfying $\alpha + \beta + \gamma = 1$. These components are defined as follows:

$$D_p(f_p) = \begin{cases} -\ln \Pr(I_p|O), & \text{if } f_p = \text{object label} \\ -\ln \Pr(I_p|B), & \text{if } f_p = \text{background label} \end{cases} \quad (3)$$

$$B_{p,q}(f_p, f_q) = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{d(p, q)} \cdot \delta(f_p, f_q) \quad (4)$$

and

$$\delta(f_p, f_q) = \begin{cases} 1, & \text{if } f_p \neq f_q \\ 0, & \text{otherwise} \end{cases}$$

where I_p is the intensity of pixel p , *object label* is the label of the object (foreground). $\Pr(I_p|O)$ and $\Pr(I_p|B)$ are the probability of intensity of pixel p belonging to object and background intensity, respectively, which are estimated from object and background intensity histograms during the training phase. $d(p, q)$ is the Euclidean distance between pixels p and q , and σ is the standard deviation of the intensity differences of neighboring voxels along the boundary. The shape prior term we have employed is similar to the one proposed by Kohli et al. [59].

$$S_p(x_{\mathbf{O}}) = 1 - \exp\left(-\frac{d(p, x_{\mathbf{O}})}{r_{\mathbf{O}}}\right) \quad (5)$$

where $d(p, x_{\mathbf{O}})$ is the distance from pixel p to the set of pixels which constitute the interior of the current shape $x_{\mathbf{O}}$ of \mathbf{O} . (Note that if p is in the interior of $x_{\mathbf{O}}$, then $d(p, x_{\mathbf{O}}) = 0$.) $r_{\mathbf{O}}$ is the radius of a circle that just encloses $x_{\mathbf{O}}$. Any one of the linear time distance transform methods [64,72–74] may be used for computing this distance.

2.3.2. Minimizing E with graph cuts

Let G be a weighted graph (V, A) , where V is a set of nodes, and A is a set of weighted arcs. The multi-way cut problem aims at finding the cut C with the minimal cost $|C|$, expressed as the sum of weights of edges in the cut. One way of minimizing GC energy is to do iterative binary segmentation based on two way graph cut. However, extracting multiple objects simultaneously would be more robust, convenient, and faster compared to the iterative binary approach [25]. In our implementation, we segment the multiple objects by using the α -expansion method [24].

The graph is designed as follows. We take $V = P \cup L$, i.e., V contains all the pixel nodes and multiple terminals corresponding to the labels in L which represent objects of interest plus the background. $A = A_N \cup A_T$, where A_N is the n -links which connect pixels p and q ($p \in P, q \in N_p$) and with a weight of w_{pq} . A_T is the set of t -links which connect each pixel p in P and terminals $\ell \in L$ and with a weight of $w_{p\ell}$. The desired graph with cut cost $|C|$ equaling $E(f)$ is constructed using the following weight assignments:

$$w_{pq} = \gamma \cdot B_{p,q} \quad (6)$$

$$w_{p\ell} = \Theta - (\alpha \cdot D_p(\ell) + \beta \cdot S_p(\ell)) \quad (7)$$

where Θ is a constant that is large enough to make the weights $w_{p\ell}$ positive.

2.4. Model building

2.4.1. T1: Specifying landmarks

One way to describe a particular instance of the shape of an object \mathbf{O} is by locating a finite number of points on its boundary, referred to as landmarks. A mathematical representation of an n -point shape in a d -dimensional space may be obtained by concatenating each dimension into a $d \times n$ component vector. In this paper, only 2D shapes are considered, hence $d = 2$. Suppose each object \mathbf{O}_i considered for inclusion in the model has l_i landmarks, $1 \leq i \leq m$. Then the vector representation for planar shapes would be:

$$x = (x_1^{O_1}, y_1^{O_1}, x_2^{O_1}, y_2^{O_1}, \dots, x_{l_1}^{O_1}, y_{l_1}^{O_1}, x_1^{O_2}, y_1^{O_2}, \dots, x_{l_2}^{O_2}, y_{l_2}^{O_2}, \dots, x_{l_m}^{O_m}, y_{l_m}^{O_m}). \quad (8)$$

In many ASM studies, a manual procedure is used to label the landmarks in a training set, although automatic methods are also available for this purpose. That is, for each image of the training set, operators locate the shape visually, and then identify significant landmarks on that shape. It is important that the landmarks are accurately located and that there is an exact correspondence among landmark labels in different instances of the training shapes. For our approach, any such method will work, although we have used the manual method in producing all presented results.

2.4.2. T2: Building the model M

To obtain a true shape representation of an object family \mathbf{O}_i , location, scale, and rotation effects within the family need to be filtered out. This is usually done by aligning the shapes within \mathbf{O}_i (in the training set) to each other by changing the pose parameters (scale, rotation, and translation) [36]. For multiple objects, the object assemblies are aligned. The model M is then constructed following the ASM procedure [36] considering the multiple objects.

2.4.3. T3: Estimating GC parameters

During the training stage, the histograms of intensity for each object are estimated from the training images. Based on this, $\Pr(I_p|O)$ and $\Pr(I_p|B)$ can be computed. As for parameters α, β and γ in Eqn. (1), since $\alpha + \beta + \gamma = 1$, we estimate only α and β by

optimization of the accuracy of segmentation as a function of α and β and set $\gamma = 1 - \alpha - \beta$. In our implementation, we use the simple but effective gradient decent method [65]. Suppose $Accu(\alpha, \beta)$ represents the algorithm's accuracy (here we use the true positive volume fraction [66]) on the training data set with parameters α, β , α and β are initialized to 0.35 each, and $Accu(\alpha, \beta)$ is optimized over the training data set to determine the best α and β .

2.5. Segmentation

The segmentation process consists of two main steps: initialization or object recognition and delineation. The initialization step aims at estimating the pose of the model, and obtaining a rough segmentation result. In the second step, an iterative GC-ASM delineation algorithm derives a finer delineation based on the initialization result. In both these steps, the shape information is integrated into the GC framework.

2.5.1. S1: Initialization/recognition

The purpose of this step is to recognize the desired objects in I ; that is, to find a shape instance x of M which is sufficiently close to the boundaries of the objects in I that we want to segment. x is subsequently modified in Steps S2 and S3 to best fit the image information in I from the consideration of both GC and ASM.

The automatic initialization method proposed here is an essential underpinning of the GC-ASM method. It relies on the fact that, corresponding to a pose of a shape instance x of M that is close to the pose of the correct boundary of the objects O_i ($1 \leq i \leq m$) in I , the GC cost (the sum of the minimum cut cost of all objects) is likely to be sharply smaller than the cost found at other poses in I . Suppose \mathbf{p} denotes the pose vector for the object assembly, with a location component x and y , scale component s , and orientation component θ . Our goal is to find the best pose in I for the model M . Our experiments indicate that, the small variations in orientation observed in clinical images can be automatically handled by the GC-ASM algorithms and thus θ can be ignored. Thus \mathbf{p} becomes 3-dimensional $(x, y, s)^t$. Let $K(M, I, \mathbf{p})$ be the GC minimum cut cost summed over all objects achieved at \mathbf{p} in I using model M . Then the recognition task is to find:

$$\mathbf{p}^* = \operatorname{argmin}_{\mathbf{p}} [K(M, I, \mathbf{p})] \quad (9)$$

The algorithm for automatic recognition runs as follows:

Algorithm for Automatic Initialization

- S1.1. Initialize the set of potential pose vectors \mathbf{p} for M over which the optimum needs to be found.
- S1.2. For each potential pose vector \mathbf{p} in the set, do steps S1.3 and S1.4.
- S1.3. ASM searching. Put model M at \mathbf{p} relative to I , then deform it as per the standard ASM method.
- S1.4. GC segmentation. Based on the cost function E (Eq. (1)), perform GC segmentation and find the minimum cut cost.
- S1.5. Select the deformed ASM shape and GC result corresponding to the lowest cut cost over all \mathbf{p} as the automatic recognition result for the subsequent steps.

The set of potential pose vectors is determined by discretizing the three-dimensional search space of $(x, y, s)^t$. If multiple objects are included, the search space becomes considerably smaller, and recognition becomes vastly more efficient because of the constraints brought on by the multiple objects. For example, any pose vector for which the objects in M extend significantly beyond the body region in I need not be included in the potential pose vector set. This underscores the importance of the actual distribution of the objects that are considered to be included in M . It is clear that

initialization efficiency and accuracy both depend on a number of factors, such as the size of the image I , the number of objects, the size of objects, and the objects' distribution. These dependencies and the actual size of the search space for different body regions are demonstrated by our results. Since the search space dealt with in the initialization method is small, we find the optimal \mathbf{p}^* in Eq. (6) by exhaustive search, which also avoids settling at local minima.

2.5.2. S2: Delineation

This step assumes that the recognized (initialized) shape instance x_p of the objects in M derived from Step S1 is sufficiently close to the actual boundaries of O_i ($1 \leq i \leq m$) in I . It then determines what the new position of the landmarks of the objects represented in x_p should be such that the minimum graph cut cost is the smallest possible. This is accomplished through an iterative GC-ASM algorithm, called IGC-ASM, presented below.

Algorithm IGC-ASM

Input: Initialized shape assembly x_p .

Output: Resulting shape x_o and the associated object boundaries.

begin
while number of iterations < $n_{iteration}$ do

1. Perform GC segmentation based on shape assembly x_p ;
2. Compute the new position of the landmarks by moving each landmark in x_p to the point closest on the GC boundary of the corresponding object; call the resulting shape assembly x_{new} ;
3. If no landmarks moved, then set x_{new} as x_o and go to 4;

otherwise, subject x_{new} to the constraints of model M , and call the result x_p .
endwhile

Perform one final GC segmentation based on x_o , and compute and output the associated object boundaries.

end

In our implementation, $n_{iteration}$ is set to 6. Also we limit the number of pixel units by which landmarks can move within any iteration to 6. A simple morphological operation (dilation and erosion) is applied to the result if there are holes in the segmentation result. In our construction of the ASM model, the shape truncation level was set as 95% (two standard deviations from the mean). During the searching stage, if any shape parameters go beyond two standard deviations from the mean, then they are constrained by applying ellipsoidal constraints in the shape parameter space.

2.5.3. S3: Subjecting to model constraints

The convergence criterion used here is a measure of the distance between two shapes encountered in two consecutive executions of Step S2. This measure is simply the maximum of the distance between corresponding landmarks in the two shapes among all objects. If this distance is greater than 0.5 pixel unit, the optimum shape found in Step S2 is subjected to the constraints of model M . Then the iterative process is continued by going back to Step S2. Otherwise, the GC-ASM process is considered to have converged and it stops with an output of the optimum shape and the optimum oriented boundaries found in Step S2.

3. Experimental results

In this section, we demonstrate both qualitatively, through image display, and quantitatively, through evaluation experiments, the extent of effectiveness of the GC-ASM method. Three clinical

image data sets, which include abdominal and chest CT, and foot MRI, have been considered. We describe some new strategies for evaluating the recognition method. For delineation evaluation, we use the framework of [66] and focus on the analysis of accuracy and efficiency of GC-ASM. We consider manual segmentation performed in these different data sets to constitute a surrogate of true segmentation for assessing the accuracy of the methods, and use the true positive, false positive, and false negative volume fractions (TPVF, FPVF, and FNVF) from that reference. We compare GC-ASM with ASM, a recently developed multi object oriented ASM method (MOASM) [48,67] and Malcolm et al.'s method [56], which incorporated a shape prior learned from training examples via kernel principal component analysis into the GC method, denoted by GCKPCA for further reference.

3.1. Image data sets

The image data sets and objects used in the experiments are described in Table 1. The CT sets constitute slices selected from 3D studies on a Siemens Sensation 16 CT scanner with a slice spacing of 5 mm, image size of 512×512 , and pixel size of $0.78 \times 0.78 \text{ mm}^2$. The MRI set consists of slices selected from 3D studies on a GE 1.5T MRI scanner with a slice spacing of 1.3 mm, image size of 256×256 , and pixel size of $0.55 \times 0.55 \text{ mm}^2$. In each set, 40 slices selected from full 3D images, acquired from fifteen different subjects are used. These slices are approximately at the same location in the body, so that, for each object, the 40 2D images in each set can be considered to represent images of a family of objects of the same shape. Two to three slices are taken on average from the same subject's data, either from the same 3D image or from different 3D images. Among them, 25 images are selected as training images, and the rest are used as testing images. In this selection, we have made sure that slices from the same subject do not appear in both the training and test sets. We considered four objects in each data set as listed in Table 1. The trained optimal GC parameters (α, β, γ) for these three data sets are $(0.25, 0.5, 0.25)$, $(0.25, 0.45, 0.30)$ and $(0.27, 0.53, 0.20)$, respectively.

3.2. Qualitative analysis

A subjective inspection revealed that, in all experiments and in all data, the GC-ASM results matched the perceived boundary very well. Some examples are displayed for each of the three data sets in Figs. 2–4. Automatic initialization based on location and scale search worked well in all cases in the sense that initialized shapes were found close to the true boundary. Our experiments indicate that, because of the orientedness nature of the GC-ASM approach, the small variations in orientation observed in clinical images can be automatically handled, and therefore, recognition based on searching in the space spanned by \mathbf{p} considering only x , y , and s was effective. In Figs. 2–4, (b) shows the original image on which the mean shape of the model is overlaid, which is the starting point for the recognition algorithm; (c) shows the recognized (initialized) shapes output by Step S1, which are the initial shapes input to Step S2 of the GC-ASM method; and (d) displays the final segmentation results. In these figures, we display in (e and f) and (g) also the delineation results for ASM, MOASM and GCKPCA. For ASM and GCKPCA, the recognition result of GC-ASM was used to

provide the initialization. For MOASM, its own optimal recognition method [48,67] was used for initialization. The same number of landmarks has been used in the models employed for all methods in the results presented in Figs. 2–4.

3.3. Quantitative analysis

It is important to assess the effectiveness of recognition separately from the overall delineation result since recognition and delineation are interdependent. Therefore, several types of tests were conducted to evaluate the two aspects separately.

3.3.1. Recognition

There are two aspects to the evaluation of the recognition method. The first has to do with the mechanism of searching for the optimum (see Eq. 6) in the pose space of \mathbf{p} , and the second relates to how close to the true pose the recognized model pose comes.

For the reasons already described, we do an explicit search for the optimum pose by discretizing the space of \mathbf{p} . So the first aspect has to do with the step size for x , y , and s for discretization of the pose space and its influence on the accuracy of recognition. In Figs. 5–7, we demonstrate the effect of step size on the ability of the method to correctly position the model so that the delineation becomes accurate. This is expressed in terms of how well the delineated objects match the true delineations (TPVF). As we can see, the recognition strategy is not very sensitive to the step size used in any of x , y , and s . TPVF actually depends on all these parameters x , y , and s . What is shown in Figs. 5–7 is a cross section of this variation passing through the optimum point; for example, while showing the variation with respect to x , the parameters y and s are fixed at their optimum values.

Coming now to the second aspect, we determine how close the optimum pose found comes to the true pose in terms of the concept of “robust region”. *Robust region* (RR) is the set of all pose vectors \mathbf{p} such that when the recognized pose is any pose within RR, the resulting delineation accuracy is acceptably “high” (say, with FPVF and FNVF $< 3\%$). The larger RR is, the more robust the recognition method is, and hence more likely the delineation algorithm's high accuracy. Fig. 8 gives an illustration of RR for the GC-ASM method for an image from the chest CT data set. The white point represents the geometric center of the actually recognized model assembly. The red region is a cross section of the 3D robust region, considering only the location components of \mathbf{p} . In this case, since the white point has fallen in RR, it implies high delineation accuracy. The cross sections of RR are shown for three different scale values $s = 0.79$, $s = 1.21$, $s = 1$, the first two representing the extreme scale values for RR, implying that the size of RR is quite large in the s dimension. Although we used the GC-ASM's recognition results for obtaining ASM results (as well as for GCKPCA), we can define RR for the ASM and GCKPCA methods also for assessing their sensitivity to initialization. Table 2 lists the robust region size for ASM, MOASM, GCKPCA and GC-ASM. We may conclude from Table 2 that MOASM, GCKPCA [56] and GC-ASM perform similarly in object recognition and ASM's RRs in these data sets are empty sets. Note that the original GCKPCA method [56] employed human interaction for recognition and GCKPCA has its own optimal recognition strategy. Therefore, the GCKPCA results actually attest to the

Table 1
Description of the image data sets used in the three segmentation experiments.

Data set	Objects	Image domain	No. of images	Landmarks used for each object (total)
Abdominal CT	Left pelvic bone, vertebra, right pelvic bone, and skin boundary	512×512	40	10, 12, 10, 12 (44)
Chest CT	Right lung, left lung, heart, and skin boundary	512×512	40	10, 9, 10, 12 (41)
Foot MRI	Talus, calcaneus, tibia, and skin boundary	256×256	40	7, 14, 15, 11 (47)

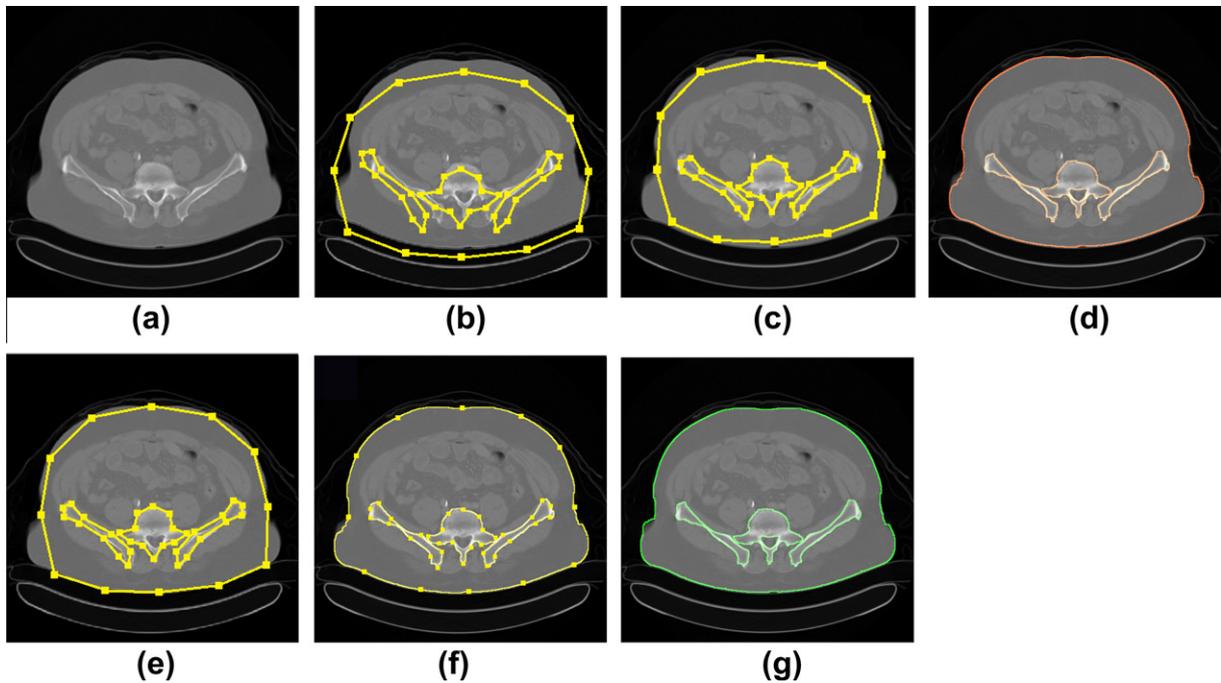


Fig. 2. One example from the abdominal CT data set: the original image and its recognition and delineation results. (a) Original Image; (b) default initial model pose; (c) automatic recognition result by GC-ASM; (d) delineation result by GC-ASM; (e) delineation result by ASM; (f) delineation result by MOASM; (g) delineation result by GCKPCA [56].

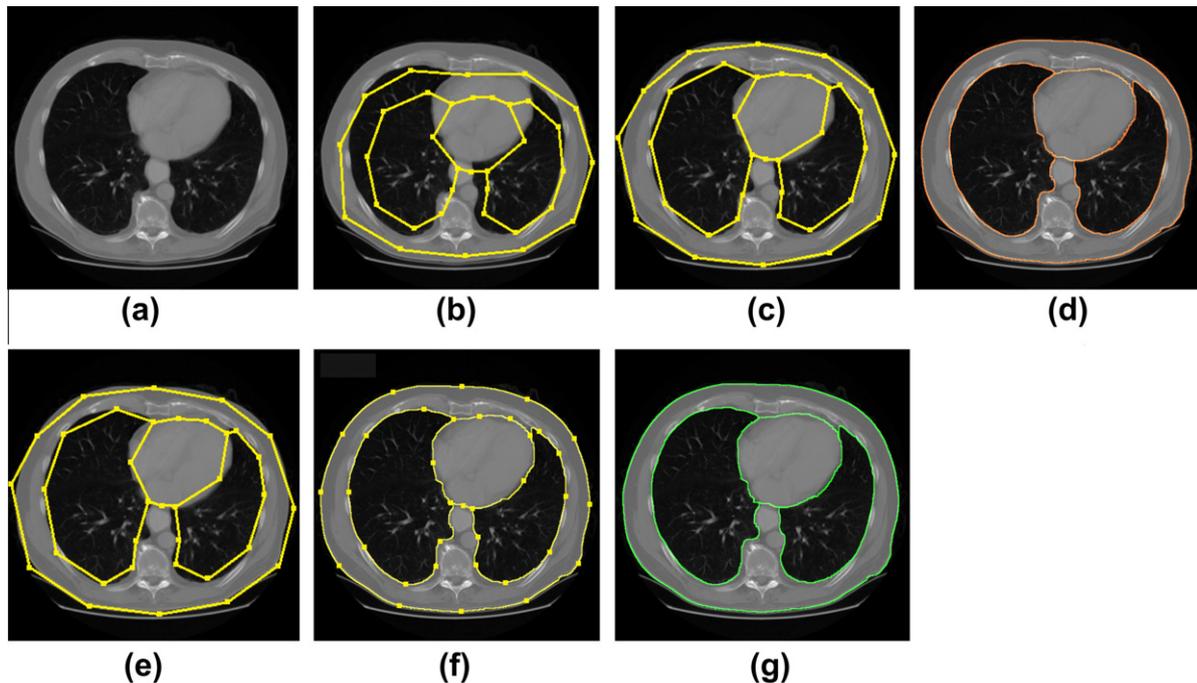


Fig. 3. One example from the chest CT data set: the original image and its recognition and delineation results. (a) Original image; (b) default initial model pose; (c) automatic recognition result by GC-ASM; (d) delineation result by GC-ASM; (e) delineation result by ASM; (f) delineation result by MOASM; (g) delineation result by GCKPCA [56].

ability of GC-ASM in automatic object recognition, in addition to asserting GCKPCA's delineation capability.

3.3.2. Delineation

For the evaluation of delineation, TPVF, FPVF and distance to “true boundary” (Dist2 TB) are used as the performance indices, and the proposed method is compared with all three methods. TPVF of a method indicates the fraction of the total amount of tis-

sue in the true delineation that is captured also by the method. FPVF denotes the amount of tissue falsely identified by the method. Dist2 TB is defined as the average distance from the delineated boundary to the boundary associated with true delineation. MOASM [67] is a hybrid method involving optimization. Its optimization is based on a three-level dynamic programming algorithm, wherein, the first level is at the pixel level which aims to find optimal oriented boundary segments between successive landmarks,

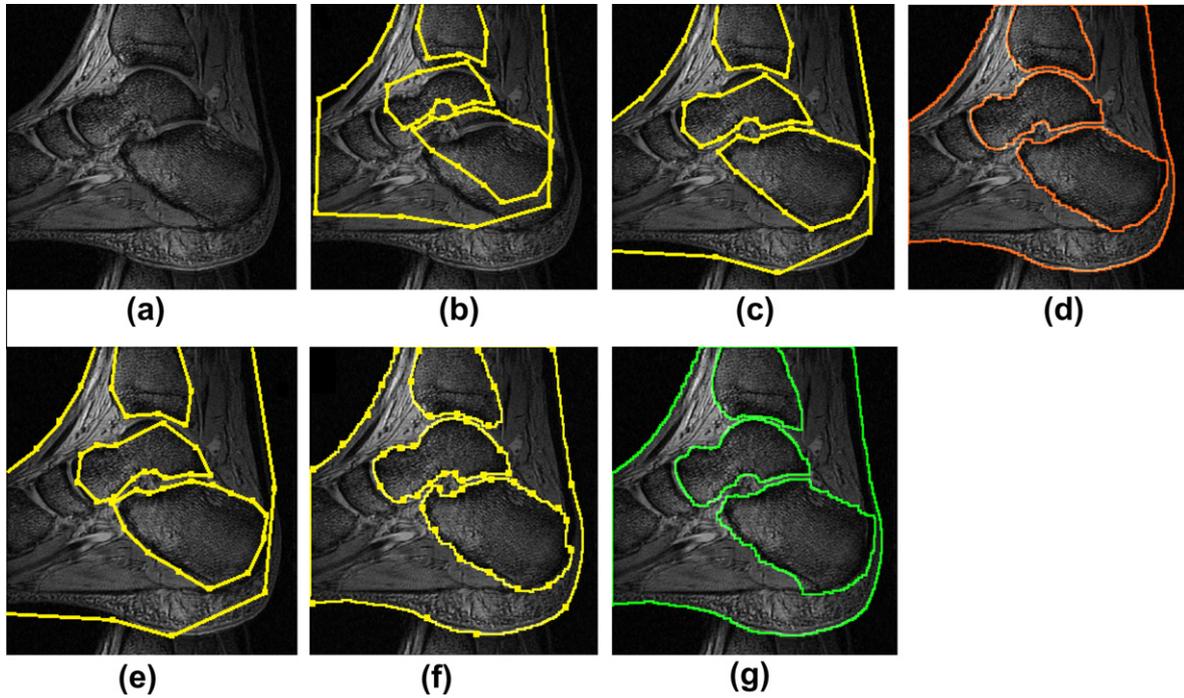


Fig. 4. One example from the foot MRI data set: the original image and its recognition and delineation results. (a) Original Image; (b) default initial model pose; (c) automatic recognition result by GC-ASM; (d) Delineation result by GC-ASM; (e) Delineation result by ASM; (f) delineation result by MOASM; (g) delineation result by GCKPCA [56].

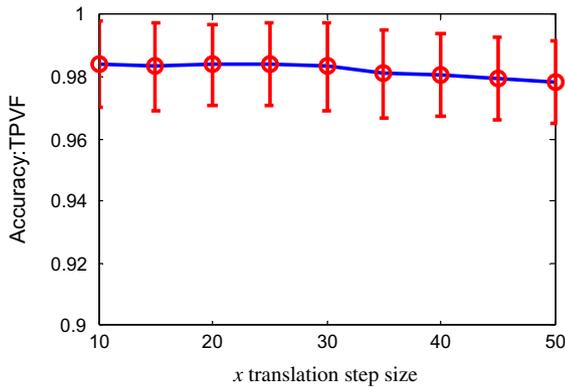


Fig. 5. The dependence of recognition accuracy on x translation step size (tested on chest CT data set).

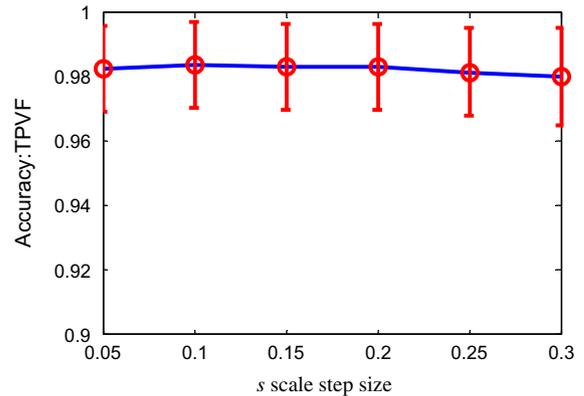


Fig. 7. The dependence of recognition accuracy on s step size (tested on chest CT data set).

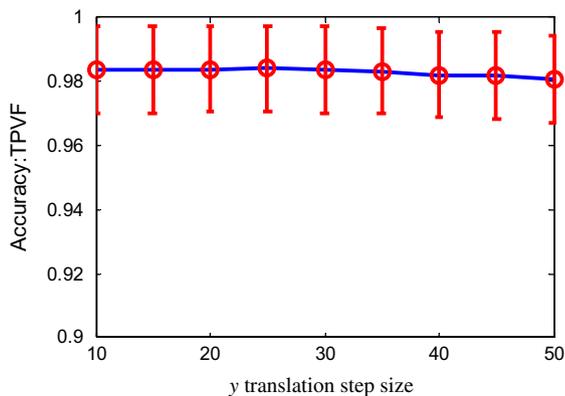


Fig. 6. The dependence of recognition accuracy on y translation step size (tested on chest CT data set).

the second level is at landmark level which aims to find optimal location for the landmarks, and the third level is at the object level which aims to find optimal arrangement of object boundaries over all objects. In GCKPCA [56], a statistical shape space using kernel principal component analysis is constructed based on the set of training shapes. The shape prior model is added into the data term and graph cut is performed iteratively starting with an initial contour. At each iteration, the pre-image of the previous labeling in this shape space is used as the prior probability map, and the negative log of this pre-image is assigned to the terminal weights. In our implementation, we set $\sigma = 20$, $\lambda = 1$, $\mu = 0.4$ as suggested in [56]; and the initialization result of our method was used as the initial contour. The objects were segmented one by one.

Table 3 lists the mean and standard deviation values of TPVF, FPVF, and Dist2TB achieved on the three data sets by all four methods. We note that GC-ASM produces considerably more accurate segmentations than the basic ASM method, and its accuracy is comparable to that of MOASM and GCKPCA [56]. Improved

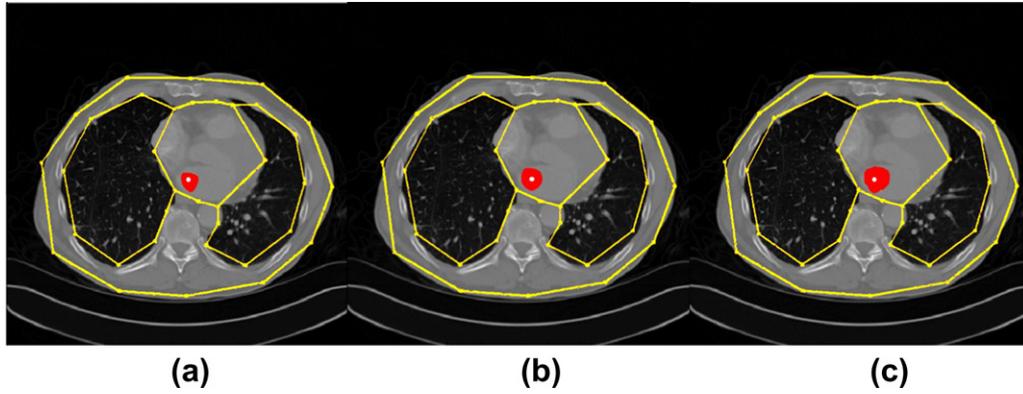


Fig. 8. Illustration of robust region RR for an image from the chest CT data set for (a) $s = 0.79$; (b) $s = 1.21$; (c) $s = 1.00$. The white point is the geometric center of the actually recognized shape assembly. In this example, the recognition is perfect. The red region is an s -cross section of the 3D robust region RR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Robust region comparison for ASM, MOASM [67], GCKPCA [56], and GC-ASM.

Data Set	ASM	MOASM	GCKPCA	GC-ASM
Abdominal CT	NULL	21 pixels \times 23 pixels \times 0.18	23 pixels \times 25 pixels \times 0.20	23 pixels \times 24 pixels \times 0.19
Chest CT	NULL	22 pixels \times 24 pixels \times 0.2	26 pixels \times 25 pixels \times 0.22	25 pixels \times 26 pixels \times 0.21
Foot MRI	NULL	10 pixels \times 11 pixels \times 0.15	10 pixels \times 12 pixels \times 0.15	9 pixels \times 11 pixels \times 0.15

Table 3
Mean and standard deviation of TPVF, FPVF and Dist2 TB for ASM [36], MOASM [67], GCKPCA [56] and GC-ASM over all objects.

Data Set	TPVF (%)			FPVF (%)			Dist2 TB (pixels)		
	Abdom-inal CT	Chest CT	Foot MRI	Abdom-inal CT	Chest CT	Foot MRI	Abdom-inal CT	Chest CT	Foot MRI
ASM	87.09 \pm 1.23	89.03 \pm 1.00	80.62 \pm 1.46	1.26 \pm 0.95	1.12 \pm 0.50	0.72 \pm 0.31	5.4 \pm 3.6	5.8 \pm 4.7	8.3 \pm 6.4
MOASM	97.32 \pm 1.13	98.32 \pm 0.22	96.83 \pm 0.31	0.13 \pm 0.01	0.17 \pm 0.01	0.26 \pm 0.02	2.6 \pm 1.4	2.4 \pm 1.7	2.3 \pm 2.1
GCKPCA	97.55 \pm 0.35	98.62 \pm 0.09	96.72 \pm 0.12	0.25 \pm 0.05	0.19 \pm 0.03	0.55 \pm 0.05	2.6 \pm 1.9	2.5 \pm 1.7	2.5 \pm 2.3
GC-ASM	97.92 \pm 0.21	98.76 \pm 0.06	96.61 \pm 0.08	0.45 \pm 0.07	0.19 \pm 0.02	0.61 \pm 0.09	2.4 \pm 2.1	2.5 \pm 1.6	2.6 \pm 2.5

Table 4
Mean operator time T_o and computational time T_c (in s) in all experiments for ASM, MOASM, GCKPCA and GC-ASM. The number of landmarks used is indicated by n .

Data set	T_o			T_c			
	ASM	MOASM	GC-ASM	ASM	MOASM	GCKPCA	GC-ASM
Abdominal CT	160 ($n = 132$)	60 ($n = 44$)	60 ($n = 44$)	25	30	86	27
Chest CT	130 ($n = 123$)	55 ($n = 41$)	55 ($n = 41$)	23	28	82	25
Foot MRI	170 ($n = 141$)	65 ($n = 47$)	65 ($n = 47$)	18	20	62	19

accuracy of OASM and MOASM over ASM has already been observed [48,67]. In all experiments, a larger number of landmarks have been used in ASM than in GC-ASM and MOASM, which is three times the number of landmarks used in GC-ASM and MOASM (see Table 4). If the same number were used in ASM as in the MOASM and GC-ASM methods, the accuracy would deteriorate considerably, as demonstrated in [48,67].

All methods have been implemented on an Intel Pentium IV PC with a 3.4 GHZ CPU using Matlab programming. In determining the efficiency of a segmentation method, two aspects should be considered - the computation time (T_c) and the human operator time (T_o). The mean T_c and T_o per data set estimated over the three data sets for each experiment are listed in Table 4. T_o measured here is the operator time required in the training step. Table 4 shows that the operator time (training) required in GC-ASM (and MOASM) is less than that of ASM since far fewer landmarks are needed in the former. The computation time required in GC-ASM is a little more than that of ASM because of the iterative nature of the

IGC-ASM algorithm. The computation time for GCKPCA [56] is longer than for GC-ASM. This is due to the fact that for multi-object segmentation, Malcolm’s method cannot segment all objects simultaneously,

A multivariate analysis of variance (MANOVA) test [76] was applied to statistically compare GC-ASM with ASM based on all six efficacy measures: RR size, TPVF, FPVF, Dist2 TB, T_o , and T_c . The MANOVA test of the 15 metric values in each data set with a p -value $p < 0.05$ indicates that the two methods produce different accuracy and efficiency in the overall level. Subsequently, paired t-tests conducted with a $p < 0.05$ indicate that the GC-ASM method produces better accuracy (RR size, TPVF, FPVF and Dist2 TB), and efficiency on operator time (T_o) than the basic ASM method, but ASM achieves better efficiency on computation time (T_c) than the GC-ASM method. Similar tests conducted in comparing GC-ASM with MOASM indicate that significant differences could not be discerned ($p > 0.35$) between the two methods on any of the factors. Similar tests also conducted in comparing GC-ASM with GCKPCA [56]

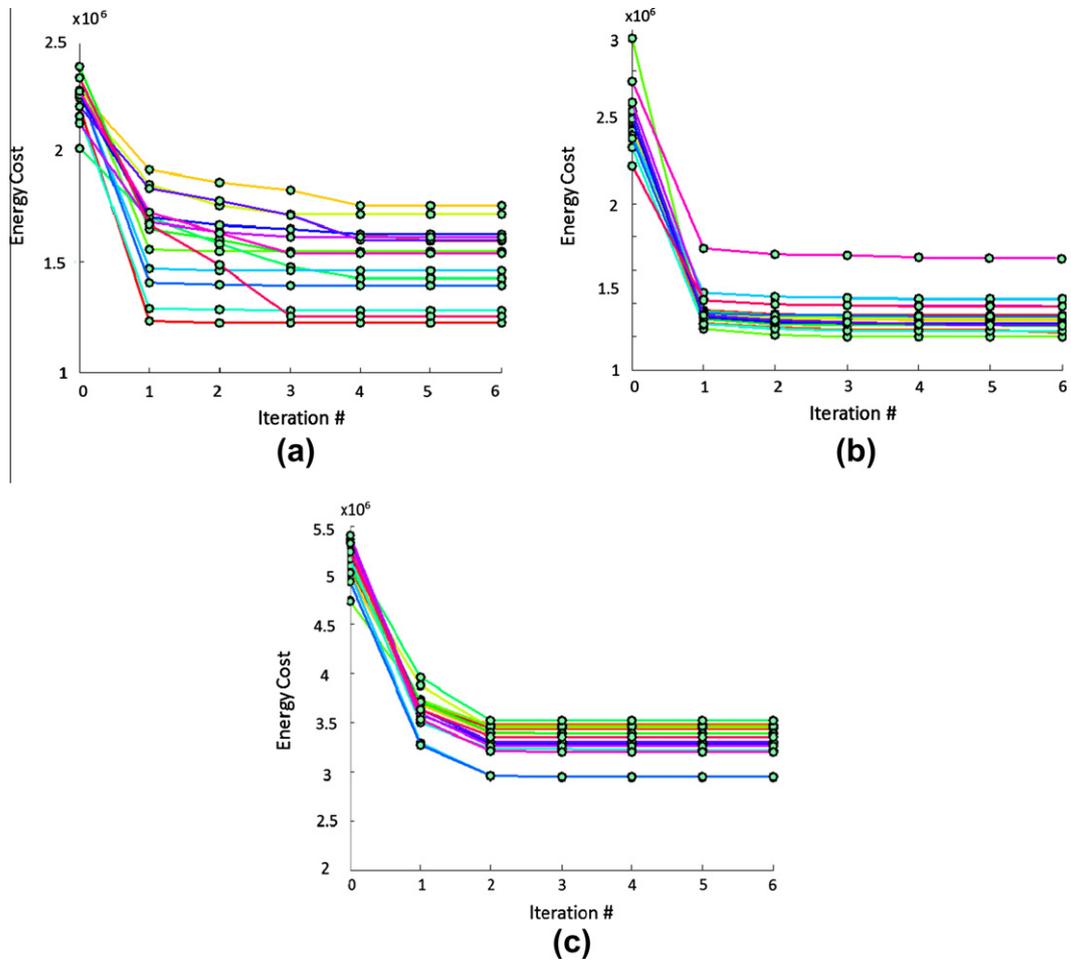


Fig. 9. The variation of energy of the proposed method on three data sets as a function of iteration number: (a) abdominal data set; (b) chest data set; and (c) foot MRI data set. There are 15 testing cases in each data set, One line represents one case. The convergence behavior on all 45 data sets is shown.

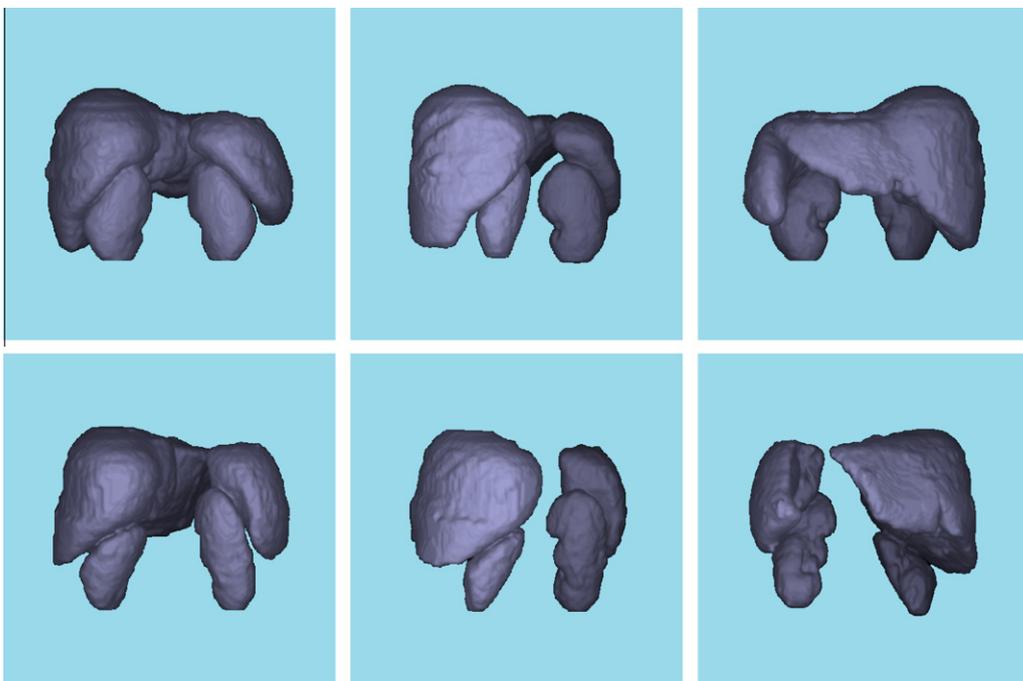


Fig. 10. Three different views of 3D delineation results from two different patient abdominal CT data sets. Objects shown are the liver, spleen, and the two kidneys.

indicate that significant differences could not be discerned ($p > 0.39$) between the two methods in terms of accuracy (RR size, TPVF, FPVF and Dist2 TB). However, GC-ASM achieves better efficiency on computation time (T_c) than GCKPCA ($p < 0.05$).

3.3.3. Analysis of the Convergence of GC-ASM

Many *pl*, *SM*, and hybrid methods in segmentation are iterative in nature and are based on searching and optimization. For most such methods, it is difficult to carry out a theoretical analysis of convergence and/or to guarantee the achievement of a globally optimal solution. This is mainly due to the inability to precisely model and predict the subject-to-subject variations in factors such as object size, shape, layout, intensity appearance, and object relationships. Obviously a theoretical guarantee of the convergence of an algorithm is desirable since this would promise a stable behavior for the method within the space analyzed by the theory. However, any optimum, local or global, does not necessarily assure the best segmentation result. In most situations, including ours, we have only an empirical recourse. Therefore, we conducted the experiments to study the convergence behavior of the proposed method on all three data sets. The results, presented in Fig. 9, show that the proposed GC-ASM algorithm generally converged within four iterations on all test images in all three data sets. The optimization converged mostly after a couple of iterations, which may be due to the fact that the images used in this paper are naturally roughly aligned.

4. Concluding remarks

In this paper, we have proposed a synergistic combination of the image-based GC and model-based ASM methods. The proposed cost function effectively integrates the ASM shape information into the GC paradigm. The automatic object recognition strategy employed by GC-ASM is based on the minimum total GC cut cost, which utilizes combined prior shape and specific image information. The proposed method was evaluated on routine clinical chest CT, abdominal CT, and foot MRI data sets. The evaluated results indicate the following: (1) An overall delineation accuracy of TPVF > 96%, FPVF < 0.6% can be achieved via the GC-ASM method for different objects, modalities, and body regions. (2) GC-ASM improves over ASM in its accuracy and precision to search region. (3) It requires far fewer landmarks (about 1/3) than ASM. (4) GC-ASM achieves full automation compared to the graph cut method. The latter requires seed specification which is commonly done by manual interaction. (5) One disadvantage of GC-ASM is its increased computational expense owing to the iterative nature of the *IGC-ASM* algorithm.

The results demonstrate that it is feasible to explicitly bring prior statistical shape information into the GC framework. This makes object recognition in GC automatic and improves delineation over ASM while maintaining or improving GC's delineation performance, especially in regions in the image where boundary information is missing. It means that a good delineation can help in recognition and a perfect recognition can make delineation most accurate. This is the spirit of the synergy established between ASM and GC by GC-ASM. The proposed method was also compared to two other hybrid methods, MOASM [67] and GCKPCA [56]. The comparison shows that GC-ASM and MOASM have comparable performances while GC-ASM improves over GCKPCA mainly in computational efficiency.

We chose a larger number of landmarks for ASM to try to bring its accuracy on par with that of GC-ASM. If we reduce these numbers from those indicated in Table 4, ASM's performance will deteriorate.

In this paper, we utilized multiple objects in the model. This can considerably facilitate recognition and in turn delineation accuracy compared to using single objects in the. This brings up an issue, specific to the GC approach, of the unavailability of a globally optimal min cut solution for simultaneously segmenting multiple objects. For multiple objects, the α -expansion method [24] can find segmentations within a known factor of the global optimum. However, it is worthy to note that this factor can be huge.

The proposed method can be easily extended to 3D segmentation once a 3D model creation facility is available. Fig. 10 shows segmentation results by GC-ASM for liver, spleen and the two kidneys based on the 3D model created from abdominal CT image data in the work of [68]. On an Intel Pentium IV PC with a 3.4 GHz CPU, the average delineation time for all four 3D objects is about 39 s.

Acknowledgment

The authors are grateful to Dr. Bruce Hirsch for the MRI foot data sets.

References

- [1] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis, and Machine Vision, Brooks/Cole, 1999.
- [2] D. Pham, C. Xu, J. Prince, Current methods in medical image segmentation, Annu. Rev. Biomed. Eng. 2 (2000) 315–337.
- [3] W. Doyle, Operations useful for similarity-invariant pattern recognition, J. ACM 9 (1962) 259–267.
- [4] U. Montanari, On the optimal detection of curves in noisy pictures, Commun. ACM 14 (1971) 335–345.
- [5] H. Liu, Two- and Three-dimensional boundary detection, Comput. Graph. Image Process. 6 (1977) 123–134.
- [6] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, Int. J. Comput. Vis. 1 (1987) 321–331.
- [7] L. Cohen, I. Cohen, Finite element methods for active contour models and balloons for 2D and 3D images, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 1131–1147.
- [8] J. Liang, T. McInerney, D. Terzopoulos, United snakes, Med. Image Anal. 10 (2) (2006) 215–233.
- [9] A. Falcao, J.K. Udupa, S. Samarasekera, S. Sharma, B. Hirsch, R. Lotufo, User-steered image segmentation paradigms: live wire and live lane, Graph. Models Image Process. 60 (1998) 233–260.
- [10] A.X. Falcao, J.K. Udupa, A 3D generalization of user-steered live wire segmentation, Med. Image Anal. 4 (2000) 389–402.
- [11] A. Schenk, G. Prause, H. Peitgen, Local cost computation for efficient segmentation of 3D objects with live wire, in: Proc. SPIE 4322 (2001) 1357–1364.
- [12] R. Malladi, J. Sethian, B. Vemuri, Shape modeling with front propagation: a level set approach, IEEE Trans. Pattern Anal. Mach. Intell. 17 (1995) 158–175.
- [13] X. Han, C. Xu, J.L. Prince, A topology preserving level set method for geometric deformable models, IEEE Trans. Pattern Anal. Mach. Intell. 25 (7) (2003) 755–768.
- [14] E.D. Angelini, Y. Jin, A.F. Laine, State-of-the-art of level set methods in segmentation and registration of medical imaging modalities, in: Suri Jasjit, David L. Wilson, Swamy Laximinarayan (Eds.), Handbook of Medical Image Analysis: Advanced Segmentation and Registration Models, Kluwer Academic Publishers, New York, NY, 2004.
- [15] J. Prewitt, M. Mendelsohn, The analysis of cell images, Annals of the New York Academy of Science 128 (1966) 1035–1053.
- [16] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man, Cyber. 9 (1979) 62–66.
- [17] P. Saha, J.K. Udupa, Optimum image thresholding via class uncertainty and region homogeneity, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2001) 689–706.
- [18] S. Zucker, Region growing: childhood and adolescence, Comput. Graph. Image Process. 5 (1976) 382–399.
- [19] Y. Chang, X. Li, Adaptive image region-growing, IEEE Trans. Image Process. 3 (1994) 868–872.
- [20] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second ed., John Wiley & Sons, New York, 2003.
- [21] J. Bezdek, L. Hall, L. Clarke, Review of MR image segmentation techniques using pattern recognition, Med. Phys. 20 (1993) 1033–1048.
- [22] D. Pham, Spatial models for fuzzy clustering, Comput. Vis. Image Underst. 84 (2001) 285–297.
- [23] Z. Wu, R. Leahy, An optimal graph theoretic approach to data clustering: theory and its application to image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 1101–1113.
- [24] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2001) 1222–1239.

- [25] Y. Boykov, G. Funka-Lea, Graph cuts and efficient N-D image segmentation, *Int. J. Comput. Vis.* 70 (2) (2006) 109–131.
- [26] J.K. Udupa, S. Samarasekera, Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation, *Graph. Models Image Process.* 58 (1996) 246–261.
- [27] J.K. Udupa, P. Saha, R. Lotufo, Relative fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002) 1485–1500.
- [28] K.C. Ciesielski, J.K. Udupa, P.K. Saha, Y. Zhuge, Iterative relative fuzzy connectedness for multiple objects with multiple seeds, *Comput. Vis. Image Underst.* 107 (3) (2007) 160–182.
- [29] H. Derin, H. Elliott, Modeling and segmentation of noisy and textured images using Gibbs random field, *IEEE Trans. Pattern Anal. Mach. Intell.* 9 (1987) 39–55.
- [30] J. Zhang, The mean field theory in EM procedures for blind Markov random field image restoration, *IEEE Trans. Image Process.* 2 (1993) 27–40.
- [31] S. Beucher, C. Lantuejoul, Use of watersheds in contour detection, in: *Proceedings of the International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, 1979, pp. 17–21.
- [32] S. Beucher, The watershed transformation applied to image segmentation, In: *10th Pfeifferkorn Conference on Signal and Image Processing in Microscopy and Microanalysis*, 1992, pp. 299–314.
- [33] J. Park, J. Keller, Snakes on the watershed, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2001) 1201–1205.
- [34] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Commun. Pure Appl. Math.* 42 (1989) 577–685.
- [35] L.A. Vese, T.F. Chan, A multiphase level set framework for image segmentation using the Mumford and Shah model, *Int. J. Comput. Vis.* 50 (3) (2002) 271–293.
- [36] T. Cootes, C. Taylor, D. Cooper, Active shape models - their training and application, *Comput. Vis. Image Underst.* 61 (1995) 38–59.
- [37] T. Cootes, G. Edwards, C. Taylor, Active appearance models, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2001) 681–685.
- [38] S.C. Mitchell, B.P.F. Lelieveldt, R.J. van der Geest, H.G. Bosch, J.H.C. Reiber, M. Sonka, Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images, *IEEE Trans. Med. Imaging* 20 (5) (2001) 415–423.
- [39] S.M. Pizer, G. Gerig, S.C. Joshi, S.R. Aylward, Multiscale medial shape-based analysis of image objects, *Proc. IEEE* 91 (10) (2003) 1670–1679.
- [40] M. Styner, G. Gerig, Automatic and robust computation of 3D medial models incorporating object variability, *Int. J. Comput. Vis.* 55 (2–3) (2003) 107–122.
- [41] C.R. Rao, H. Toutenburg, A. Fieger, C. Heumann, T. Nittner, S. Scheid, *Linear Models: Least Squares and Alternatives*, Springer Series in Statistics, 1999.
- [42] G. Christensen, R. Rabbitt, M. Miller, 3-D brain mapping using a deformable neuroanatomy, *Phys. Med. Biol.* 39 (1994) 609–618.
- [43] P. Thompson, A. Toga, Detection, visualization and animation of abnormal anatomic structures with a probabilistic brain atlas based on random vector field transformations, *Med. Image Anal.* 1 (1997) 271–294.
- [44] X. Fan, J. Yang, L. Cheng, A novel segmentation method for MR brain images based on fuzzy connectedness and FCM, *Lect. Notes Comput. Sci.* 3613 (2005) 505–513.
- [45] F. Admasu, S. Al-Zubi, K. Toennies, N. Bodammer, H. Hinrichs, Segmentation of multiple sclerosis lesions from MR brain images using the principles of fuzzy-connectedness and artificial neural networks, in: *International Conference on Image Processing, ICIP 2003 Proceedings*, volume 2 on page(s) II-1081–1084 volume 3, 2003.
- [46] H.T. Nguyen, M. Worring, R. Boomgaard, Watersnakes: energy-driven watershed segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (3) (2003) 330–342.
- [47] J. Yao, D. Chen, Live level set: a hybrid method of livewire and level set for medical image segmentation, *Med. Phys.* 35 (9) (2008) 4112–4120.
- [48] J. Liu, J.K. Udupa, Oriented active shape models, *IEEE Trans. Med. Imaging* 28 (4) (2009) 571–584.
- [49] J. Hanssegard, S. Urheim, K. Lunde, S.I. Rabben, Constrained active appearance models for segmentation of triplane echocardiograms, *IEEE Trans. Med. Imaging* 26 (10) (2007) 1391–1400.
- [50] Y. Zhou, J. Bai, Atlas-based fuzzy connectedness segmentation and intensity nonuniformity correction applied to brain MRI, *IEEE Trans. Biomed. Eng.* 54 (2007) 122–129.
- [51] M. Rousson, N. Paragios, Prior knowledge, level set representations and visual grouping, *Int. J. Comput. Vis.* 76 (2008) 231–243.
- [52] A. Besbes, N. Komodakis, G. Lings, N. Paragios, Shape priors and discrete mrfs for knowledge-based segmentation. in: *IEEE Conference in Computer Vision and Pattern Recognition (CVPR'09)*.
- [53] D. Freedman, T. Zhang, Interactive Graph Cut Based Segmentation With Shape Priors, *CVPR*, 2005, pp. 755–762.
- [54] A. Ayvaci, D. Freedman, Joint Segmentation-Registration of Organs Using Geometric Models, *EMBS*, 2007, pp. 5251–5254.
- [55] M. Kumar, P. Torr, A. Zisserman, "ObjCut", in *Computer Vision and Pattern Recognition*, 2005, pp. 18–23.
- [56] J. Malcolm, Y. Rathi, A. Tannenbaum, Graph Cut Segmentation with Nonlinear Shape Priors, *ICIP07*, IV, 2007, pp. 365–368.
- [57] C. Allene, J.Y. Audibert, M. Couprie, R. Keriven, Some links between extremum spanning forests, watersheds and min-cuts, *Image Vis. Comput.* 28 (10) (2010) 1460–1471.
- [58] C. Couprie, L. Grady, L. Najman, H. Talbot, Power watershed: a unifying graph-based optimization framework, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (7) (2011) 1384–1399.
- [59] P. Kohli, J. Rihan, M. Bray, P.H.S. Torr, Simultaneous segmentation and pose estimation of humans using dynamic graph cuts, *Int. J. Comput. Vision* 79 (3) (2008) 285–298.
- [60] N. Vu, B.S. Manjunath, Shape Prior Segmentation of Multiple Objects with Graph Cuts, *CVPR 2008* (2008) 1–8.
- [61] X.J. Chen, J.K. Udupa, D. Torigian, A. Alavi, GC-ASM: Synergistic Integration of Active Shape Modeling and Graph-Cut Methods, *SPIE Medical Imaging*, 2009, Paper 7259-11.
- [62] V. Kolmogorov, R. Zabih, What energy functions can be minimized via graph cuts?, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (2) (2004) 147–159.
- [63] E. Dahlhaus, D.S. Johnson, C.H. Papadimitriou, P.D. Seymour, M. Yannakakis, The complexity of multiterminal cuts, *SIAM J. Comput.* 23 (4) (1994) 864–894.
- [64] C.R. Maurer Jr., R.S. Qi, V. Raghavan, A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (2) (2003) 265–270.
- [65] J.A. Nymman, *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. Springer Publishing, ISBN 0-387-24348-8, 2005.
- [66] J.K. Udupa, V.R. Leblanc, Y. Zhuge, C. Imielinska, H. Schmidt, L.M. Currie, B.E. Hirsch, J. Woodburn, A framework for evaluating image segmentation algorithms, *Comput. Med. Imaging Graph.* 30 (2) (2006) 75–87.
- [67] Xinjian. Chen, Jayaram.K. Udupa, Abass. Alavi, Drew.A. Torigian, Automatic anatomy recognition via multiobject oriented active shape models, *Med. Phys.* 37 (12) (2010) 6390–6401.
- [68] Bağıcı, U., Udupa, J.K. and Chen, X.J., "3D Scale Based Multi-Object Recognition in Hierarchical Framework", *Proc. SPIE*, paper 7263–149 (2010).
- [69] Slabaugh, Greg, Unal, Gozde, Graph Cut Segmentation Using an Elliptical Shape Prior, *ICIP 2005*, pp. 1222–1225.
- [70] Zhu-Jacquot Jie, R. Zabih, Graph Cuts Segmentation with Statistical Shape Priors for Medical Images, *SITIS'07*, pp. 631–635.
- [71] J. Zhu-Jacquot, R. Zabih, Segmentation of the Left Ventricle in Cardiac MR Images Using Graph Cuts with Parametric Shape Priors, *ICASSP 2008*, pp. 521–524.
- [72] T. Saito, J.I. Toriwaki, New algorithms for Euclidean distance transformation of an n -dimensional digitized picture with applications, *Pattern Recog.* 27 (1994) 1551–1565.
- [73] T. Hirata, Unified linear-time algorithm for computing distance maps, *Inform. Proc. Lett.* 58 (3) (1996) 129–133.
- [74] Krzysztof C. Ciesielski, Xinjian Chen, Jayaram K. Udupa, George J. Grevera, Linear time algorithms for exact distance transform, *J. Math. Imaging Vision* 39 (3) (2010) 193–209.
- [75] H. Ishikawa, Exact optimization for markov random fields with convex priors, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (10) (2003) 1333–1336.
- [76] W.J. Krzanowski, *Principles of Multivariate Analysis*, Oxford University Press, 1988.
- [77] Bo Xiang, Chaohui Wang, Jean-François Deux, Alain Rahmouni, Nikos Paragios: Tagged cardiac MR image segmentation using boundary and regional-support and graph-based deformable priors, *ISBI 2011*, pp. 1706–1711.
- [78] Bo Xiang, Chaohui Wang, Jean-François Deux, Alain Rahmouni, Nikos Paragios: 3D cardiac Segmentation with Pose-Invariant Higher-Order MRFs, *ISBI 2012*, pp. 1425–1428.