Synthesis of intensity gradient and texture information for efficient three-dimensional segmentation of medical volumes

Sreenath Rao Vantaram
Eli Saber
Sohail A. Dianat
Yang Hu

Synthesis of intensity gradient and texture information for efficient three-dimensional segmentation of medical volumes

Sreenath Rao Vantaram,^{a,*} Eli Saber,^{a,b} Sohail A. Dianat,^{a,b} and Yang Hu^a

^aRochester Institute of Technology, Chester F. Carlson Center for Imaging Science, 1 Lomb Memorial Drive, Rochester, New York 14623, United States

^bRochester Institute of Technology, Department of Electrical and Microelectronic Engineering, 1 Lomb Memorial Drive, Rochester, New York 14623, United States

Abstract. We propose a framework that efficiently employs intensity, gradient, and textural features for three-dimensional (3-D) segmentation of medical (MRI/CT) volumes. Our methodology commences by determining the magnitude of intensity variations across the input volume using a 3-D gradient detection scheme. The resultant gradient volume is utilized in a dynamic volume growing/formation process that is initiated in voxel locations with small gradient magnitudes and is concluded at sites with large gradient magnitudes, yielding a map comprising an initial set of partitions (or subvolumes). This partition map is combined with an entropy-based texture descriptor along with intensity and gradient attributes in a multivariate analysis-based volume merging procedure that fuses subvolumes with similar characteristics to yield a final/refined segmentation output. Additionally, a semiautomated version of the aforesaid algorithm that allows a user to interactively segment a desired subvolume of interest as opposed to the entire volume is also discussed. Our approach was tested on several MRI and CT datasets and the results show favorable performance in comparison to the state-of-the-art ITK-SNAP technique. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.2.2.024003]

Keywords: medical image segmentation; 3-D gradient detection; volume growing; multivariate analysis; volumetric segmentation.

Paper 14160R received Dec. 1, 2014; accepted for publication Apr. 16, 2015; published online May 8, 2015.

1 Introduction

Volumetric segmentation from a medical imaging viewpoint is defined as the process of segregating all volumetric pixels (or voxels) in a medical volume into distinct subvolumes (also known as segments or groups or partitions) exhibiting homogeneous or nearly homogeneous attributes. This fundamental data analysis procedure is employed as a preprocessing step for enhancing the performance of numerous higher level operations such as three-dimensional (3-D) reconstruction, modeling, visualization and navigation, by providing a pertinent representation of the input volume. The increase in the number of medical applications has necessitated the development of algorithms that effectively perform segmentation in a reasonable timeframe.

Many automated/semiautomated methodologies^{1–3} have been developed to confront the segmentation problem for magnetic resonance imaging (MRI) and computed tomography (CT) volumes in 3-D. Approaches involving surface/contour detection clustering, thresholding, region growing, watershed-based, active contours/surfaces, texture-based, and multiresolution statistics, as well as model/atlas-based methods have had wide spread use for volumetric segmentation of medical data.

Early approaches developed for 3-D segmentation in the medical realm were mostly targeted at MRI volumes. Bomans et al.⁴ proposed an approach for segmentation of MRI data based on a 3-D extension of the Marr–Hildreth edge detection scheme.

The fundamental principle of their approach was that the zero crossings of the aforesaid edge detector in 3-D have close correspondences to the anatomical surfaces inherent in data. In their work, Yan and Karp⁵ devised a method based on the work in Ref. 6 for segmentation of MRI volumes of the brain into K disparate tissue types using an adaptive K -means clustering algorithm, wherein each tissue was modeled by a Gibbs random fields. Furthermore, Choi et al.⁷ used a 3-D Markov random fields model for volumetric object extraction, which was demonstrated on MRI's volumes of the knee joint. Justice et al.⁸ developed a semiautomated seed region growing approach demonstrated for MR brain segmentation. Though the method is simple and easy to implement, it suffers from the drawback that all volumetric elements are processed sequentially using a specific sorted order. Sijbers et al.⁹ propounded a two-step procedure for segmentation of MR data. Initially, the entire volume was segmented using the watershed transform followed by a volume merging process based on the minimum description length principle. Hastreiter and Ertl¹⁰ exploited the concept of “Intelligent Scissors” which enables propagation of contours detected in two-dimensional (2-D) to adjacent slices in 3-D, to be applied with a methodology for brain segmentation in MRI volumes.

Over the past decade, the advancement of CT and ultrasound technology for medical imaging has led to the development of several 3-D segmentation algorithms for these modalities, while still being compatible with MRI data. Yezzi et al.¹¹ came up with an image feature metrics driven parametric active contour model

*Address all correspondence to: Sreenath Rao Vantaram, E-mail: sreenath.rit@gmail.com

for segmentation of MRI, CT, and ultrasound imagery. The work was novel in the sense that it unified geometric curve evolution formalisms with traditional parametric energy models to yield the final segmentation outcome. Baillard et al.¹² designed an energy-based geometric active surface evolution model using a level set framework for robust segmentation of anatomical structures in volumetric MRI and ultrasound data. Two major advantages of their algorithm were: (1) its potential for segmenting complex anatomical structures without extensive *a priori* information and (2) its generic nature that required minimal parametric tuning. Tek et al.¹³ proposed a method for segmentation of 3-D structures in CT and MRI volumes based on ray propagation by mean shift analysis that incorporates a smoothness constraint. Freedman et al.¹⁴ founded a model-based segmentation algorithm grounded on probability distributions of photometric variables for deformable objects in CT volumes. On the other hand, Ding et al.¹⁵ established a single 2-D atlas-guided segmentation algorithm for 3-D segmentation of CT volumes to overcome problems of complexity and control germane to 3-D atlases. A prominent semiautomated method for efficient segmentation of anatomical structures was proposed by Yushkevich et al.¹⁶ based on the notions of geodesic active contours and region completion. Recently, Shen et al.¹⁷ proposed a 3-D segmentation procedure applicable to CT and MRI modalities using a novel volumetric deformable model also known as the active volume model (AVM), which is a self-constrained generative object model that does not require any manual training and focuses on accurate modeling of the attributes associated with the volumetric object of interest. The work in Ref. 17 was enhanced by Shen et al. in Ref. 18, where a multiple-surface AVM was proposed to simultaneously segment coupled medical objects. Olowoyeye et al.¹⁹ pioneered a medical (CT and MRI) volume segmentation methodology using its textural properties derived via a 3-D Gabor filter bank whose number depended on the size of the volume. The algorithm commences by convolving the input volume with the aforesaid filter bank. Subsequently, the obtained filter responses were clustered to arrive at the final segmentation output. Campadelli et al.²⁰ developed a gray-level based segmentation framework for abdominal organs from CT scans based on a multiplanar fast marching method. Medical volume segmentation using multiresolution analysis via the 3-D discrete wavelet and ridgelet transforms integrated with hidden Markov models (HMMs) was first seen in the work of AlZu'bi and Amira.²¹ In a similar context using HMMs, Huang et al.²² established a 3-D segmentation algorithm utilizing a rotationally invariant 3-D region-based HMM. Meanwhile, an emerging number of machine learning processes focusing on feature selection and learning methods are engaged as well. Islam et al.²³ proposed a stochastic model to extract brain tumor features based on multifractional Brownian motion and extended the AdaBoost algorithm to segment the brain tumor in an MR image. Yaqub et al.²⁴ employed a random forests framework for 3-D volumetric segmentation by emphasizing “strong” features and neglecting “weak” ones, coupled with a weighting tree techniques in the forest to improve the segmentation accuracy for MRI.

Besides the aforementioned approaches, several segmentation methods for specific abdominal components (organs as well as vertebral structure) have been developed for CT volumes. Hu et al.²⁵ developed a method for accurate segmentation of the lungs from CT imagery using optimal gray level thresholding and morphological operations. Kang et al.²⁶ instituted

a multistep approach for precise segmentation of skeletal structures using 3-D region growing and boundary refinement. Ecabert et al.²⁷ established an effective model-based approach for 3-D segmentation of the heart from CT images using the generalized Hough transform. In their work, Ji et al.²⁸ utilized a learning-based algorithm for liver CT segmentation by using an autocontext model to generate multiple auto-context model classifiers from various atlases, with the final segmentation results obtained by a multiclassifier fusion technique. Li et al.²⁹ employed a probabilistic shape model with a global intensity distribution for CT segmentation of liver, spleen and kidneys.

In this paper, we propose a novel multifeature 3-D volumetric segmentation algorithm that efficiently synthesizes the intensity (or radiodensity in the case of CT), gradient, and textural cues inherent in medical imagery. At the start of the partitioning process, a 3-D gradient detection step is utilized to divulge intensity variations across the volume. Next, voxels with small gradient magnitudes (signifying gradually varying intensity) are grouped together and individually labeled to identify a set of seeds that initialize the formation of various partitions. The obtained seeds are grown into volumetric portions of higher gradient magnitudes accompanied by the addition of new seeds in sections that remain unsegmented as the algorithm progresses. The aforesaid procedure generates an initial partition map that may or may not comprise meaningful groups with reference to the content in the input volume. Consequently, the resultant partition map is refined by integrating it with an entropy-based texture descriptor along with intensity and gradient features in a statistical 3-D volume merging procedure which highlights group differences involving multiple variates, and subsequently fusing the ones with high similarity to achieve the final segmentation. Our approach was entirely implemented in MATLAB® and its performance was qualitatively and quantitatively evaluated on various MRI and CT datasets. The novelty of the proposed segmentation methodology relative to our prior work in the color domain³⁰ is threefold. First, our 3-D volume growing process: (1) concurrently places emphasis on the homogeneous and heterogeneous characteristics of a volume (through intensity and gradient attributes), (2) does not entail manual selection of seeds, (3) abstains from pixel-by-pixel agglomeration, (4) avoids processing order dependency of various volumetric zones, and (5) incorporates a parallel design (simultaneously grows multiple portions of the input volume). Second, our 3-D volume merging approach facilitates a framework that fuses distinct features to define meaningful volumetric segments even in the presence of diverse texture disparities. Finally, we provide a generic mathematical framework for effective gradient detection in 3-D data as well as a semiautomated architecture of the aforementioned algorithm that allows a user to selectively partition a desired subvolume of interest (SVOI) within the input volume.

The remainder of this paper is organized as follows. Section 2 furnishes a detailed illustration of the proposed algorithm. Section 3 discusses a semiautomatic version of the proposed segmentation framework. Experimental results and quantitative evaluations are discussed in Sec. 4, whereas conclusions are drawn in Sec. 5.

2 Proposed 3-D Segmentation Methodology

An overview of our four module 3-D segmentation scheme is shown in Fig. 1. The first module performs a 3-D gradient

Fig. 1 Overview of the proposed three-dimensional (3-D) segmentation algorithm.

estimation technique on the input volume. The second module is responsible for the formation of an initial set of partitions via a 3-D volume growing scheme. The third module is a texture characterization module which discriminates different textures through an entropy-based descriptor. Finally, the algorithm concludes its operation in the fourth module which performs a 3-D volume merging procedure to arrive at the final segmentation result. The following subsections describe each of the aforementioned modules in detail.

2.1 Module 1: 3-D Vector Field Gradient Estimation

In order to ensure that the proposed segmentation algorithm yields results are consistent with the intensity variations across the volume, we formulate a 3-D gradient detection scheme based on the work in Ref. 31 which guides the directions in which the volume growing process transpires. In what follows, let $V_{d \rightarrow 3}$ be a 3-D vector field with dimensions $K \times L \times M$ with d attributes, where $a_t \in \{1, 2, \dots, d\}$ denotes the attribute index. Moreover, let x, y, z represent voxel coordinates, such that $1 \leq x \leq K$, $1 \leq y \leq L$, and $1 \leq z \leq M$. Consequently, a volumetric gradient matrix (J) also known as the 3-D Jacobian, at voxel location (x, y, z) can be defined as

$$J(x, y, z) = \begin{bmatrix} \frac{\partial V_1(x, y, z)}{\partial x} & \frac{\partial V_1(x, y, z)}{\partial y} & \frac{\partial V_1(x, y, z)}{\partial z} \\ \vdots & \vdots & \vdots \\ \frac{\partial V_d(x, y, z)}{\partial x} & \frac{\partial V_d(x, y, z)}{\partial y} & \frac{\partial V_d(x, y, z)}{\partial z} \end{bmatrix}. \quad (1)$$

The gradient magnitude is acquired by first constructing an outer product matrix $J^T J$, given by

$$J^T J(x, y, z) = \begin{bmatrix} a(x, y, z) & d(x, y, z) & e(x, y, z) \\ d(x, y, z) & b(x, y, z) & f(x, y, z) \\ e(x, y, z) & f(x, y, z) & c(x, y, z) \end{bmatrix}, \quad (2)$$

where

$$a(x, y, z) = \sum_{a_t=1}^d \left[\frac{\partial V_{a_t}(x, y, z)}{\partial x} \right]^2, \quad (3)$$

$$b(x, y, z) = \sum_{a_t=1}^d \left[\frac{\partial V_{a_t}(x, y, z)}{\partial y} \right]^2, \quad (4)$$

$$c(x, y, z) = \sum_{a_t=1}^d \left[\frac{\partial V_{a_t}(x, y, z)}{\partial z} \right]^2, \quad (5)$$

$$d(x, y, z) = \sum_{a_t=1}^d \left[\frac{\partial V_{a_t}(x, y, z)}{\partial x} \right] \left[\frac{\partial V_{a_t}(x, y, z)}{\partial y} \right], \quad (6)$$

$$e(x, y, z) = \sum_{a_t=1}^d \left[\frac{\partial V_{a_t}(x, y, z)}{\partial x} \right] \left[\frac{\partial V_{a_t}(x, y, z)}{\partial z} \right], \quad (7)$$

$$f(x, y, z) = \sum_{a_t=1}^d \left[\frac{\partial V_{a_t}(x, y, z)}{\partial y} \right] \left[\frac{\partial V_{a_t}(x, y, z)}{\partial z} \right]. \quad (8)$$

Following this, the magnitude of the gradient (G) at voxel location (x, y, z) is computed by finding the square root of the largest eigenvalue of $J^T J(x, y, z)$, given by

$$G(x, y, z) = \sqrt{\arg \max_i \{\lambda_i [J^T J(x, y, z)]\}}, \quad (9)$$

where $\lambda_i [J^T J(x, y, z)]$ is the i 'th eigenvalue of $J^T J(x, y, z)$.

In Eq. (9), it can be seen that $J^T J(x, y, z)$ cumulatively has three eigenvalues ($\lambda_i, i = 1, 2, 3$) calculated specifically at the voxel location (x, y, z) . Also, since $J^T J(x, y, z)$ is a real, symmetric and positive semidefinite 3×3 matrix for every voxel location, the closed form solutions of its three eigenvalues can be given as in Refs. 32 and 33:

$$\lambda_1(x, y, z) = m(x, y, z) + 2\sqrt{p(x, y, z)} \cos \phi(x, y, z), \quad (10)$$

$$\lambda_2(x, y, z) = m(x, y, z) - \sqrt{p(x, y, z)} [\cos \phi(x, y, z) + \sqrt{3} \sin \phi(x, y, z)], \quad (11)$$

$$\lambda_3(x, y, z) = m(x, y, z) - \sqrt{p(x, y, z)} [\cos \phi(x, y, z) - \sqrt{3} \sin \phi(x, y, z)], \quad (12)$$

where

$$m(x, y, z) = (1/3)[a(x, y, z) + b(x, y, z) + c(x, y, z)] \quad (13)$$

$$p(x, y, z) = (1/6) \sum_{i=1}^3 \sum_{j=1}^3 \{[J^T J(x, y, z) - m(x, y, z)I_3]^2\}_{ij} \quad (14)$$

$$\phi(x, y, z) = (1/3) \tan^{-1} [\sqrt{p(x, y, z)^3 - q(x, y, z)^2} / q(x, y, z)], \quad (15)$$

$$q(x, y, z) = (1/2) \det[J^T J(x, y, z) - m(x, y, z)I_3]. \quad (16)$$

Here I_3 is a 3×3 identity matrix, $p(x, y, z)$ represents the sum of the squares of the elements of $J^T J(x, y, z) - m(x, y, z)I_3$ weighted by a factor of $1/6$, and $\det$ denotes the determinant operator. It can be shown that Eq. (10) is indeed the closed form solution of the largest eigenvalue of $J^T J(x, y, z)$. Figure 2(a) shows two slices from a sample MRI dataset of dimensions $256 \times 256 \times 63$. Figure 2(b) reveals the corresponding

edge information derived by subjecting the entire volume to the above described vector field gradient estimation algorithm. Figure 2(c) portrays the edge information rendered as 3-D iso-surfaces (left) in three gradient ranges (see legend) along the x, y, z axes accompanied by a cross-sectional view (right) of 10 slices (#30 through #40) in the $x - y$ plane. Having established the gradient magnitude (G) at all voxel locations in V , a one-dimensional (1-D) histogram of these gradient values is generated for subsequent processing (see Fig. 1).

2.2 Module 2: Initial Clustering and 3-D Volume Growing

The proposed algorithm delineates various subvolumes within V in two stages, namely initial clustering and volume growing, as shown in Fig. 3.

The initial clustering step commences the partitioning process by identifying voxel-congregations in low gradient “hollow areas” possessing gradually varying intensity. Such a collection of homogeneous voxel-groupings are identified by choosing an

Fig. 2 3-D gradient estimation: (a) original MRI slices (#30 and #40), (b) corresponding edge maps and (c) 3-D view (left) accompanied by a cross-sectional view (right).

initialization threshold (g_{init}) in a low gradient range, accomplished by computing g_{init} as a small percentage (α) of the maximum value (G_{max}) in G (represented as $g_{init} = \alpha G_{max}$). Next, voxels in G that satisfy $G \leq g_{init}$ are made to undergo connected component analysis³⁴ to yield a parent seeds (PSs) map [see Fig. 4(a)] comprised of several uniquely labeled voxel-clusters that serve as starting points for growth. It is imperative to note that in Fig. 4, seeds are randomly color coded for distinction. Furthermore, to avoid managing numerous small seeds in homogeneous areas, a minimum seed volume (MSV) criterion is utilized to restrict the volumetric extent of PSs, as depicted in Fig. 4(b). MSV defines the minimum number of voxels that every seed should mandatorily contain. This measure is set (empirically at five voxels) such that it is small enough to ensure that reasonably fine entities in the volume are captured as the algorithm progresses. Additionally, to ascertain proper subvolume formation, PSs aimed at capturing low gradient locations of the volume are restricted to a large constraint of ~ 50 times MSV, whereas this criterion is gradually reduced for the addition of new seeds at subsequent higher gradient densities at which finer details in the volume manifest themselves. Generation of PSs concludes the initial clustering stage, which is succeeded by a 3-D iterative volume growing stage.

Volume growing is defined as a technique which starts from a single voxel or a set of individual voxels (termed seeds), and based on *a-priori* user-defined similarity rule, iteratively accumulates voxels around each of them. The growth of a volumetric entity ceases when no voxels satisfying the aforesaid similarity principle are found, resulting in compact 3-D partitions. Most

Fig. 3 Flowchart of the proposed volume growing method.

Fig. 4 Initial clustering: (a) initial parent seeds (PSs) map using $g_{\text{init}} = 2.5\% \{G_{\text{max}}\}$ and (b) spatially constrained PSs map using $50 \times \text{MSV} = 250$ voxels. Volume growing: (c) child seeds (CSs) map, (d) parent seed surfaces ($\text{PS}_{\text{surfaces}}$), (e) adjacent CSs, (f) partition map obtained after parent-child similarities has been processed, (g) partition map obtained at the end of a few iterations of the growth process, (h) new seeds nonadjacent to existent PSs, (i) PSs for next stage of volume growing and (j) final volume growth map.

conventional volume growing approaches suffer from some or all of the following short comings: (1) they require manual selection of seeds, (2) they perform voxel-by-voxel agglomerations, (3) they have a sequential architecture (handle only one component of the volume at a time), and (4) they tend to have processing order dependency. Conversely, we propose a volume growing approach that simultaneously lays emphasis on the homogeneous and heterogeneous characteristics of voxels, using intensity and gradient attributes. Our growth routine commences at voxels with small gradient magnitudes (intensity homogeneous sections) and terminates at locations with large gradient magnitudes (abrupt intensity transitions) with no dependency on the order in which they are handled.

Moreover, at various stages of processing, our growing scheme is integrated with an active seed addition mechanism that automatically identifies new seeds in unsegmented portions of V , consequently facilitating parallel growth of multiple adjacent and/or nonadjacent subvolumes. The volume growing process proceeds to grow PSs by first identifying the next higher gradient magnitude, $g_h > g_{init}$, in the histogram of G . Following this, voxels whose gradient values are less than or equal to g_h and that are not tagged as PSs, are subjected to connected component labeling to yield a map comprising of new seeds [Fig. 4(c)], called child seeds (CSs). At this stage, only the CSs adjacent to existent PSs are operated on, identified by distinguishing the ones that share voxels with the peripheral 3-D surfaces of PSs. Parent seed surfaces ($PS_{surfaces}$) are morphologically derived [Fig. 4(d)] as a binary volume by subtracting PSs from their dilated equivalents, represented as

$$PS(x, y, z)_{surfaces} = \begin{cases} 1 & \forall (x, y, z) \text{ where } \{(PSs \oplus \psi) - PSs\} > 0 \\ 0 & \forall (x, y, z) \text{ where } \{(PSs \oplus \psi) - PSs\} = 0 \end{cases} \quad (17)$$

In Eq. (17), $PSs \oplus \psi$ represents the 3-D dilation of the PSs map with a uniform (all 1's) $3 \times 3 \times 3$ cubic structuring element ψ . Following this, the labels of CSs bearing adjacency with PSs are obtained [Fig. 4(e)] by multiplying the $PS_{surfaces}$ and CSs maps on a per-voxel basis. Furthermore, to avoid computational burden, adjacent CSs whose sizes are smaller than the MSV criterion are directly assigned to their associated parent labels, whereas the larger ones are spectrally compared to their parents. Parent-child similarity is assessed by computing the Euclidean distance ($\Delta E_{PSs-CSs}$) between their mean intensity values (m_{PSs}) and (m_{CSs}), denoted as

$$\Delta E_{PSs-CSs} = \sqrt{(m_{PSs} - m_{CSs})^2}. \quad (18)$$

To this effect, the adjacent CSs (larger than the MSV measure) whose mean intensity values are found to be comparable to their parents are merged with them, whereas the rest are discarded, as shown in Fig. 3. Specifically, the aforementioned parent-child similarity is inferred by assessing whether the obtained $\Delta E_{PSs-CSs}$ values are less than a maximum allowed distance ΔE_{max} (set equal to 1). This culminates in a single iteration of volume growing at a gradient threshold g_h , resulting in the gradual growth of PSs [see Fig. 4(f) relative to Fig. 4(b)]. Additionally, as the aforementioned growing procedure advances iteratively into higher gradient densities affecting the parallel growth of all PSs, volumetric compartments that are nonadjacent

to them remain unsegmented in the partition map, as portrayed in Fig. 4(g). Therefore, at specific processing intervals, our growth mechanism is interleaved with an active seed addition step to facilitate the addition of new seeds in nonadjacent expanses to the existent PSs. These intervals, termed as volume growth intervals, are derived by dividing the entire gradient histogram into M components demarcated using $M - 1$ gradient values, such that each component spans an approximately equal area (or number of voxels). Figure 5(a) shows the 1-D gradient histogram of an MRI volume previously displayed in Fig. 2(a), fragmented into 10 volume growth intervals (distinguished using red lines) each spanning an equal ($\sim 10\%$) area. Additionally, Fig. 5(b) portrays a zoomed portion of the histogram [boxed in green in Fig. 5(a)] for a clear view of the various growth intervals. The aforementioned choice of the number of intervals was found to optimally control (i.e., not being too many or too few) the number of seeds being added to the partition map while accounting for the exponential nature of the gradient histogram. Thus, on conclusion of the growth process in an arbitrary interval (when g_h is equal to its upper bound), voxel-clusters in expanses that are not adjacent to existing seeds and are relatively larger than the MSV criterion [Fig. 4(h)] are added to the region map. The resulting agglomeration of seeds becomes the PSs map for the next interval of volume growing, as seen in Figs. 3 and 4(i). This interlaced mechanism involving volume growing and active seed addition culminates when the gradient threshold that is equivalent to the highest magnitude G_{max} is dealt with to arrive at the final volume growth map [Fig. 4(j)].

2.3 Module 3: 3-D Entropy-Based Texture Modeling

Contrasting intensity fluctuations ranging from multiple shades of intensity to repeating patterns of local intensity variations, pertinent to organs, soft-tissues, bones, air, and liquid substances are abundantly present in medical volumes. These fluctuations pose significant problems in achieving the segmentation objective, often resulting in over-segmentation and misinterpretation or perceptual ambiguity of boundaries/surfaces surrounding them. Such volumetric zones of intensity variations are referred to as textures and may contain structured or stochastic patterns/distributions. To mitigate this problem, the proposed algorithm was furnished with a texture modeling component (see Fig. 1), which characterizes different textures via a 3-D entropy-based descriptor.

A popular hypothesis that stems from the field of information theory³⁵ is that the presence of information can be probabilistically modeled and that the information manifested by a random event is inversely proportional to its probability of occurrence. Thus, if $X = \{x_1, x_2, \dots, x_J\}$ is a random set of intensity values present in a 3-D scalar field such as a CT/MRI volume, and $\{P(x_1), P(x_2), \dots, P(x_J)\}$ are their corresponding probabilities of occurrences, then an arbitrary intensity $\{x_i\}$ is said to divulge:

$$I(x_i) = \log_2 \frac{1}{P(x_i)} = -\log_2 P(x_i) \quad (19)$$

binary units or bits of information (since the base of the logarithm is 2). Furthermore, as per the law of large numbers, an intensity level $\{x_i\}$ will on average will exist $kP(x_i)$ times in a volume comprising of k voxels. Consequently, the total information (I) in these k voxels, whose intensities are modeled as a discrete random variable X , is given by

Fig. 5 (a) Gradient histogram of a sample MRI dataset [shown in Fig. 2(a)] divided into 10 volume growth intervals each spanning equal area and (b) portion of the gradient histogram boxed in green (displayed for viewing ease).

$$I(X) = -kP(x_1)\log_2 P(x_1) \dots -kP(x_J)\log_2 P(x_J). \quad (20)$$

Equation (20) algebraically simplifies to

$$I(X) = -k \sum_{i=1}^J P(x_i)\log_2 P(x_i). \quad (21)$$

Therefore, the average information evident per voxel is

$$H(X) = - \sum_{i=1}^J P(x_i)\log_2 P(x_i). \quad (22)$$

The measure $H(X)$ also symbolizes the degree of randomness present in the volume and is popularly known as entropy. The entropy calculation in Eq. (22) which is defined for a single random variable (3-D scalar field) can be extended to multiple

random variables $X_1, X_2, \dots, X_d$ (3-D vector field with d attributes) by computing the joint entropy, defined as

$$H(X_1, X_2, \dots, X_d) = - \sum_i \sum_j \dots \sum_k P(x_{1i}, x_{2j}, \dots, x_{dk}) \times \log_2 P(x_{1i}, x_{2j}, \dots, x_{dk}). \quad (23)$$

Thus, for a 3-D input MRI/CT scalar field, utilizing Eq. (22), an entropy calculation is done locally in a $3 \times 3 \times 3$ neighborhood around every voxel location to arrive at a descriptor that distinguishes various textures within the input volume. Additionally, since medical image data typically contain high bit-depths (12 to 16 bits) of information per voxel, the resultant probability distributions of various intensities tend to be quite large and sparsely populated. To this effect, the input volume is quantized to low bit-depths or a few gray levels (typically 8 bits) before being employed for texture characterization to reduce the computational load. The above mentioned quantization step can also be employed in the case of 3-D vector fields to avoid the computation of the joint probability distribution of intensities across attributes for entropy calculation, as seen in Eq. (23).

2.4 Module 4: 3-D Volume Merging Using Analysis of Variance

The set of subvolumes formed at the culmination of the growing process is typically an oversegmented result with reference to the composition of the input volume [see Figs. 2(a) and 4(j)]. This is due to the fact that these partitions are assembled across various gradient densities in a process that, although accounting for the intensity similarity of voxels, is oblivious to the abundant presence of texture disparities among them. Hence, the last module in our segmentation framework involves a 3-D volume merging routine that unifies subvolumes with similar characteristics in a multivariate intensity-texture space, to arrive at the final/refined set of segments. The aforesaid functionality is accomplished through a statistically motivated procedure called the multivariate analysis of variance (MANOVA),³⁶ which helps to construct a canonical variate space to emphasize between-group differences, and ultimately scrutinizes the possibility of relating various subvolume pairs to a single factor that enables them to be combined together. The following discussion provides a generic overview of the MANOVA method.

In what follows, let the data pertinent to all subvolumes be structured in the form of an $n \times p$ matrix, in which n samples (voxels) are partitioned into g groups, while each sample is associated with p attributes (e.g., intensity and textural features) $x_1, x_2, \dots, x_p$. The eventual goal of the MANOVA procedure is to define a single optimal direction in the p -dimensional space that efficiently highlights differences between multivariate populations. It is imperative to observe that any arbitrary direction in p -dimensions can be obtained as a weighted aggregate of particular vectors $a^T = (a_1, a_2, \dots, a_p)$ and along this direction, every p -variate observation can be projected to a univariate observation using a transformation $y_i = a^T x_i$. Additionally, since the data comprises of n samples divided into g groups, the obtained univariate observations can be represented based on their sample ($j = 1 \dots n_j$) and group ($i = 1 \dots g$) index as y_{ij} denoting a y value for the j 'th sample in the i 'th group. To analyze whether y_{ij} 's in a given direction divulge any significant differences among various groups, the total sum-of-squares of

y_{ij} 's is defined in terms of its sum-of-squares between (SSB)-groups and sum-of-squares within (SSW)-groups' components as

$$\text{SSB}(a) = \sum_{i=1}^g n_i (\bar{y}_i - \bar{y})^2, \quad (24)$$

where

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^g \sum_{j=1}^{n_i} y_{ij} = \frac{1}{n} \sum_{i=1}^g n_i \bar{y}_i$$

and

$$\text{SSW}(a) = \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2. \quad (25)$$

In Eqs. (24) and (25), the notation (a) is utilized to emphasize that the SSB and SSW components are directionally dependent or vary with the choice of a . The aforementioned components are used to define a mean square ratio (F) that highlights between-groups differences, given as

$$F = \left\{ \frac{1}{(g-1)} \text{SSB}(a) \right\} / \left\{ \frac{1}{(n-g)} \text{SSW}(a) \right\}. \quad (26)$$

In Eq. (26), observe that the value of F increases as the variability between-groups becomes larger relative to that within-groups. Thus, it can be inferred that the optimal direction that accentuates between-groups differences is the values of $a^T = (a_1, a_2, \dots, a_p)$ that maximize F . The optimal values of a using the original multivariate observations as opposed to their univariate/1-D projections are obtained by defining the multivariate equivalents of the SSB and SSW components as

$$B_0 = \sum_{i=1}^g n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T, \quad (27)$$

$$W_0 = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T, \quad (28)$$

where x_{ij} , $\bar{x}_i$, $\bar{x}$ correspond to y_{ij} , $\bar{y}_i$, $\bar{y}$, whereas B_0 and W_0 are known as the between-groups and within-groups sum-of-squares and -products matrix, respectively. Moreover, since $y_{ij} = a^T x_{ij}$, Eqs. (24) to (26) can be revised as

$$\text{SSB}(a) = a^T B_0 a, \quad (29)$$

$$\text{SSW}(a) = a^T W_0 a, \quad (30)$$

$$F = \left\{ \frac{1}{(g-1)} a^T B_0 a \right\} / \left\{ \frac{1}{(n-g)} a^T W_0 a \right\} = \frac{a^T B a}{a^T W a}. \quad (31)$$

In Eq. (31), $B = B_0/(g-1)$ and $W = W_0/(n-g)$ are known as the between-groups and within-groups covariance matrices. Thus, the choice of a that maximizes the value of

F in Eq. (31) can be derived by evaluating $\partial F / \partial a = 0$, represented as

$$B a - \left(\frac{a^T B a}{a^T W a} \right) W a = 0 \Rightarrow B a - l W a = 0, \quad (32)$$

where $l = \frac{a^T B a}{a^T W a}$.

This can be algebraically simplified to

$$(B - l W) a = 0 \Rightarrow (W^{-1} B - l I) a = 0. \quad (33)$$

From Eq. (32), it can be inferred that l , which represents the ratio of the between-group to within-group variance, is a constant value at the maximum of F and will be significantly greater than unity when there is distinct group separation. Additionally, from Eq. (33), observe that a maximized F value ascertains that l is the largest eigenvalue of $W^{-1} B$ (possessing multiple eigenvalues), while a is the eigenvector corresponding to the largest eigenvalue that determines the appropriate linear combination $y = a^T x$.

So far, the goal has been to determine a single optimal direction to view group-differences. However, a single direction may often render an oversimplification of the multivariate space being analyzed, especially when the dimensionality of the space (p) or the number of groups (g) is large. Since one or both conditions may be true for medical volumes, several directions may be generated using Eq. (33), whose efficiency in delineating groups-differences is dependent on the magnitudes of their corresponding eigenvalue/eigenvector pairs. Hence, if $W^{-1} B$ possesses s nonzero eigenvalues ($l_1 > l_2 > \dots > l_s > 0$) with corresponding eigenvectors $(a_1, a_2, \dots, a_s)$, a set of variates $(y_1, y_2, \dots, y_s)$ known as canonical variates can be obtained that spans a space called a canonical variate space. Equation (33) may now be represented in matrix form as $BA = WAL$, where A is a matrix of all eigenvectors, whereas L is the diagonal matrix of all eigenvalues, with dimensions $(p \times s)$ and $(s \times s)$, respectively.

An appropriate metric to quantify the variability between two random groups of data (i and j) in a canonical variate space is the distance between their corresponding group means $\bar{y}_i = A^T \bar{x}_i$ and $\bar{y}_j = A^T \bar{x}_j$, given as

$$D^2 = [(\bar{y}_i - \bar{y}_j)^T (\bar{y}_i - \bar{y}_j)]. \quad (34)$$

Equation (34) can be rewritten as

$$D^2 = [(\bar{x}_i - \bar{x}_j)^T A A^T (\bar{x}_i - \bar{x}_j)]. \quad (35)$$

The term AA^T that modifies the influence of each variate in the above shown distance computation can be evaluated by considering two arbitrary eigenvalue/eigenvectors pairs (l_i, a_i) and (l_j, a_j) . As per Eq. (33), they satisfy

$$B a_i = l_i W a_i, \quad (36)$$

$$B a_j = l_j W a_j. \quad (37)$$

Premultiplying Eq. (36) by a_j^T and Eq. (37) by a_i^T

$$a_j^T B a_i = l_i a_j^T W a_i \quad (38)$$

$$a_i^T B a_j = l_j a_i^T W a_j. \quad (39)$$

Since $a_j^T B a_i$, $a_i^T B a_j$, $a_j^T W a_i$, and $a_i^T W a_j$ are all scalars, and B , W are symmetric, it can be concluded that $a_j^T B a_i = a_i^T B a_j$, $a_j^T W a_i = a_i^T W a_j$. Using the first condition, we have

$$l_i a_j^T W a_i = l_j a_i^T W a_j. \quad (40)$$

If $l_i \neq l_j$, Eq. (40) holds if-and-only-if $a_j^T W a_i = a_i^T W a_j = 0$, or in matrix form $A^T W A$ is diagonal. Furthermore, when the normalization $A^T W A = I$ is incorporated, canonical variates can be arranged to be uncorrelated and of equal within-group variance. From this relation, $A A^T$ can be derived as

$$A^T W A = I \Rightarrow A A^T W A A^T = A A^T \Rightarrow A A^T = W^{-1}. \quad (41)$$

Equation (35) can now be rewritten as

$$D^2 = [x(\bar{x}_i - \bar{x}_j)^T W^{-1} (\bar{x}_i - \bar{x}_j)]. \quad (42)$$

Thus, a canonical variate space is constructed to highlight differences among groups by maximizing a mean square ratio of the between-group to within-group variability. Additionally, in this space, the Euclidean distance between the group means is equivalent to the Mahalanobis distance in the original data space [as seen from Eqs. (34) and (42)] and is the preferred measure variability between two multivariate populations.

Figure 6 shows the flow-diagram of our MANOVA-driven 3-D volume-merging process that revamps the partition map obtained at the end of the growing stage to a useful segmentation outcome that is consistent with the content of the input volume. Note that the aforementioned volume growth map determines the g groups into which the n voxels are divided. Initially, intensity and texture features are individually converted to $n \times 1$ column vectors and are concatenated with each other to cast the data to a $n \times p$ matrix form (where $p = 2$ attributes). This resultant matrix is evaluated via the MANOVA method that transforms the data to a canonical variate space, optimized to highlight differences among the existent partitions characterized by multiple variates. In this space, the Euclidean distances (ΔE) between all possible subvolume pairs are computed as a measure of similarity between them.

At this point, the merging process is commenced by identifying the subvolume pair with the minimum distance, signifying the maximum similarity. However, instead of fusing only a single subvolume pair per merging iteration, we adopt a “greedy” algorithm for computational efficiency, where the ΔE between the two most similar partitions is gradually increased ($\Delta E + \delta$) until a larger set of “highly similar” subvolume pairs are obtained before merging is performed. In this subset of similar subvolume pairs, partitions are fused with each other from the most similar pair to the least similar one, which culminates a single-volume merging iteration.

Next, the distances are recomputed for all possible pairs in the new/refined set of subvolumes. This process is repeated hierarchically until, in the specified order of priority, either the ΔE value exceeds a predefined threshold ($T = 1$) or the desired number of segments (S set based on complexity of the volume) is obtained, yielding the final 3-D segmentation outcome. While the first criterion ensures automatic acquisition of a practicable number of segments for higher level tasks, the second

Fig. 6 Statistical 3-D volume merging using analysis of variance.

rule ascertains that the proposed algorithm accounts for the complexity of the input data.

3 Semiautomatic 3-D Segmentation Using the Proposed Framework

Although MRI/CT and other medical imaging systems continuously capture tremendous volumes of data, extraction of meaningful diagnostic information from the acquired datasets rests squarely on the shoulders of the end users. For example, radiologists are often interested in tools that facilitate quick and effective 3-D visualization of a desired volumetric entity such as an abdominal organ or vertebral structure, to aid sound medical reasoning and judgment. Consequently, in this section, we discuss a semiautomatic segmentation algorithm derived as a subset of the segmentation framework proposed in Sec. 2 that enables a user to interactively segment a desired SVOI as opposed to the entire volume. An overview of the algorithm consisting of three modules is shown in Fig. 7. The first module is responsible for the SVOI selection procedure. The second module performs 3-D gradient detection across the volume. Finally, the third module performs volume growing to arrive at the final segmentation result of the specified SVOI. The following subsections describe the functionalities of each of the three modules.

3.1 Module 1: SVOI Selection

The algorithm commences interactively, wherein a user (or radiologist) indicates the SVOI by pairwise selecting the extremities of a few line markings (called RECIST measurements) within its cross-section in an arbitrary slice of the input volume. Figure 8(a) shows user-selected start (x_1, y_1, z_1) and end (x_2, y_2, z_1) voxels of a single line marking (l_1) in slice z_1 , inside the left lung of a CT volume. Given two such extremities in a plane/slice z_1 , the general equation of the line traversing them can be obtained as

$$y - y_1 = m(x - x_1) \Rightarrow mx - y + (y_1 - mx_1) = 0, \quad (43)$$

where the slope of the line $m = (y_2 - y_1)/(x_2 - x_1)$. Consequently, a voxel (x_i, y_i, z_1) in the plane z_1 lies on the line represented by Eq. (43) if and only if

$$mx_i - y_i + (y_1 - mx_1) = 0. \quad (44)$$

The set of all voxels satisfying Eq. (44) [shown as green dotted lines in Fig. 8(b)] and constrained within (x_1, y_1, z_1) and (x_2, y_2, z_1) define the line marking l_1 , as seen in Fig. 8(c). Subject to the complexity of the SVOI, such multiple line markings can be utilized to effectively initialize the volume growing procedure.

3.2 Modules 2-3: Gradient Estimation and Volume Growing

Once the SVOI is selected by a user, a 3-D gradient volume (G) is computed from the input data using the scheme described in Sec. 2.1. This is followed by the generation of its histogram, which is employed in the subsequent initial clustering and volume growing steps, utilized to segment the SVOI from the input volume. The initial clustering step instigates formation of the SVOI by identifying a group of voxels possessing similar or

Fig. 7 Overview of the proposed semiautomatic segmentation algorithm.

Fig. 8 Subvolume of interest selection: (a) user-selected start and end voxels; (b) the set of all voxels satisfying Eq. (44); (c) defining the line marking.

gradually varying intensity in the vicinity of the user-specified “line markings” (internal to the SVOI). To ensure that such a set of homogeneous voxels manifesting small gradient magnitudes among them are isolated, an initialization threshold (g_{init}), equivalent to the most commonly prevalent gradient value (given by the mode) in all line markings is chosen. In general, if $l = \{l_1, l_2, \dots, l_N\}$ denotes a set of N user indicated line markings each of length $\{n_1, n_2, \dots, n_N\}$ number of voxels, respectively, and $G_l = \{G_1, G_2, \dots, G_N\}$ represents the set of gradient magnitudes associated with l , where G_i signifies the gradient magnitudes of the voxels that define l_i , then g_{init} is selected such that

$$g_{init} = \begin{cases} \text{mode}(G_l) & \text{if mode}(G_l) \text{ is unimodal} \\ \gamma : \gamma \in \text{mode}(G_l) & \text{if mode}(G_l) \text{ is multimodal} \end{cases} \quad (45)$$

When the mode of G_l is not unique (multimodal), a random choice γ from the acquired multiple mode values is used for g_{init} , as represented in Eq. (45). From this point forward, the 3-D volume growing routine, described in Sec. 2.2, is iteratively carried out over all gradient magnitudes (higher than g_{init}), which culminates when the gradient threshold equivalent to the maximum magnitude in G is processed to arrive at the final 3-D segmentation result of the SVOI. It is important to observe that the active seed addition component of the volume growing process as well as the texture characterization and volume merging modules is not mandated, given that the semiautomatic algorithm is targeted only at a single SVOI uniquely seeded and labeled via the user-defined line markings.

4 Results and Discussions

The results of the proposed algorithm were evaluated on multiple MRI/CT scans provided by the Massachusetts General Hospital (MGH), National Biomedical Imaging Archive (NBIA), and the DataPhysics Research (DPR) Incorporation. Specifically, the unsupervised version of the proposed algorithm (discussed in Sec. 2) was qualitatively and quantitatively evaluated on 20 normal 8-bit MR brain datasets and their manual segmentations provided by the Center for Morphometric Analysis at the MGH.³⁷ The MRI datasets were of spatial dimensions 256×256 while the number of slices varied from 60 to 64. On the other hand, the supervised segmentation methodology (illustrated in Sec. 3) was demonstrated using two different CT studies on three target subvolumes (lungs, spleen, and vertebral structure) provided by the NBIA and DPR. The datasets were of spatial dimensions 512×512 while the number of slices varied from 100 to 150. Additionally, our results were compared

against a state-of-the-art 3-D active contour-based segmentation algorithm developed by Yushkevich et al.¹⁶ and made available as an open source software tool called ITK-SNAP imaging. ITK-SNAP is a user-friendly interface software utilized for the segmentation of anatomical structures with active contour methods. It implements approaches of geodesic active contours by Caselles et al.^{38,39} and region competition by Zhu and Yuille.⁴⁰ Both methods express an evolving contour as a partial differential equation, as shown below:

$$\frac{\partial}{\partial t} C(t, u, v) = F\vec{N}, \quad (46)$$

where C represents the closed surface determined by spatial coordinates u, v and time t . $\vec{N}$ is the unit normal of the surface and F defines both internal and external forces in the normal direction. Caselles proposed that the external forces can be expressed as the following:^{33,34}

$$F = \alpha g_I + \beta_{K_{g_I}} + \gamma(\nabla g_I \vec{N}), \quad (47)$$

where g_I is the speed function and α, β , and γ are weighting parameters for different force contributions. On the other hand, Zhu and Yuille evaluated the external forces with a probabilistic module, which investigated the probability that a voxel belongs to the object of interest or background, as given by Ref. 38:

$$F = \alpha(P_{\text{obj}} - P_{\text{bg}}) + \beta_k. \quad (48)$$

Both methods define the mean curvature of C as the internal force. In our evaluation, the ITK-SNAP algorithm using Zhu and Yuille's approach with default parametric settings was executed utilizing a downloadable file provided by the Penn Image Computing and Science Laboratory (PICS) in the Department of Radiology at the University of Pennsylvania. The results of our algorithm were run from MATLAB® files developed at the Rochester Institute of Technology, with S empirically set equal to 15 segments (for the automatic segmentation framework). Furthermore, our algorithm was run on a machine having an Intel (R) Core (TM) i7 CPU processor, 2.79 GHz with 6 GB RAM.

4.1 Qualitative Evaluation

The results of the proposed algorithm at interim stages of processing as well as in comparison to the ITK-SNAP method are presented in Figs. 9 to 13 for different MRI/CT datasets. Figure 9(a) shows two slices (#25 and 30) from a sample MRI volume named "11_3" (11 and 3 are the patient identification and scan number, respectively) of dimensions $256 \times 256 \times 63$. Figure 9(b) represents the corresponding gradient maps acquired using the 3-D vector field gradient estimation technique discussed in Sec. 2.1. Figure 9(c) shows the initial set of subvolumes acquired at the end of the volume growing process. It is important to note that the output of the proposed algorithm is a unique label assignment for every volumetric entity (in a 3-D matrix) each of which is associated with a randomly generated pseudocolor for visual distinction, as seen in Fig. 9(c). From this figure, it can be seen that the set of subvolumes formed at the conclusion of the growing process is oversegmented due to reasons specified in Sec. 2.4. Figure 9(d) exhibits the outcomes of the 3-D entropy-based texture descriptor computed from the input volume. The output segmentation result

acquired at the end of the MANOVA-based volume merging algorithm is displayed in Fig. 9(e). As mentioned previously, Fig. 9(e) portrays the final segmentation as a pseudocolor map where each color corresponds to an individual segment.

A big challenge with segmenting MRI data is the segregation of gray-matter and white-matter areas of the volume due to their intricate layout in the brain. Gray matter is a major component of the brain that predominantly comprises neuronal cell bodies and is positioned in the peripheral areas of the brain. Conversely, the white-matter section contains myelinated axon tracts and is located in the central portion of the brain engulfed within the gray matter. In order to visually analyze the performance of the proposed approach against the ITK-SNAP algorithm, we extracted the partitions associated with the gray- and white-matter sections of the input volume from the automatically segmented result of the proposed algorithm and rendered them as 3-D entities in Fig. 10(a). In this figure, the subvolume on the left (rendered in red) is the gray matter, whereas the subvolume on the right (reproduced in green) is the white matter.

Figure 10(b) shows the corresponding 3-D views of the gray- and white-matter subvolumes derived using the ITK-SNAP algorithm as described in Ref. 16, whereas Fig. 10(c) portrays the reference/ground-truth segmentations for these areas provided by the MGH, performed manually using the techniques described in Refs. 41 and 42. From Fig. 10(b), it can be observed that the ITK-SNAP algorithm suffers from significant errors in partitioning gray matter, whereas the segmented outcome of the white matter comprises much less detail in comparison to the proposed scheme. The aforementioned phenomenon is observed due to the fact that in the case of geodesic active contour-based procedures such as ITK-SNAP, the obtained partitions are formed at the expense of a few errors and minor image features for efficiently minimizing the cost function being utilized are ignored. Conversely, our method that utilizes a 3-D volume growing procedure, simultaneously laying emphasis on the homogeneous and heterogeneous characteristics of the data via intensity and 3-D gradient information, in conjunction with a 3-D entropy-based texture descriptor in an information fusion mechanism (3-D volume merging), facilitates the formation of volumetrically compact gray and white matter subvolumes, as seen in Fig. 10(a). Moreover, from Fig. 10, it can be observed that the results of the proposed algorithm are visually in much closer proximity to the ground truths than the ITK-SNAP outputs.

Contrary to performing segmentation in a fully automated process as discussed in Figs. 9 and 10 for MRI volumes, Figs. 11 to 13 demonstrate the semiautomated use of our segmentation framework for selective segmentation of three target subvolumes—lungs, spleen, and vertebral structure in CT data—in line with the work flow of modern day radiologists.

Abstraction of the lungs is a relatively complex task due to the presence of numerous nodules contained within them. In Fig. 11(a), on the left is shown an arbitrary slice (#60) obtained from a sample study of dimensions $512 \times 512 \times 104$, publicly available at the NBIA. The right image of Fig. 11(a) depicts two user-defined line markings made in the lungs. Figure 11(b) illustrates the planar view of the segmented area (in red) across two arbitrarily chosen slices (#33 and #60) of the same volume. The aforementioned images illustrate the competence of our method for lung segmentation by avoiding the grouping of large nodules as being part of them. In Fig. 11(c), the segmented lungs in the 3-D using our framework (left) and the ITK-SNAP algorithm (right) are shown. This result illustrates the shortfall of

Fig. 9 (a) Slice #25 and 30 of an MRI volume "11_3", with corresponding (b) gradient maps, (c) volume growth maps, (d) texture maps and (e) final segmentation outcomes.

Fig. 10 Gray-matter (left) and white-matter (right) segmentation results: (a) ITK-SNAP, (b) proposed algorithm and (c) ground-truth segmentations.

the ITK-SNAP algorithm in furnishing well-defined lung regions in the vicinity of numerous structural variations (like the ribs) encompassing them, as well as a tumor resting on top of the left lung. Conversely, these complex surroundings have been efficiently dealt with using the proposed algorithm. The proficiency of the proposed algorithm can also be observed in Fig. 12 where the segmentation outcomes of the spleen acquired using a sample study of dimensions $512 \times 512 \times 136$, was provided by the DPR Inc.

Segmentation of the vertebral structure in the abdomen is a difficult task due to its complex structure. In Fig. 13(a), 16 user initializations (on the right) demarcated within the cross section of the bones in slice #29 of the previously mentioned study used for spleen segmentation are shown. Figure 13(b) furnishes the 2-D view of the segmented abdominal bone area (red) for two slices (#29 and #100) of the volume. Finally, Fig. 13(c) details the 3-D segmentation results using the proposed algorithm on the left and the ITK-SNAP algorithm on the right. The results in Figs. 13(b) and 13(c) demonstrate the potential of the proposed method to segment complex bone structures pertinent

Fig. 11 Target subvolume—lungs: (a) arbitrary CT slice (left) and user-defined line markings in the SVOI (right), (b) 2-D view of the corresponding segmentation using the proposed algorithm across two different slices and (c) 3-D view of the segmentation results using the proposed method (left) and ITK-SNAP (right).

to the ribs, spine, and the pelvic girdle. On the other hand, the ITK-SNAP algorithm, in spite of using several contour initializations across multiple different slices, fails to segment the entire bone structure.

Overall, Figs. 9 to 13 not only indicate our algorithm's efficacy in achieving desirable segmentations with a competitive qualitative performance against a state-of-the-art technique described in Ref. 16, but also demonstrate the generalization potential of the proposed algorithm across different modalities such as MRI/CT for the segmentation task in both an automated as well as a semiautomated architecture (as preferred by radiologists). Additionally, in ITK-SNAP framework, several different parameters (e.g., number of balloons and their size, algorithm iterations, location and number of slices in which balloon should be placed, etc., to name a few) and stringent convergence criteria have to be set for each run of the algorithm, which is not the case with the proposed algorithm, whereas in the automated version, at most S may be changed based on the complexity of the volume.

Fig. 12 Target subvolume—spleen: (a) Arbitrary CT slice (left) and user-defined line markings in the subvolume of interest (SVOI) (right), (b) two-dimensional (2-D) view of the corresponding segmentation using the proposed algorithm across two different slices, and (c) 3-D view of the segmentation results using the proposed method (left) and ITK-SNAP (right).

We take the spleen and the vertebral structure cases for illustrations to demonstrate the segmentation differences between the proposed method and ITK-SNAP in Figs. 14 and 15. The green part in Fig. 14(a) is the segmentation differences between the proposed algorithm and ITK-SNAP, whereas the red part only denotes the results using the proposed algorithm. Figure 14(b) presents the segmentation differences in 2-D of an arbitrary slice 64 and Figs. 14(c) and 14(d) display the segmentation results in slice 64 from the two approaches, respectively. It is observed that the ITK-SNAP generates the scattered small regions, whereas the proposed algorithm shows a better performance for obtaining a reasonable homogenous region.

In Fig. 15(a), the green part again shows the segmentation differences for the vertebral structure and the red region denotes the segmentation provided by ITK-SNAP. According to Figs. 15(c) and 15(d), the segmentation result from ITK-SNAP

Fig. 13 Target subvolume—vertebral structure: (a) arbitrary CT slice (left) and user-defined line markings in the SVOI (right), (b) 2-D view of the corresponding segmentation using the proposed algorithm across two different slices and (c) 3-D view of the segmentation results using the proposed method (left) and ITK-SNAP (right).

appears incomplete, whereas the proposed algorithm generates an integral structure of the bones.

4.2 Quantitative Evaluation

The growth of applications for segmentation in the medical imaging community has resulted in the development of several quantitative performance evaluation methodologies/measures for the obtained outcomes. In his work, Zhang⁴³ provided a comprehensive review of analytical and empirical approaches that have been instituted to perform segmentation evaluation from a quantitative viewpoint. Analytical methods evaluate segmentations by taking into consideration the properties and underlying principles of the algorithms that generated them. In contrast, empirical techniques quantify the segmentation performance of an algorithm indirectly by estimating the goodness and/or conflict between test datasets and their corresponding manually segmented ground truths.

Fig. 14 Spleen: (a) red volumes represent results using the proposed algorithm and combining green and red volumes are the results using ITK-SNAP, (b) 2-D view of the segmentation differences from the proposed method and ITK-SNAP in slice 64, (c) 2-D view of segmentation from the proposed algorithm in slice 64 and (d) 2-D view of segmentation from ITK-SNAP in slice 64.

Fig. 15 Vertebral structure: (a) red volumes represent results using ITK-SNAP and combining green and red volumes are the results using the proposed algorithm, (b) 2-D view of the segmentation differences from the proposed method and ITK-SNAP in slice 88, (c) 2-D view of segmentation from the proposed algorithm in slice 88 and (d) 2-D view of segmentation from ITK-SNAP in slice 88.

Table 1 Gray-matter segmentation accuracies(%).

Patient ID_Scan #	Proposed		UK-SNAP	
	OA	Mis	OA	Ms
100_23	94.43	5.57	89.64	7.36
110_3	95.57	4.43	89.45	7.55
111_2	94.60	5.40	90.58	6.42
112_2	96.30	3.70	91.47	5.53
11_3	95.29	4.71	90.79	6.21
12_3	94.79	5.21	91.16	5.84
13_3	94.86	5.14	90.31	6.69
15_3	89.10	10.90	90.79	6.21
16_3	95.68	4.32	91.11	5.89
17_3	93.99	6.01	91.33	5.67
191_3	96.31	3.69	91.83	5.17
1_24	95.95	4.05	91.03	5.97
202_3	95.50	4.50	91.80	5.20
205_3	96.04	3.96	92.22	4.78
2_4	91.23	8.77	90.78	6.22
4_8	93.10	6.90	90.08	6.92
5_8	93.52	6.48	90.35	6.65
6_10	91.52	8.48	89.77	7.23
7_8	88.13	11.87	90.72	6.28
8_4	95.68	4.32	91.12	5.88
Mean	94.08	5.92	90.82	6.18

Table 2 White-matter segmentation accuracies(%).

Patient ID_Scan #	Proposed		UK-SNAP	
	OA	Mis	OA	Ms
100_23	94.39	2.61	92.83	4.17
110_3	94.07	2.93	92.20	4.80
111_2	93.57	3.43	92.30	4.70
112_2	94.71	2.29	93.17	3.83
11_3	94.12	2.88	92.83	4.17
12_3	93.74	3.26	93.10	3.90
13_3	93.51	3.49	92.67	4.33
15_3	94.24	2.76	92.87	4.13
16_3	94.65	2.35	92.85	4.15
17_3	93.99	3.01	92.88	4.12
191_3	94.87	2.13	93.46	3.54
1_24	94.52	2.48	92.64	4.36
202_3	94.52	2.48	93.25	3.75
205_3	94.63	2.37	93.49	3.51
2_4	94.80	2.20	92.40	4.60
4_8	93.60	3.40	91.23	5.77
5_8	91.01	5.99	92.74	4.26
6_10	94.01	2.99	91.28	5.72
7_8	95.08	1.92	93.37	3.63
8_4	94.93	2.07	92.78	4.22
Mean	94.15	2.85	92.72	4.28

Given the availability of 20 normal MR brain datasets and their corresponding manual segmentations for gray/white-matter areas provided by the MGH, we quantitatively evaluated the segmentation performance of the proposed algorithm (unsupervised framework) against ITK-SNAP in an empirical fashion. More specifically, the overall accuracy—percentage of correctly segmented pixels and error rate—percentage of erroneously segmented pixels for both categories across all 20 datasets were calculated as depicted in Tables 1 and 2. From these tables, it can be seen that for both gray matter as well as white matter, the segmentation outputs acquired from our algorithm have the highest overall percentage accuracy with the lowest error rate, illustrating that the outcomes generated using our approach have better agreement with the reference segmentations.

5 Conclusions

This paper provides an illustration of a novel multiattribute 3-D medical volume segmentation technique in automated and semiautomated frameworks for applications such as 3-D

reconstruction, visualization, and navigation to name a few. The first module of the algorithm performs an effective 3-D gradient computation scheme with the potential of accommodating 3-D vector field data. The resultant gradient is employed to dynamically parse through the data from subvolumes of contiguous pixels that display homogeneous intensity to yield an initial set of partitions. These partitions are refined using a multi-variate statistical analysis process that fuses intensity and texture information modeled via local 3-D neighborhood-based entropy calculations, to arrive at the final segmentation result. As mentioned previously, this work also demonstrates a semiautomated version of the proposed algorithm derived as a subset of its unsupervised counterpart. Our segmentation framework was benchmarked against the method in Ref. 16 with favorable qualitative and quantitative performances over multiple MRI/CT volumes. It is our intention to perform a quantitative evaluation on CT volumes using manually segmented ground truths over a diverse set of studies and target subvolumes. Furthermore, another future direction of research we envision is the integration of

the proposed algorithm with model-based techniques for the segmentation of complex cardiac/aortic anatomical structures.

Acknowledgments

This work was supported by the Chester F. Carlson Center for Imaging Science and the Department of Electrical and Micro-electronic Engineering, Rochester Institute of Technology, Rochester, New York. The authors would like to thank Dr. Paul Yushkevich from the Penn Image Computing and Science Laboratory at the University of Pennsylvania for providing the ITK-SNAP algorithm and Dr. Andrew J. Worth from the Center for Morphometric Analysis at Massachusetts General Hospital, for providing the 20 MRI datasets and their reference segmentations utilized in the quantitative evaluation component of this work.

References

1. D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.* **2**(1), 315–337 (2000).
2. H. Bulu and A. Alpkocak, "Comparison of 3D segmentation algorithms for medical imaging," in *IEEE Int. Symp. on Computer-Based Medical Systems*, Vol. 15, pp. 269–274 (2007).
3. D. J. Withey and Z. J. Koles, "Medical image segmentation: methods and software," in *IEEE Joint Meeting of the 6th Int. Symp. on Noninvasive Functional Source Imaging (NFSI) of the Brain and Heart and the Int. Conf. on Functional Biomedical Imaging (ICFBI)*, pp. 140–143 (2007).
4. M. Bomans et al., "3-D segmentation of MR images of the head for 3-D display," *IEEE Trans. Med. Imaging* **9**(2), 177–183 (1990).
5. M. X. H. Yan and J. S. Karp, "Segmentation of 3D brain MR using an adaptive K-means clustering algorithm," in *IEEE Nuclear Science Symp. and Medical Imaging Conf.*, Vol. 4, pp. 1529–1533 (1994).
6. T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Process.* **40**(4), 901–914 (1992).
7. S. M. Choi et al., "Volumetric object reconstruction using the 3D-MRF model-based segmentation," *IEEE Trans. Med. Imaging* **16**(6), 887–892 (1997).
8. R. K. Justice et al., "Medical image segmentation using 3-D seeded region growing," *Proc. SPIE* **3034**, 900–910 (1997).
9. J. Sijbers et al., "Watershed-based segmentation of 3D MR data for volume quantization," *Magn. Reson. Imaging* **15**(6), 679–688 (1997).
10. P. Hastreiter and T. Ertl, "Fast and interactive 3D-segmentation of medical volume data," in *Proceedings of Image and Multidimensional Digital Signal Processing (IMDSP)*, pp. 41–44 (1998).
11. A. Yezzi, Jr. et al., "A geometric snake model for segmentation of medical imagery," *IEEE Trans. Med. Imaging* **16**(2), 199–209 (1997).
12. C. Baillard, C. Barillot, and P. Bouthemy, "Robust adaptive segmentation of 3D medical images with level sets," Technical Report INRIA (2000).
13. H. Tek et al., "Segmentation of 3D medical structures using robust ray propagation," *Lec. Notes Comput. Sci.* **2488**, 572–579 (2002).
14. D. Freedman et al., "Model-based segmentation of medical imagery by matching distributions," *IEEE Trans. Med. Imaging* **24**(3), 281–292 (2005).
15. F. Ding, W. K. Leow, and S. Wang, "Segmentation of 3D CT volume images using a single 2D atlas," in *Proc. of the 1st Int. Workshop on Computer Vision for Biomedical Image Applications*, pp. 1–10 (2005).
16. P. A. Yushkevich et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage* **31**(3), 1116–1128 (2006).
17. T. Shen et al., "Active volume models for 3D medical image segmentation," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 707–714 (2009).
18. T. Shen and X. Huang, "3D medical image segmentation by multiple-surface active volume models," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, pp. 1059–1066 (2009).
19. A. Olowoyeye, M. Tuceryan, and S. Fang, "Medical volume segmentation using bank of gabor filters," in *24th Annual ACM Symp. on Applied Computing*, pp. 826–830 (2009).
20. P. Campadelli, E. Casiraghi, and S. Pratioli, "A segmentation framework for abdominal organs from CT scans," *Artif. Intell. Med.* **50**(1), 3–11 (2010).
21. S. AlZu'bi and A. Amira, "3D medical volume segmentation using hybrid multiresolution statistical approaches," *Adv. Artif. Intell.* **2010**, 1–15 (2010).
22. A. Huang, R. Abugharbieh, and R. Tam, "A novel rotationally invariant region-based hidden Markov model for efficient 3-D image segmentation," *IEEE Trans. Image Process.* **19**(10), 2737–2748 (2010).
23. A. Islam, S. M. S. Reza, and K. M. Iftikharuddin, "Multifractal texture estimation for detection and segmentation of brain tumors," *IEEE Trans. Biomed. Eng.* **60**(11), 3204–3215 (2013).
24. M. Yaqub et al., "Investigation of the role of feature selection and weighted voting in random forests for 3-D volumetric segmentation," *IEEE Trans. Med. Imaging* **33**(2), 258–271 (2014).
25. S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric x-ray CT images," *IEEE Trans. Med. Imaging* **20**(6), 490–498 (2001).
26. Y. Kang, K. Engelke, and W. A. Kalender, "A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data," *IEEE Trans. Med. Imaging* **22**(5), 586–598 (2003).
27. O. Ecabert et al., "Automatic model-based segmentation of the heart in CT images," *IEEE Trans. Med. Imaging* **27**(9), 1189–1201 (2008).
28. H. Ji et al., "ACM-based automatic liver segmentation from 3-D CT images by combining multiple atlases and improved mean-shift techniques," *IEEE J. Biomed. Health Inform.* **17**(3), 690–698 (2013).
29. C. Li et al., "Joint probabilistic model of shape and intensity for multiple abdominal organ segmentation from volumetric CT images," *IEEE J. Biomed. Health Inform.* **17**(1), 92–102 (2013).
30. L. G. Ugariza et al., "Automatic image segmentation by dynamic region growth and multiresolution merging," *IEEE Trans. Image Process.* **18**(10), 2275–2288 (2009).
31. H. C. Lee and D. R. Cok, "Detecting boundaries in a vector field," *IEEE Trans. Signal Process.* **39**(5), 1181–1194 (1991).
32. O. K. Smith, "Eigenvalues of a symmetric 3×3 matrix," *Commun. Assoc. Comput. Mach.* **4**(4), 168 (1961).
33. D. Zwillinger, *Standard Mathematical Tables*, 26th ed., pp. 9–11, CRC Press, Boca Raton, Florida (1981).
34. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Vol. 1, pp. 28–48, Addison-Wesley, Boston, Massachusetts (1991).
35. T. A. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley-Interscience, Hoboken, New Jersey (1991).
36. W. J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective*, pp. 291–301, Oxford University Press, Oxford, United Kingdom (1988).
37. The Internet Brain Segmentation Repository (IBSR), <http://www.cma.mgh.harvard.edu/ibsr/>.
38. V. Caselles et al., "A geometric model for active contours in imaging processing," *Numer. Math.* **66**(1), 1–31 (1993).
39. V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comput. Vision* **22**(1), 61–79 (1997).
40. S. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(9), 884–900 (1996).
41. D. N. Kennedy, P. A. Filipek, and V. S. Caviness, "Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging," *IEEE Trans. Med. Imaging* **8**(1), 1–7 (1989).
42. P. A. Filipek et al., "The young adult human brain: an MRI-based morphometric analysis," *Cereb. Cortex* **4**(4), 344–360 (1994).
43. Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition* **29**(8), 1335–1346 (1996).

Sreenath Rao Vantaram received his BTech (equivalent BS) degree in electronics and communication engineering from the Jawaharlal Nehru Technological University, Andhra Pradesh, India, in 2006 and his MS and PhD degrees in electrical engineering and imaging science, from the Rochester Institute of Technology in 2009 and 2012, respectively. He is currently a part of Intel Corporation in Oregon. His research interests are computer vision, digital image/video processing, understanding, and analysis.

Eli Saber is a professor in the Electrical Engineering Department at the Rochester Institute of Technology. He received his BS degree in electrical and computer engineering from the University of Buffalo in 1988 and his MS and PhD degrees in the same discipline from the University of Rochester in 1992 and 1996, respectively. His research interests are in the areas of digital image/video processing, including image/video segmentation, object tracking, content-based image/video analysis, and understanding.

Sohail A. Dianat is a professor in the Electrical and Microelectronic Engineering Department at the Imaging Science Center at the Rochester Institute of Technology. He received his BS degree in electrical engineering from the Arya-Mehr University of Technology,

Tehran, Iran, in 1973, and his MS and DSc degrees in electrical engineering from George Washington University in 1977 and 1981, respectively. His current research interests include digital signal/image processing and wireless communication.

Yang Hu received his BS degree in telecommunications from Wuhan University, Wuhan, China, in 2005 and his MS degree in electrical engineering from Rochester Institute of Technology, Rochester, New York in 2012. He is currently working toward his PhD in imaging science at the Chester F. Carlson Center for Imaging Science, at the Rochester Institute of Technology. His research interests include digital image/video processing, computer vision, and medical imaging.