

Optimum Image Thresholding via Class Uncertainty and Region Homogeneity

Punam K. Saha, *Member, IEEE*, and Jayaram K. Udupa, *Senior Member, IEEE*

Abstract—Thresholding is a popular image segmentation method that converts a gray-level image into a binary image. The selection of optimum thresholds has remained a challenge over decades. Besides being a segmentation tool on its own, often it is also a step in many advanced image segmentation techniques in spaces other than the image space. Most of the thresholding methods reported to date are based on histogram analysis using information-theoretic approaches. These methods have not harnessed the information captured in image morphology. Here, we introduce a novel thresholding method that accounts for both intensity-based class uncertainty—a histogram-based property—and region homogeneity—an image morphology-based property. A scale-based formulation is used for region homogeneity computation. At any threshold, intensity-based class uncertainty is computed by fitting a Gaussian to the intensity distribution of each of the two regions segmented at that threshold. The theory of the optimum thresholding method is based on the postulate that objects manifest themselves with fuzzy boundaries in any digital image acquired by an imaging device. The main idea here is to select that threshold at which pixels with high class uncertainty accumulate mostly around object boundaries. To achieve this, a new threshold energy criterion is formulated using class-uncertainty and region homogeneity such that, at any image location, a high energy is created when both class uncertainty and region homogeneity are high or both are low. Finally, the method selects that threshold which corresponds to the minimum overall energy. The method has been compared to a recently-published maximum segmented image information (*MSII*) method. Superiority of the proposed method was observed both qualitatively on clinical medical images as well as quantitatively on 250 realistic phantom images generated by adding different degrees of blurring, noise, and background variation to real objects segmented from clinical images.

Index Terms—Image segmentation, thresholding, uncertainty, scale, homogeneity, threshold energy, segmented image information.

1 INTRODUCTION

TWO- and higher-dimensional images are currently available through sensing devices that operate on a wide range of frequency in the electromagnetic spectrum—from ultrasound to visible light to X- and γ -rays [1]. Defining objects in these image data is fundamental to most computerized image-related applications. It is obvious that defining objects is essential prior to their visualization, manipulation, and analysis. Even operations such as image interpolation and filtering, seemingly unrelated to object definition, can be made more effective with object knowledge. Thus, defining objects in images seems to be vital to any computerized imaging applications. This activity, generally referred to as *image segmentation*, spans over three decades. Pal and Pal [2] reviewed various methods for gray-level image segmentation. Thresholding is undoubtedly one of the most popular segmentation approaches because of its simplicity. However, the automatic selection of a robust, optimum threshold has remained a challenge in image segmentation. Besides being a segmentation tool on its own, thresholding is frequently used as one of the steps in many advanced segmentation methods. In these applications, thresholding is not applied on the original images, but

applied in a space generated by the segmentation method. For example, in fuzzy connectedness segmentation [3], a threshold is applied on the strength of connectedness among image elements to produce a final segmentation. Thus, the methods to determine effective thresholds have wide-spread applications.

Many methods for automatic threshold selection have been reported [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23] over the past four decades. An early review of thresholding methods was reported in [4]. A comparative performance study of global thresholding techniques was reported by Lee et al. [5]. Another comparative analysis of the performance of eleven histogram-based thresholding algorithms was carried out by Glasbely [6]. Among early works on thresholding, Prewitt and Mendelson [7] suggested to select the threshold at the valleys of the histogram, while Doyle [8] advocated the choice of the median. Otsu [9] described a method that maximizes the between-class variance. Ridler and Calvard [10] proposed an iterative method of thresholding (also see [11]). Tsai [12] suggested selecting a threshold at which the resulting binary images have the same first three moments.

Pun [13] described a method that maximizes the upper bound of the a posteriori entropy derived from the histogram. Wong and Sahoo's method [14] determines the optimum threshold by maximizing the a posteriori entropy subject to certain inequality constraints that characterize the uniformity and shape of the segmented regions. Pal and Pal [15] developed another entropy-based method by considering the joint probability distribution of the neighboring pixels which they further modified [16] with a new definition of

• The authors are with the Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Fourth Floor, Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021.
E-mail: {saha, jay}@mipg.upenn.edu.

Manuscript received 15 Mar. 2000; revised 14 Nov. 2000; accepted 7 Apr. 2001.

Recommended for acceptance by S. Sarkar.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 111711.

entropy. The method of Kapur et al. [17] selects the optimum threshold by maximizing the sum of entropies of the segmented regions. A similar approach was reported by Abutaleb [18] which maximizes the 2D entropy. Brink's method [19] maximizes the sum of the entropies computed from two autocorrelation functions of the thresholded image histograms. Li and Lee's method [20] minimizes relative cross entropy or Kullback-Leibler (KL) distance between the probability distribution functions of the original and the thresholded image. Kitler and Illingworth [21] developed a thresholding method that minimizes the notion of a segmentation error derived using an information-theoretic approach, while Dunn et al.'s method [22] is based on a uniform error criterion. Recently, Leung and Lam [23] developed a method that maximizes segmented image information derived using information-theoretic approaches and demonstrated that their method is better than the methods based on minimum error [21] and uniform error [22] criteria.

While the methods of Wong and Sahoo [14] and Pal and Pal [15], [16] incorporate some spatial image information in their methods, others are mostly histogram-based techniques. Although researchers have used spatial image information in several nonthresholding image segmentation methods [25], [26], [27], thresholding is a fundamentally different and simple operation. Further, final segmentations using these methods depend on initial segmentations which are not needed by thresholding. One common trait of all histogram-based approaches is that they do not utilize the considerable amount of information that is captured in the spatial distribution of intensities and in image morphology. It is obvious that, in real-life imaging applications, it is very difficult to select a threshold from the histogram only without seeing the image, while the latter has a clear object morphology. This observation motivated us to develop a method that uses image morphology, as further amplified below.

Object boundaries in any acquired images exhibit fuzziness because of the blurring or the ubiquitous partial voluming effect introduced by the imaging device. As seen from the description above, most thresholding methods utilize some sort of class uncertainty or entropy criteria to dictate the selection of the optimum threshold. It is then clear that, if the optimality criterion is set up properly, then at the optimum threshold, image elements in the vicinity of the boundaries should have a high value of class uncertainty or entropy. This is very difficult to achieve without some object knowledge or without extracting some object information from the image. (This is the classic paradox repeatedly encountered in the development of segmentation methods in the literature: "segmentation is needed for segmentation.") We argue that some simple intensity homogeneity criteria can be utilized to generate rough object knowledge. We postulate that an optimality criterion that combines the previously utilized class uncertainty/entropy criterion with these region homogeneity criteria will provide significantly improved thresholds by making the image elements in the vicinity of boundaries have high values of the optimality function at the optimum threshold.

This paper is organized as follows: In Section 2.1, we describe our notations and definitions. In Section 2.2, we give a general formulation of the class uncertainty principle

of selecting optimum thresholds, which is followed, in Section 2.3, by the formulation of the scale-based region homogeneity principle. The proposed optimum thresholding method utilizing class uncertainty and region homogeneity is next presented in Section 2.5. In Section 3, we discuss the results of applications, first, qualitative comparisons with a recently published thresholding method [23] utilizing medical images, and then a quantitative statistical comparison with the same method utilizing realistic phantoms. Finally, we state our conclusions in Section 4. A preliminary version of this paper was presented at the SPIE conference on Medical Imaging 2000 [24].

2 THEORY AND ALGORITHMS

2.1 Notations and Definitions

Let n -dimensional Euclidean space \mathbb{R}^n be divided into hypercubes by n orthogonal families, each consisting of equally distant parallel planes. Although the method presented here is readily extendable to handle resolution anisotropy of the imaging grid system, for simplicity, we assume that the resolution of the underlying grid is isotropic. We refer to the hypercubes as *spels* (an abbreviation for spatial elements) and assume that the successive planes from the same family are unit distance apart. We determine a coordinate system such that the centroid of each spel has coordinates (c_1, c_2, \dots, c_n) , where c_i s are integers. Thus, we have a one-to-one and onto mapping between the set of spels and Z^n as follows: An n -tuple c in Z^n is mapped to a spel such that c_j for $1 \leq j \leq n$, gives the j th coordinate of the center of the spel represented by c . With this interpretation of spels, we shall think of Z^n itself as the set of all spels in \mathbb{R}^n and shall use the concepts of spels and points denoting their centroids interchangeably.

In this paper, we follow the convention that, for a fuzzy subset (and a fuzzy relation) [28] \mathcal{A} of a reference set X , $\mu_{\mathcal{A}}$ is the membership function of \mathcal{A} in X (resp., in $X \times X$). The range of $\mu_{\mathcal{A}}$ is $[0, 1]$. A fuzzy relation α on Z^n is said to be a *fuzzy spel adjacency* if it is reflexive and symmetric. It is desirable that α be such that its membership value $\mu_{\alpha}(c, d)$ for any $c, d \in Z^n$ is a nonincreasing function of the distance $\|c - d\|$ between c and d , where $\|\cdot\|$ denotes any $L2$ norm in \mathbb{R}^n . For simplicity of computation, the functional form of μ_{α} used for the image processing results presented in this paper is as follows: For any spel c , $\mu_{\alpha}(c, c) = 1$. Further, for any spels $c, d \in Z^n$, $\mu_{\alpha}(c, d) = 1$ if c and d differ in exactly one coordinate by 1. Otherwise, $\mu_{\alpha}(c, d) = 0$. We call the pair (Z^n, α) , where α is any fuzzy spel adjacency, a *fuzzy digital space*. A fuzzy digital space is a concept that characterizes the underlying digital grid system independent of any image-related notions.

A *scene over a fuzzy digital space* (Z^n, α) is a pair $\mathcal{C} = (C, f)$, where $C = \{c \mid -b_j \leq c_j \leq b_j \text{ for some } b \in Z_+^n\}$, Z_+^n is the set of n -tuples of positive integers, f is a function whose domain is C , called the *scene domain*, and whose range is a set of integers $\Gamma = [MIN, MAX]$. To avoid ill-defined conditions, we need to confine the search for optimum threshold within Γ leaving out intensities at the ends of Γ . Accordingly, for scenes with at least four gray levels, we define $\Gamma_- = [MIN + 2, MAX - 1]$. Its need will

become clear in (2.6)-(2.12). All that is presented in this paper generalizes in a straightforward manner to vector-valued scenes, that is, when f is a vector-valued function. However, for the sake of simplicity of description, we will treat f as a scalar-valued function. A scene \mathcal{C} over (Z^n, α) will be called a *binary scene* if $\Gamma = \{0, 1\}$.

2.2 Intensity-Based Class Uncertainty

Intensity-based class uncertainty principles have been used by several researchers in the past for optimum threshold selection [21], [22], [23]. Here, we give its generic formulation as we have used in our proposed new method.

The idea behind intensity-based class uncertainty is to determine the uncertainty of classification of a spel into a certain object class based on its scene intensity and the a priori knowledge of the intensity probability distributions of different classes. Although the framework can be generalized to the problem of determining multiple thresholds to classify spels into multiple object classes, in this paper, we confine ourselves to the two-class problem of separating objects (foreground) from background, but will indicate some extensions in Section 3 for the case of multiple objects. To determine class uncertainty, we assume that we have a priori knowledge of the intensity probability distributions for both the object and the background and also of the probability of a spel belonging to the object. Let F_o and F_b represent the true object and background classes, respectively. Let θ denote the probability of a spel belonging to the object class so that $1 - \theta$ is the probability of a spel belonging to the background class. Let $p_o(g)$ denote the a priori probability that an object spel has the intensity value g . That is,

$$p_o(g) = P(f(c) = g \mid c \in F_o), \quad (2.1)$$

where P represents "probability." Let $p_b(g)$ denote the a priori probability that a background spel has the intensity value g . That is,

$$p_b(g) = P(f(c) = g \mid c \in F_b). \quad (2.2)$$

Let $p(g)$ denote the a priori probability that any spel has the intensity value g . Then,

$$p(g) = \theta p_o(g) + (1 - \theta) p_b(g).$$

The a posteriori probability that a spel having a gray value g belongs to the object class is, by Bayes rule,

$$P(c \in F_o \mid f(c) = g) = \frac{\theta p_o(g)}{p(g)}. \quad (2.3)$$

Similarly, the a posteriori probability that a spel having a gray value g belongs to the background class is given by

$$P(c \in F_b \mid f(c) = g) = \frac{(1 - \theta) p_b(g)}{p(g)}. \quad (2.4)$$

Following Shannon and Weaver [29], after knowing that the intensity value at a spel c is g , the uncertainty of classifying c into the object or background class is the entropy of the two a posteriori probabilities defined in (2.3) and (2.4). We refer to this as *intensity-based class uncertainty* or *class*

uncertainty for short. The *class uncertainty* $H(g)$ at an intensity g is then given by

$$H(g) = -\frac{\theta p_o(g)}{p(g)} \log \frac{\theta p_o(g)}{p(g)} - \frac{(1 - \theta) p_b(g)}{p(g)} \log \frac{(1 - \theta) p_b(g)}{p(g)}. \quad (2.5)$$

Following (2.3)-(2.5), once we know the a priori entities θ , p_o , and p_b , class uncertainty can be computed for every intensity. For any scene $\mathcal{C} = (C, f)$ over (Z^n, α) since the range Γ of f is a finite set of integers, the probabilities p_o and p_b defined in (2.1) and (2.2) are also the probability distributions (density functions) of the underlying discrete random variable g . In past approaches [21], [22], [23], these have been assumed to have Gaussian forms G_{m_o, σ_o} and G_{m_b, σ_b} , and the means and standard deviations m_o , m_b , σ_o , and σ_b have been estimated from the given scene as a function of the threshold as described below.

Suppose $t \in \Gamma_-$ is any threshold intensity. Let $F_{o,t}$ and $F_{b,t}$ denote the set of foreground and background spels induced by the threshold t . That is,

$$F_{o,t} = \{c \mid c \in C \text{ and } f(c) \geq t\}, \quad (2.6)$$

$$\text{and, } F_{b,t} = \{c \mid c \in C \text{ and } f(c) < t\}. \quad (2.7)$$

Let $m_o(t)$ and $m_b(t)$ denote the means of the spel intensities in \mathcal{C} within the sets $F_{o,t}$ and $F_{b,t}$, respectively, let $\sigma_o(t)$ and $\sigma_b(t)$ denote the standard deviations of the spel intensities in \mathcal{C} in these respective regions, and let $\theta(t)$ be defined as

$$\theta(t) = \frac{|F_{o,t}|}{|C|}, \quad (2.8)$$

where $|X|$ denotes the cardinality of the set X . (Note that, $F_{b,t} = \phi$ when $t = MIN$, $\sigma_b(t) = 0$ when $t = MIN + 1$, and $\sigma_o(t) = 0$ when $t = MAX - 1$. This is why we need $t \in \Gamma_-$.) Then, the above probability density functions, denoted with the subscript t to indicate their dependence on the threshold for any $g \in \Gamma$, are given by

$$p_{o,t}(g) = \frac{1}{\sqrt{2\pi} \sigma_o(t)} e^{-\frac{(g - m_o(t))^2}{2\sigma_o(t)^2}}, \text{ for } g \in \Gamma, \quad (2.9)$$

$$p_{b,t}(g) = \frac{1}{\sqrt{2\pi} \sigma_b(t)} e^{-\frac{(g - m_b(t))^2}{2\sigma_b(t)^2}}, \text{ for } g \in \Gamma, \quad (2.10)$$

$$\text{and } p_t(g) = \theta(t) p_{o,t}(g) + (1 - \theta(t)) p_{b,t}(g), \text{ for } g \in \Gamma. \quad (2.11)$$

From (2.5), the class uncertainty as a function of t for $g \in \Gamma$ is given by

$$H_t(g) = -\frac{\theta(t) p_{o,t}(g)}{p_t(g)} \log \frac{\theta(t) p_{o,t}(g)}{p_t(g)} \quad (2.12)$$

$$- \frac{(1 - \theta(t)) p_{b,t}(g)}{p_t(g)} \log \frac{(1 - \theta(t)) p_{b,t}(g)}{p_t(g)}.$$

Thus, (2.6)-(2.12) provide an algorithm for computing class uncertainty $H_t(g)$ for every intensity g in Γ at every threshold t in Γ_- .

2.3 Region Homogeneity

We think of *region homogeneity* (or *homogeneity* for short) as a property of every spel in the given scene $\mathcal{C} = (C, f)$ over

(Z^n, α) . That is, homogeneity μ_τ is a function $\mu_\tau : C \rightarrow [0, 1]$. It depends on spel adjacency α and another fuzzy relation ψ on C called *fuzzy affinity* that was introduced in [3] in connection with the development of fuzzy connectedness principles. The idea behind affinity is to capture the notion of local “hanging-togetherness” of spels. The strength $\mu_\psi(c, d)$ of the relation between any two spels c and d in C is greater when the scene intensities in the vicinity of c and d are more similar. In this paper, we will utilize a scale-based formulation of affinity suggested in [31], [32] for defining ψ , as described in the rest of this section. The functional form of μ_τ we employ is given by, for any $c \in C$,

$$\mu_\tau(c) = \frac{\sum_{d \in C} \mu_\alpha(c, d) \mu_\psi(c, d)}{\sum_{d \in C} \mu_\alpha(c, d)}, \quad (2.13)$$

In words, the homogeneity at a spel c is a weighted average of the affinities with c of spels in C . Typically, since the adjacency between spels that are far apart is 0, the spels that actually matter in (2.13) are those that are in the close vicinity of c .

The fundamental idea behind the scale-based formulation of affinity between two spels c and d is to utilize the concept of a local structure size, or *scale*, at c and d . The *scale* in a scene C at any spel $c \in C$ is defined as the radius $r(c)$ of the largest hyperball centered at c which lies entirely in the same object region defined under an appropriately defined intensity uniformity criterion. Intuitively, it appears that image segmentation is first needed to determine object scale. A simple and effective algorithm has been described in [31] which estimates $r(c)$ at every spel $c \in C$ in any scene C without explicit image segmentation but based on continuity of intensity uniformity. This concept of scale is similar (only in spirit) to the notion of “scale” used in scale-space principles for computer vision [33], but as is readily seen, the associated techniques are fundamentally different (see [31], [32] for further details).

2.3.1 Determination of Scale

A hyperball $B_k(c)$ with center at $c \in C$ and of radius k is defined by

$$B_k(c) = \{e \in C \mid \|c - e\| \leq k\}. \quad (2.14)$$

For a hyperball $B_k(c)$ centered at c and of radius k , we define a fraction, $FO_k(c)$, that indicates the fraction of the set of the spels in the ball boundary whose intensities are sufficiently uniform with that of c , by

$$FO_k(c) = \frac{\sum_{e \in B_k(c) - B_{k-1}(c)} W_\psi(|f(c) - f(e)|)}{|B_k(c) - B_{k-1}(c)|}, \quad (2.15)$$

where W_ψ is a membership function corresponding to the fuzzy proposition “ x is small.” Many choices for this function are available [28]. In this paper, a zero mean unnormalized Gaussian function with standard deviation σ_ψ is used for W_ψ . The algorithm for *object scale estimation* is summarized below.

Algorithm OSE

Input: C , $c \in C$, W_ψ , a fixed threshold t_s .

Output: $r(c)$.

begin

```

set  $k = 1$ ;
while  $FO_k(c) \geq t_s$  do
    set  $k$  to  $k + 1$ ;
endwhile;
set  $r(c)$  to  $k - 1$ ;
output  $r(c)$ ;

```

end

The algorithm iteratively increases the ball radius k by 1, starting from 1, and checks for $FO_k(c)$, the fraction of the object containing c that is contained in the ball boundary. The first time when this fraction falls below the tolerance parameter t_s , we consider that the ball enters into an object region different from that to which c belongs. Following the recommendation in [31], we have used $t_s = 0.85$.

The parameter σ_ψ is determined as follows: Over the entire scene domain C , local intensity differences $|f(c) - f(d)|$ are computed for all possible pairs (c, d) of spels such that $c \neq d$ and $\mu_\alpha(c, d) > 0$. The upper 10 percentile values of these differences are discarded to account for interobject boundaries. The mean (M_h) and the standard deviation (σ_h) are computed over the lower 90 percentile values. These estimates are then used in setting up the value of σ_ψ as follows:

$$\sigma_\psi = M_h + 3\sigma_h. \quad (2.16)$$

The rationale for this choice is that, in a normal distribution, three standard deviations on both sides of the mean cover 99.7 percent of the population.

2.4 Determination of Affinity

In order to determine affinity between two spels $c, d \in C$, two digital hyperballs, one centered at c and the other at d , denoted $B_{cd}(c)$ and $B_{cd}(d)$, both of radius $\min[r(c), r(d)]$, are defined by

$$B_{cd}(c) = \{e \in C \mid \|c - e\| \leq \min[r(c), r(d)]\}, \quad (2.17)$$

$$B_{cd}(d) = \{e \in C \mid \|d - e\| \leq \min[r(c), r(d)]\}. \quad (2.18)$$

These two hyperballs are used as neighborhoods for c and d , respectively. Note that the size of these neighborhoods is defined by local scales $r(c)$ and $r(d)$.

For defining μ_ψ , consider any two spels $c, d \in C$ such that $\mu_\alpha(c, d) > 0$. Consider any spels $e \in B_{cd}(c)$ and $e' \in B_{cd}(d)$ such that they represent the corresponding spels within $B_{cd}(c)$ and $B_{cd}(d)$. In other words, $c - e = d - e'$. We will define two weighted sums $D^+(c, d)$ and $D^-(c, d)$ of intensity differences between the two hyperballs as follows: Let

$$\delta_{cd}^+(e, e') = \begin{cases} f(e) - f(e'), & \text{if } f(e) - f(e') > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (2.19)$$

$$\delta_{cd}^-(e, e') = \begin{cases} f(e') - f(e), & \text{if } f(e) - f(e') < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (2.20)$$

Then,

$$D^+(c, d) = \sum_{\substack{e \in B_{cd}(c) \\ e' \in B_{cd}(d) \\ \text{s.t. } c - e = d - e'}} [1 - W_\psi(\delta_{cd}^+(e, e'))] \omega_{cd}(\|c - e\|), \quad (2.21)$$

$$D^-(c, d) = \sum_{\substack{e \in B_{cd}(c) \\ e' \in B_{cd}(d) \\ \text{s.t. } c - e = d - e'}} [1 - W_\psi(\delta_{cd}^-(e, e'))] \omega_{cd}(\|c - e\|), \quad (2.22)$$

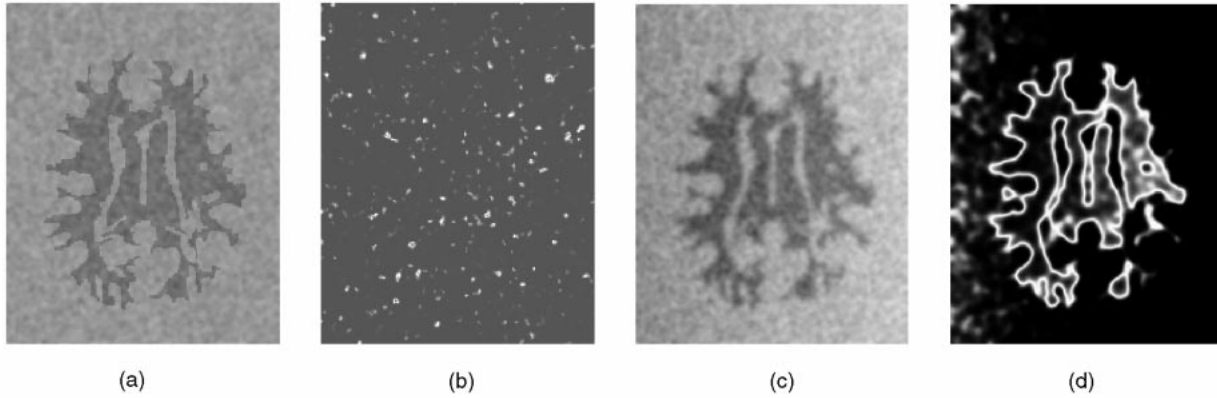


Fig. 1. Illustrations of principles in Postulate A. (a) A 2D phantom scene obtained by segmenting brain WM regions in an MRI slice and then adding a zero-mean Gaussian noise to this segmented result (see Section 3.2). (b) The class uncertainty scene at the theoretical optimum threshold. Brightness at a pixel is proportional to its class uncertainty value. (c) Same as (a) except that a blurring and an intensity ramp across columns were added prior to adding noise. (d) Same as (b) but for (c).

where ω_{cd} is a membership function corresponding to the fuzzy proposition “ x is small.” Similar to W_ψ , a zero-mean Gaussian with a standard deviation $\sigma_{\omega_{cd}} = \min[r(c), r(d)]$ is used for ω_{cd} .

The connection of the above equations to the affinity relation ψ is as follows: There are two types of intensity variations surrounding c and d —intra- and interobject variations. The intraobject component is generally random, and, therefore, is likely to be near zero overall. The interobject component, however, has a direction. It either increases or decreases along the direction given by $c - d$ and is likely to be larger than the intraobject variation. It is reasonable, therefore, to assume that the smaller of $D^+(c, d)$ and $D^-(c, d)$ represents the intraobject component and the other represents the combined effect of the two components. (Note that when the values of δ_{cd}^+ (respectively, δ_{cd}^-) are small, $D^+(c, d)$ (respectively, $D^-(c, d)$) also becomes small.) If there is a slow background component of variation within the small neighborhood considered, this component is unlikely to cause a variation comparable to the interobject component. This strategy leads us to the following functional form for μ_ψ :

$$\mu_\psi(c, d) = 1 - \frac{|D^+(c, d) - D^-(c, d)|}{\sum_{e \in B_{cd}(c)} \omega_{cd}(\|c - e\|)}. \quad (2.23)$$

Note that $|D^+(c, d) - D^-(c, d)|$ represents the degree of local non uniformity of the regions containing c and d . Its value is low when both c and d are inside an (uniform) object region. Its value is high when c and d are in the vicinity of (or across) a boundary. The denominator in (2.23) is a normalization factor. From (2.13) and (2.17)–(2.23), the definition of homogeneity $\mu_\tau(c)$ at any spel c in a given scene is complete. Just to re-emphasize, this definition depends on only one parameter σ_ψ .

2.5 Optimum Threshold Selection

In this section, we describe the proposed thresholding method that takes into account both class uncertainty and region homogeneity. The motivation and the basis for the method come from the following postulate.

Postulate A. *In any scene with fuzzy boundaries, at optimum partitioning of object classes, spels with high class uncertainty appear in the vicinity of object boundaries.*

It is, indeed, difficult to prove or disprove this postulate mathematically due to the nature of its statement. However, an intuitive justification for it can be given as follows: When p_o and p_b are Gaussian distributions, the optimum threshold t_0 is at the point of intersection of the two distributions, i.e., t_0 is such that $p_o(t_0) = p_b(t_0)$. This follows from the pattern recognition theory [34, chapter 2] for the classification of features based on minimum-error-rate. It is clear from (2.12) that scene intensities in the vicinity of t_0 produce high uncertainty at threshold t_0 . Let us first consider a scene with *nonfuzzy* boundaries, as demonstrated in Fig. 1a. The scene in this figure is generated by adding zero-mean Gaussian noise to a scene created by manually segmenting the white matter (WM) region on an MRI slice of a patient’s head. Before adding noise, the gray values in the two regions are set to correspond to the mean gray values in the WM and gray matter (GM) regions, respectively. Note that no blurring is added here. Since the truth is known here, the theoretical best threshold can be determined by an exhaustive search looking for a maximum overlap between the true (original binary) regions and the regions resulting from thresholding. Fig. 1b shows the class uncertainty scene for the scene in Fig. 1a at the best threshold, with high brightness indicating high class uncertainty. As shown in the figure, high uncertainty regions are randomly scattered over the scene and they do not happen to be in the vicinity of object boundaries. But, this example does not contradict the postulate because the hypothesis of the postulate requires that the scene has *fuzzy* boundaries. That is, the intensity changes smoothly (and not like a step function) across interfaces between object regions. Now, consider a scene with fuzzy boundaries as illustrated in Fig. 1c. The scene in Fig. 1c is generated in the same way as that of Fig. 1a except that a blurring and a background intensity ramp across columns were added prior to adding noise. Now, it is readily seen that the pixels in the vicinity of

boundaries have intensities in between the averages of the object intensities and the background intensities and, therefore, have high class uncertainty at a proper threshold. Fig. 1d, the class uncertainty scene at the best threshold, illustrates this fact that pixels with high uncertainty (high intensities) appear in the vicinity of boundaries. Although the postulate as stated is constrained to images with blurred or fuzzy boundaries, it holds good for scenes acquired by most imaging devices since such scenes exhibit fuzzy boundaries due to finite resolution, partial volume effect, and material inhomogeneity.

It is clear from Postulate A that, to determine an optimum threshold which is as close as possible to the theoretical best threshold, some form of region homogeneity information is needed. We may then devise criteria (similar to class uncertainty) that will assume high values when both class uncertainty and homogeneity are low or both are high. Our aim is to devise a threshold energy criterion function E whose minimum will produce the optimum threshold. The functional form of E we have chosen consists of a class uncertainty factor ((2.12)) and a homogeneity factor. In the rest of this section, we describe the formulation of E under these guidelines.

Instead of taking $\mu_\tau(c)$ ((2.13)) directly, we will use the normalized rank $\mu_\varphi(c)$ of these values in the scene to represent homogeneity at c .

$$\mu_\varphi(c) = \frac{LC(\mu_\tau(c))}{LC(1)}, \quad (2.24)$$

where

$$LC(x) = \sum_{y \in \Upsilon \text{ and } y \leq x} L(y), \quad \text{for } x \in [0, 1], \quad (2.25)$$

where $L(y)$ is the number of spels $c \in C$ in the scene with $\mu_\tau(c) = y$ and Υ is the set of all $z \in [0, 1]$ such that there is a spel $c \in C$ for which $\mu_\tau(c) = z$. Note that Υ is a discrete set with finite cardinality. The rank-based normalization of homogeneity expressed by $\mu_\varphi(c)$ makes it less sensitive to the actual choice of the parameter σ_ψ . Finally, we define a *threshold energy* $E(t)$ for any $t \in \Gamma_-$ for any scene $C = (C, f)$ over (Z^n, α) to be

$$E(t) = \sum_{c \in C} H_t(f(c))\mu_\varphi(c) + (1 - H_t(f(c)))(1 - \mu_\varphi(c)). \quad (2.26)$$

When both class uncertainty and homogeneity are high, $E(t)$ is high and the first term in (2.26) dominates. When both these entities are low, $E(t)$ is high and the second term dominates. Analogously, when one of them is low and the other is high, $E(t)$ is low. Thus, $E(t)$ expresses, at threshold t , the total disagreement of the spel values over the scene domain to the underlying intensity distribution and to the homogeneity of the regions as per the principle of Postulate A that object boundaries are associated with high chaos. Our aim is to find a threshold t at which $E(t)$ is minimum. That is, to find

$$t_{MHUE} = \operatorname{argmin}\{E(t) \mid t \in \Gamma_-\}. \quad (2.27)$$

Since $E(t)$ represents energy based on both homogeneity and uncertainty, we refer to this process of determining the optimum threshold as the method of *minimization of homogeneity- and uncertainty-based energy (MHUE)*. We shall carefully compare it qualitatively in Section 3 and objectively with a recently published method [23] that maximizes the histogram-based segmentation information. Before closing this section, we present an example in Fig. 2 to illustrate the *MHUE* method. Fig. 2a shows $E(t)$ as a function of threshold t for the phantom scene displayed in Fig. 1a. Fig. 2b shows the location of the optimum threshold (the global minimum in Fig. 2a) on the scene intensity histogram. Fig. 2c displays the optimally thresholded binary scene. Note that, since the object region is darker, the binary scene is determined by applying (2.7), where t is the optimum threshold.

3 RESULTS AND DISCUSSION

In this section, we describe the results of both qualitative and quantitative experiments to assess the performance of the proposed *MHUE* method compared to a recently published *maximum segmented image information (MSII)* method [23]. This method determines the threshold at which histogram-based segmented image information gain is maximum. We first present a short description of this method.

Prior to segmenting (by thresholding) an image into classes, average uncertainty of classifying a spel c in a scene C into the object and background classes is measured by Shannon and Weaver's equation [29] as

$$H_X = -\theta \log \theta - (1 - \theta) \log(1 - \theta), \quad (3.28)$$

where θ is as defined in Section 2. H_X is called the *initial scene uncertainty*. Let C_o and C_b represent the set of spels in the segmented object and background classes, respectively. Let $p_{o,b}$ denote the probability of a spel belonging to the true object F_o knowing that it belongs to the segmented background class C_b . Three other probabilities, namely, $p_{o,o}$, $p_{b,o}$, and $p_{b,b}$ are defined analogously. (The details of how to estimate these four probabilities at any threshold t are given in [23].) Note that, by definition,

$$p_{o,o} + p_{b,o} = 1, \quad (3.29)$$

$$p_{o,b} + p_{b,b} = 1. \quad (3.30)$$

The average uncertainty H_o of true classification of all segmented object spels is given by

$$H_o = -p_{o,o} \log p_{o,o} - p_{b,o} \log p_{b,o}, \quad (3.31)$$

while for segmented background spels, the average uncertainty H_b is given by

$$H_b = -p_{o,b} \log p_{o,b} - p_{b,b} \log p_{b,b}. \quad (3.32)$$

Therefore, average uncertainty H_S of true classification knowing the thresholded classes is given by

$$H_S = \theta H_o + (1 - \theta) H_b. \quad (3.33)$$

H_S is called *residual uncertainty*. Segmented image information *SII*, defined as the difference between initial scene

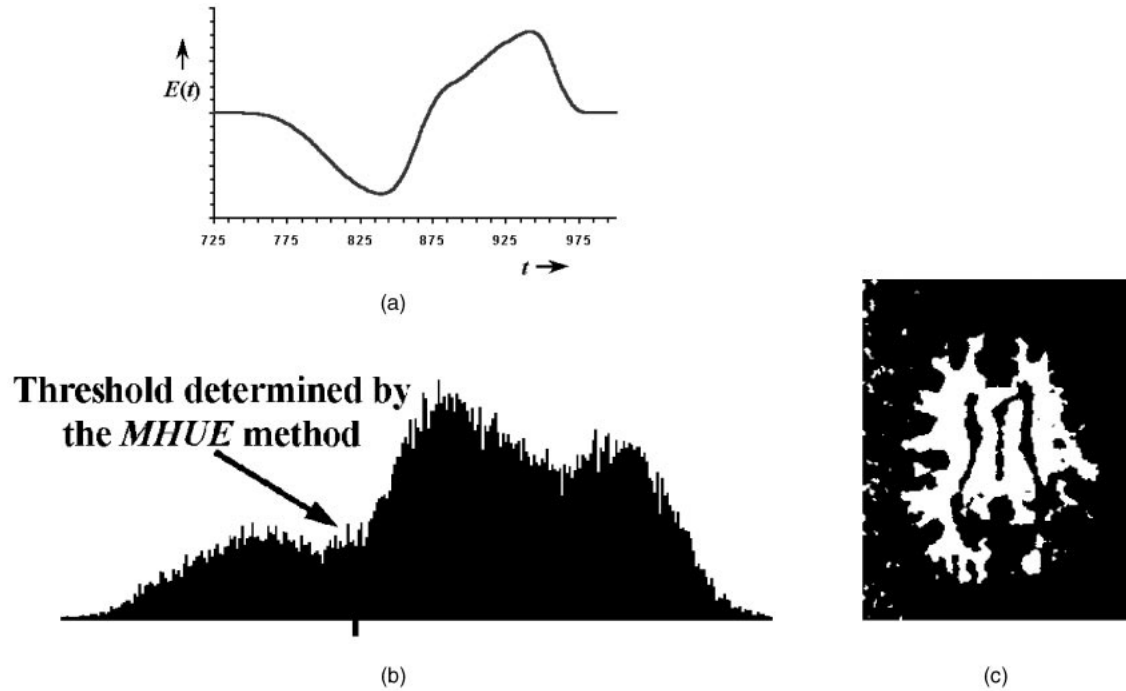


Fig. 2. (a) The distribution of threshold energy $E(t)$ for the phantom scene of Fig. 1a. (b) The location of the optimum threshold is indicated on the scene intensity histogram. (c) The binary scene resulting from thresholding the scene in Fig. 1a at the optimum value indicated in (b).

entropy H_X and the residual uncertainty after segmentation, is

$$SII = H_X - H_S. \quad (3.34)$$

The *MSII* thresholding method selects the threshold associated with the largest *SII*. Fig. 3 illustrates the results

of application of this method to the phantom scene of Fig. 1a. Fig. 3a shows the plot of *SII* as a function of threshold t . Fig. 3b demonstrates the location of the optimum threshold on intensity histogram. Fig. 3c displays the thresholded binary scene. As for Fig. 3, (2.7) is used for determining this binary scene.

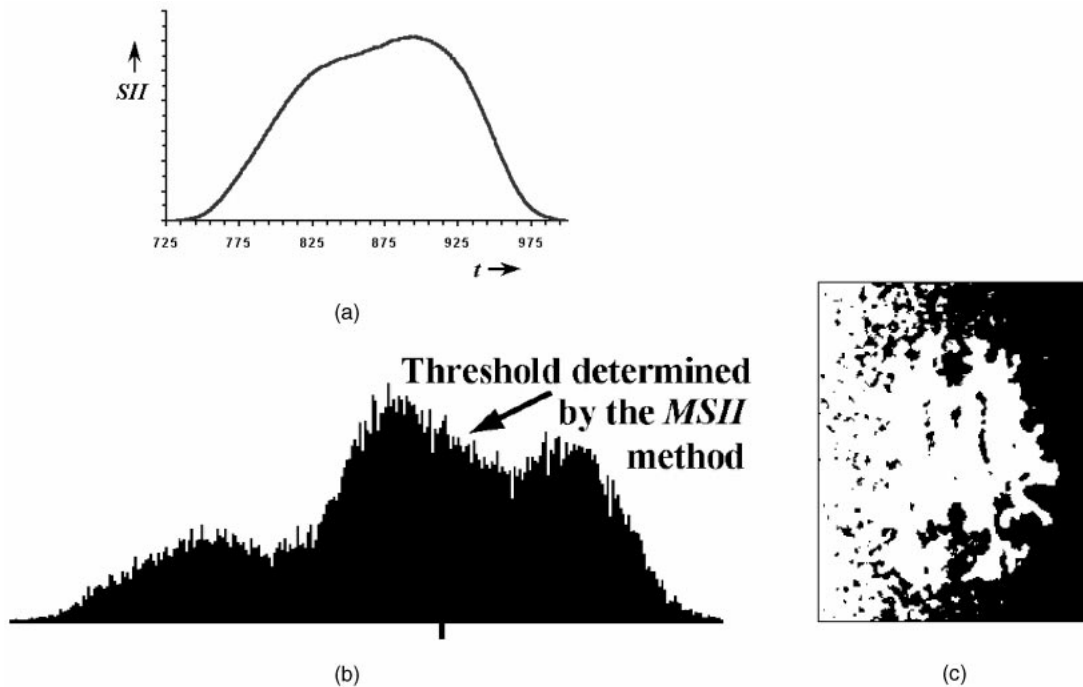


Fig. 3. (a) The distribution of segmented image information (*SII*) for the phantom scene of Fig. 1a. (b) The location of the optimum threshold is indicated on the scene intensity histogram. (c) The binary scene resulting from thresholding the scene in Fig. 1a at the optimum value indicated in (b).

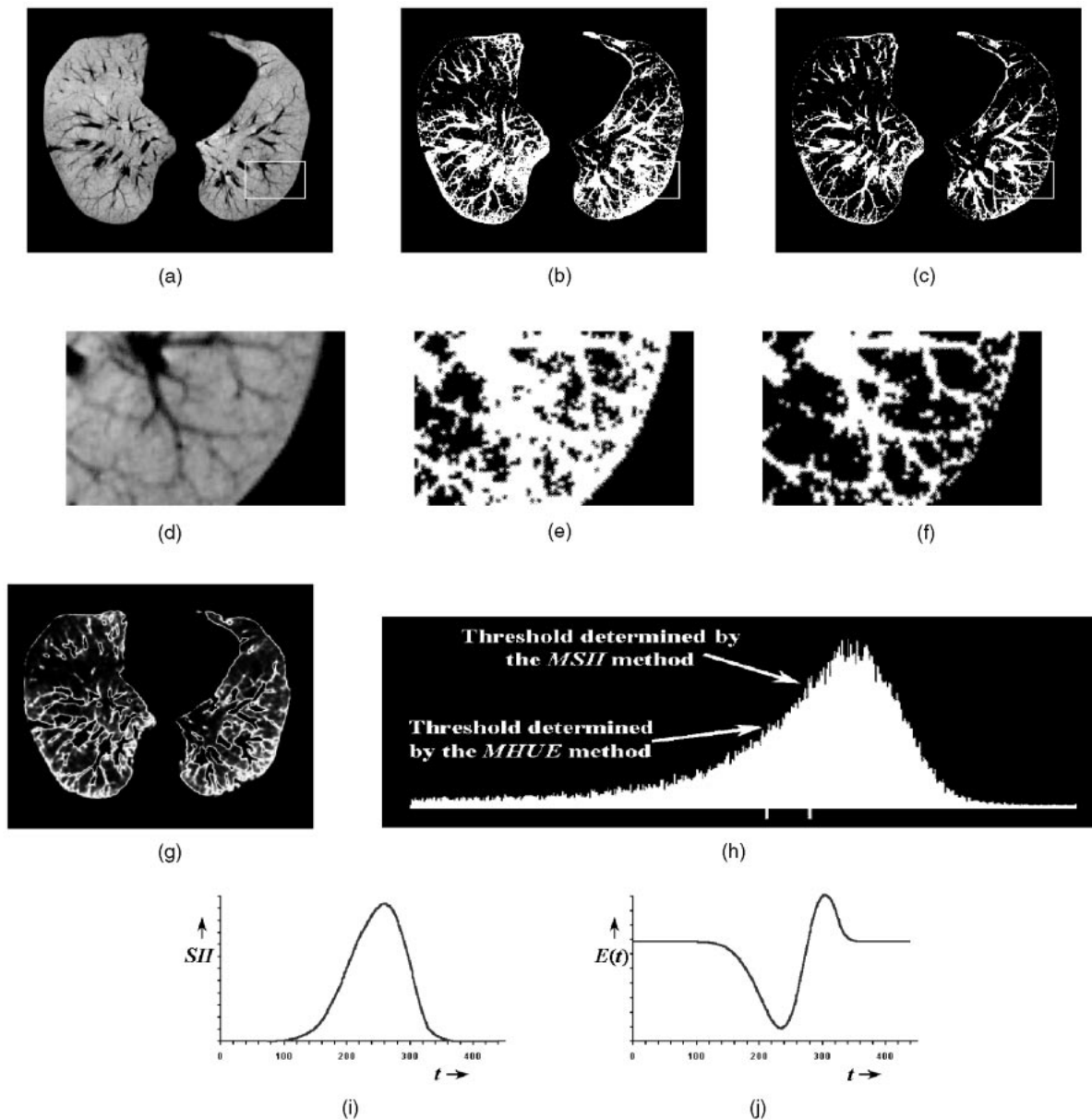


Fig. 4. A comparison between the *MSII* and *MHUE* methods. (a) A 2D slice of a 3D CT scene of an emphysema patient's lung after segmenting out surrounding tissues and then inverting intensities within the lung regions. The background is assigned a constant zero intensity. The object of interest is the dark appearing vessels. (b) Binary scene resulting from the *MSII* method. (c) Binary scene resulting from the *MHUE* method. (d), (e), and (f) Three subscenes taken at the same region of interest from (a), (b), and (c), respectively. (g) The scene of class uncertainty at the optimum threshold estimated using the *MHUE* method. (h) The locations of the selected thresholds on the scene intensity histogram. (i) The plot of *SII* as a function of threshold t . (j) The plot of $E(t)$.

3.1 Qualitative Comparison

A qualitative phantom example has already been demonstrated in Figs. 1, 2, and 3. It is clear that the segmentation in Fig. 2c is visually more acceptable than that in Fig. 3c. Further examples based on clinical images are illustrated in this section. *MSII* and *MHUE* methods return a threshold. In the following examples, which of (2.6) or (2.7) is used to obtain the binary scene is obvious from the context.

3.1.1 Two-Object Segmentation

In this section, we present two examples—one of a CT scene of an emphysema patient's lung and the other of a patient's digitized mammogram. Here, we concentrate on the

problem of segmenting a single object from the background. Although the original segmentation problem in these cases involves multiple objects, we did a presegmentation in the first example to separate lung parenchyma and, in the second example, to separate breast parenchyma. Within the segmented parenchymal region, a two-object segmentation task is considered—in the lungs (Fig. 4), to separate vessels, and in the breast (Fig. 5), to delineate the dense fibroglandular regions.

The scene shown in Fig. 4a was prepared by first segmenting the lung region from the surrounding tissues using the fuzzy connectedness method [3], [31] and then by inverting scene intensities within the lungs so that lung

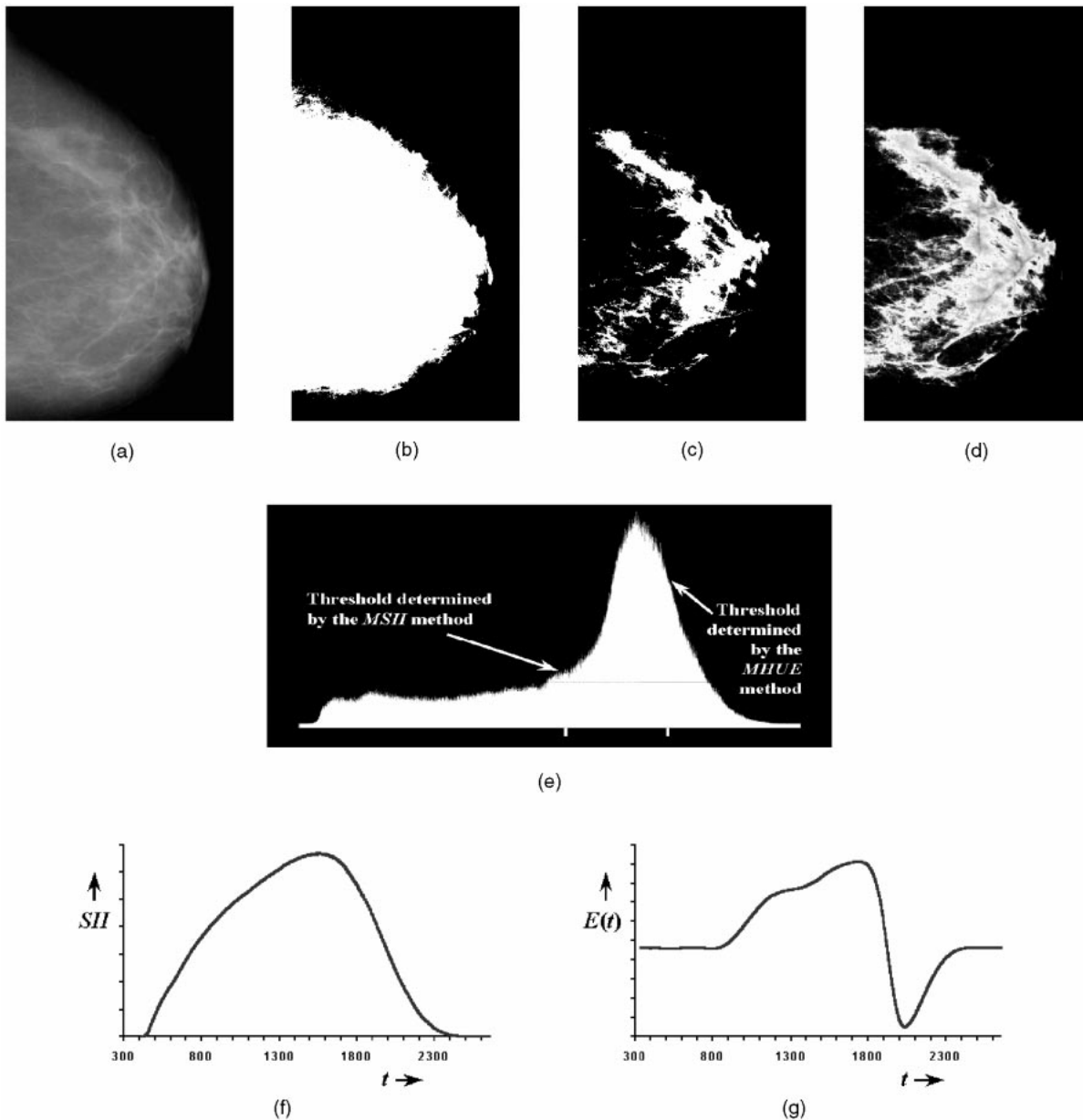


Fig. 5. A comparison between the *MSII* and *MHUE* methods. (a) A 2D X-ray projection scene of a patient's breast after segmenting out the background and then assigning a constant zero intensity to it. (b) Binary scene resulting from the *MSII* method. The object of interest is the bright fibroglandular region in (a). (c) Binary scene resulting from the *MHUE* method. (d) The scene of class uncertainty at the optimum threshold estimated using the *MHUE* method. (e) The locations of the selected thresholds on the scene intensity histogram. (f) The plot of SII as a function of threshold t . (g) The plot of $E(t)$.

parenchymal tissues appear brighter while the vessels appear darker. In Fig. 4a, the nonlung region is shown with zero intensity. The size of the scene is 410×328 with a pixel size of $0.78 \times 0.78 \text{ mm}^2$. Figs. 4b and 4c show segmented vessel regions using the *MSII* and *MHUE* methods, respectively. Note in Fig. 4b that the segmented region includes more lung parenchymal region than in Fig. 4c. This is clearly visible in the three magnified subscenes shown in Figs. 4d, 4e, and 4f of a region of interest selected in Figs. 4a, 4b, and 4c. Fig. 4g shows the class uncertainty values as a scene at the optimum threshold selected using the *MHUE* method, which further asserts the validity of Postulate A. Fig. 4h shows the

locations of the thresholds selected by the two methods on the intensity histogram. Figs. 4i and 4j show the plots of SII and $E(t)$ values as functions of threshold t .

The scene displayed in Fig. 5a was prepared by segmenting the breast region from the background in a digitized mammogram using a method based on fuzzy connectedness [35]. Here, the task is to segment the fibroglandular tissue (appearing bright) from breast parenchyma (appearing relatively dark). The size of the scene is 1088×1969 with a pixel size of $100 \times 100 \text{ micron}^2$. Figs. 5b and 5c show the segmented fibroglandular tissue regions using the *MSII* and the *MHUE* methods. Clearly, the segmentation in Fig. 5c is much better than that in Fig. 5b.

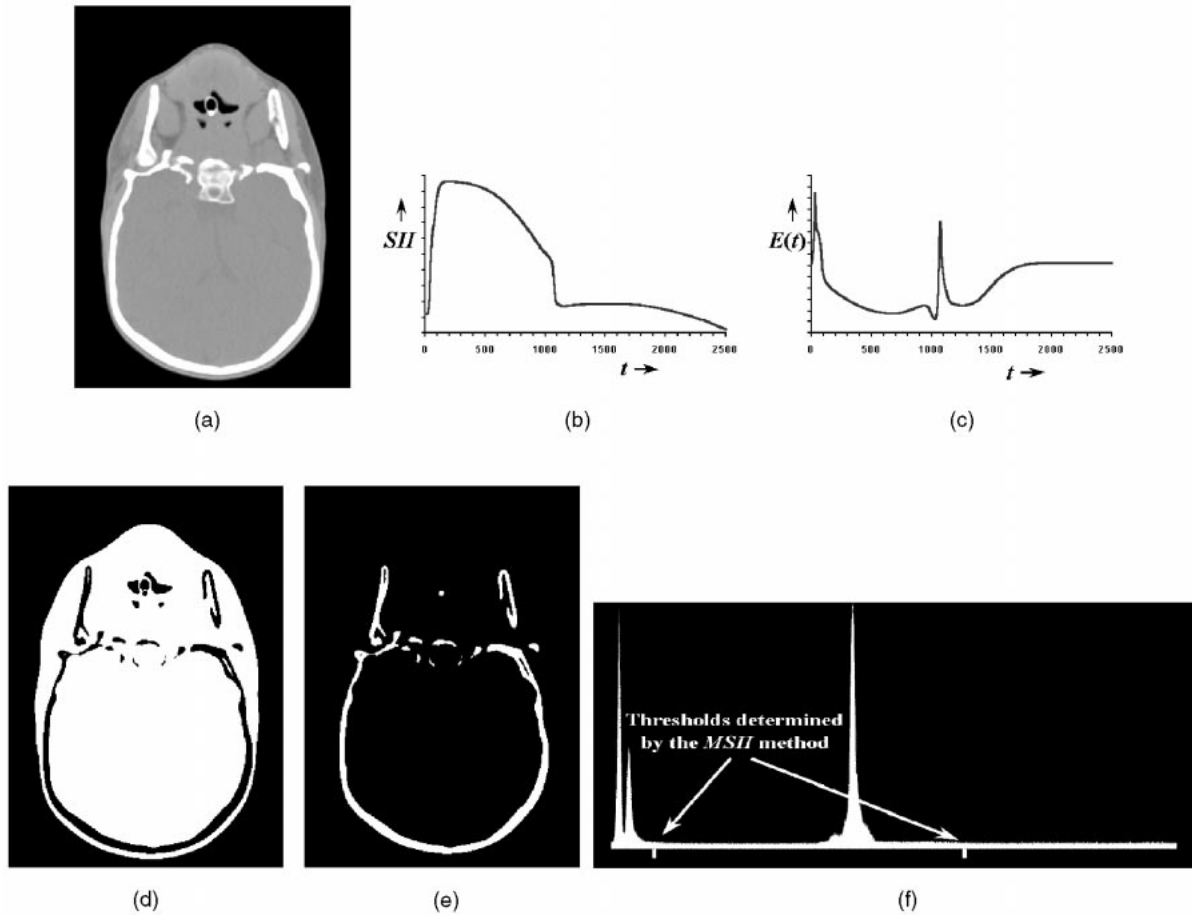


Fig. 6. A comparison between the *MSII* and *MHUE* methods for multiple object segmentation. (a) A 2D slice of a 3D CT scene of a child's head with craniofacial anomalies. (b) and (c) Plots of SII and $E(t)$ as functions of threshold t . (d) and (e) Two binary scenes resulting from thresholding within the two intervals (see (f))—from the first to the second threshold selected by the *MSII* method and from the second threshold to the largest value in Γ . (f) Locations of the local maxima of SII are shown on the scene intensity histogram. (g), (h), and (i) Three binary scenes resulting from thresholding within three intervals (see (j))—from the first to the second, the second to the third and from the third to the largest value in Γ . (j) Locations of the local minima of $E(t)$ on the scene intensity histogram. (k), (l), and (m) Three class uncertainty scenes at the threshold intensities at the three minima of $E(t)$.

Fig. 5d displays the class uncertainty values as a scene at the optimum threshold selected by the *MHUE* method. The object boundaries in this example are more fuzzy than those in Figs. 1 and 4 and this is correctly reflected in the uncertainty scene. Fig. 5e demonstrates the locations of the thresholds selected by the two methods on the intensity histogram of the scene in Fig. 5a. Figs. 5f and 5g are the plots of SII and $E(t)$ as functions of threshold t .

3.1.2 Multiple-Object Segmentation

Here, we compare the two methods in segmenting multiple objects in the CT scene of a patient with craniofacial anomalies. To identify optimum thresholds for the different objects—bone, soft tissue, and fat—we use the local minima of $E(t)$ and the local maxima of SII (this strategy was followed in [23] for SII). A 2D slice from the 3D CT scene is displayed in Fig. 6a. Bone appears with the brightest pixel values, soft tissues with intermediate, and fat with the lowest (higher than black which represents background or air) values. Figs. 6b and 6c show the plots of SII and $E(t)$ as functions of threshold t . Note that, in Fig. 6b, there are only

two local maxima of SII , while there are three local minima of $E(t)$ in Fig. 6c. Fig. 6d shows the tissue region segmented by thresholding using the intensity interval between the two local maxima of SII , while Fig. 6e shows the region with intensities greater than the rightmost local maximum of SII . The segmented background is not shown here. Fig. 6f shows locations of the two maxima of SII on the scene intensity histogram. Figs. 6g, 6h, and 6i display the three tissue regions segmented by thresholding using intensity intervals between every two successive local minima of $E(t)$. Fig. 6i shows segmentation at the threshold corresponding to the rightmost local minimum of $E(t)$. Fig. 6j shows the locations of the three local minima of $E(t)$ on the scene intensity histogram. Here, too, the segmented background is not shown. Three scenes representing class uncertainty values at each of the three local minima of $E(t)$ are displayed in Figs. 6k, 6l, and 6m. Again, it is clear that the *MHUE* method produced an acceptable segmentation of bone, soft tissue, and fat regions outperforming the *MSII* method. The latter failed to make a distinction

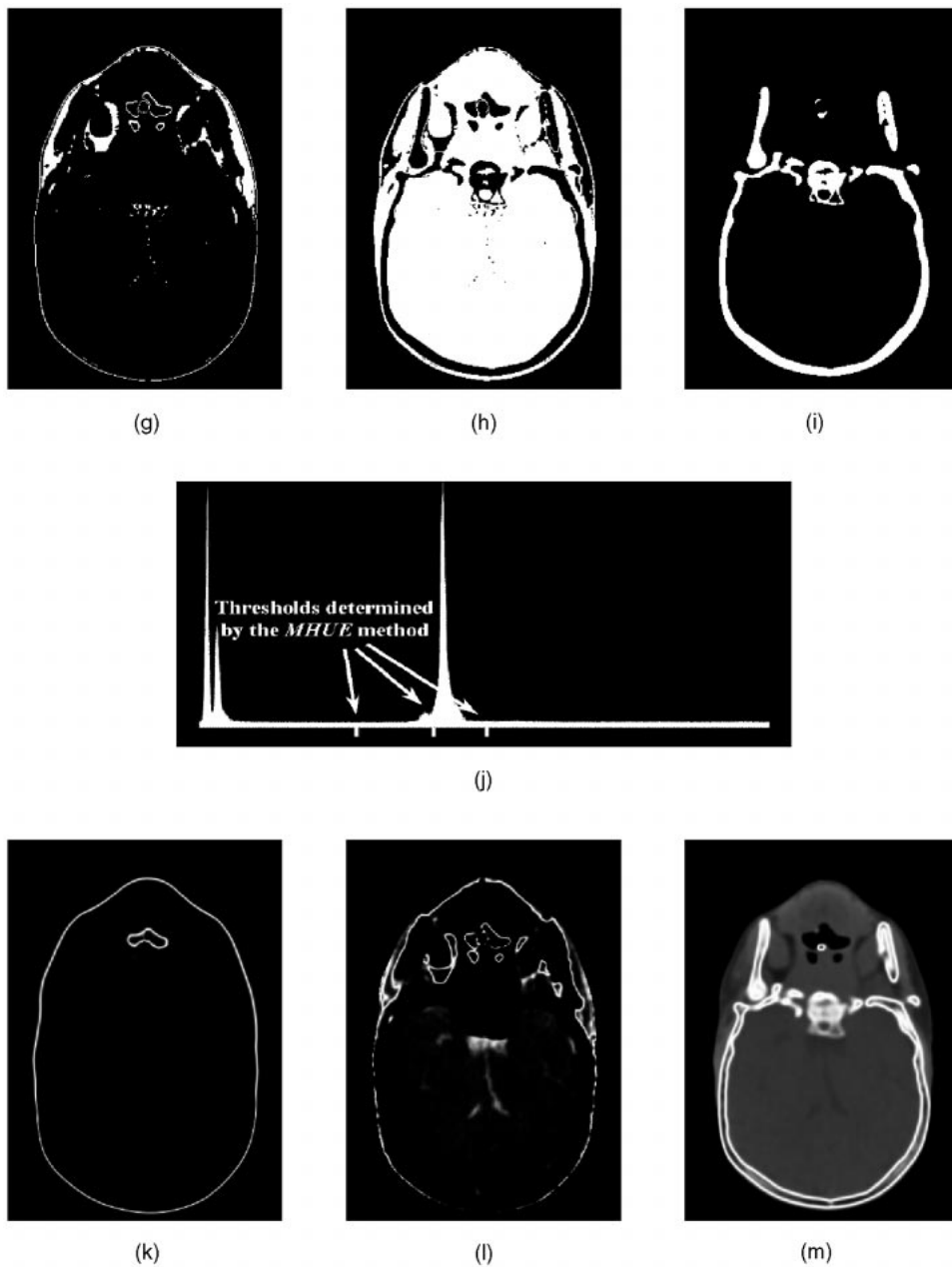


Fig. 6. (cont.).

between fat and soft tissue. Further, its segmentation of bone also seems unsatisfactory in this example.

As a final qualitative example, we applied the two methods on 10 2D slices each taken from a 3D MR scene, acquired as a FLAIR sequence, of a different patient's head with brain tumor. In each case, *MSII* detected one threshold while *MHUE* detected three thresholds—1) between background and foreground, 2) between WM/GM, and 3) for tumors. Two representative images are displayed in Figs. 7 and 8, each of which shows: (a) the original slice and (b) the binary scene segmented at the threshold selected by the *MSII* method; (c), (d), and (e) three binary scenes segmented at the thresholds selected by the *MHUE* method (again, for both methods, the binary scene (background) corresponding to the threshold

interval from the minimum of Γ to the first threshold is not shown); (f), (g), and (h) three class uncertainty scenes at the thresholds selected by the *MHUE* method; (i) and (j) plots of SII and $E(t)$ as functions of threshold t ; and (k) locations of the thresholds on the scene intensity histogram. It should be pointed out that, in the intensity histograms, illustrated in Figs. 7k and 8k, there is hardly any information illustrating a threshold between WM and GM, while the proposed method successfully detected it in each case.

All six visual examples illustrated in this section in Figs. 2, 3, 4, 5, 6, 7, and 8 overwhelmingly indicate the correct behavior and the superiority of the *MHUE* method over *MSII*. Further, Figs. 1c, 4g, 5d, 6k, 6l, 6m, 7f, 7g, 7h, 8f, 8g, and 8h support the validity of Postulate A. The

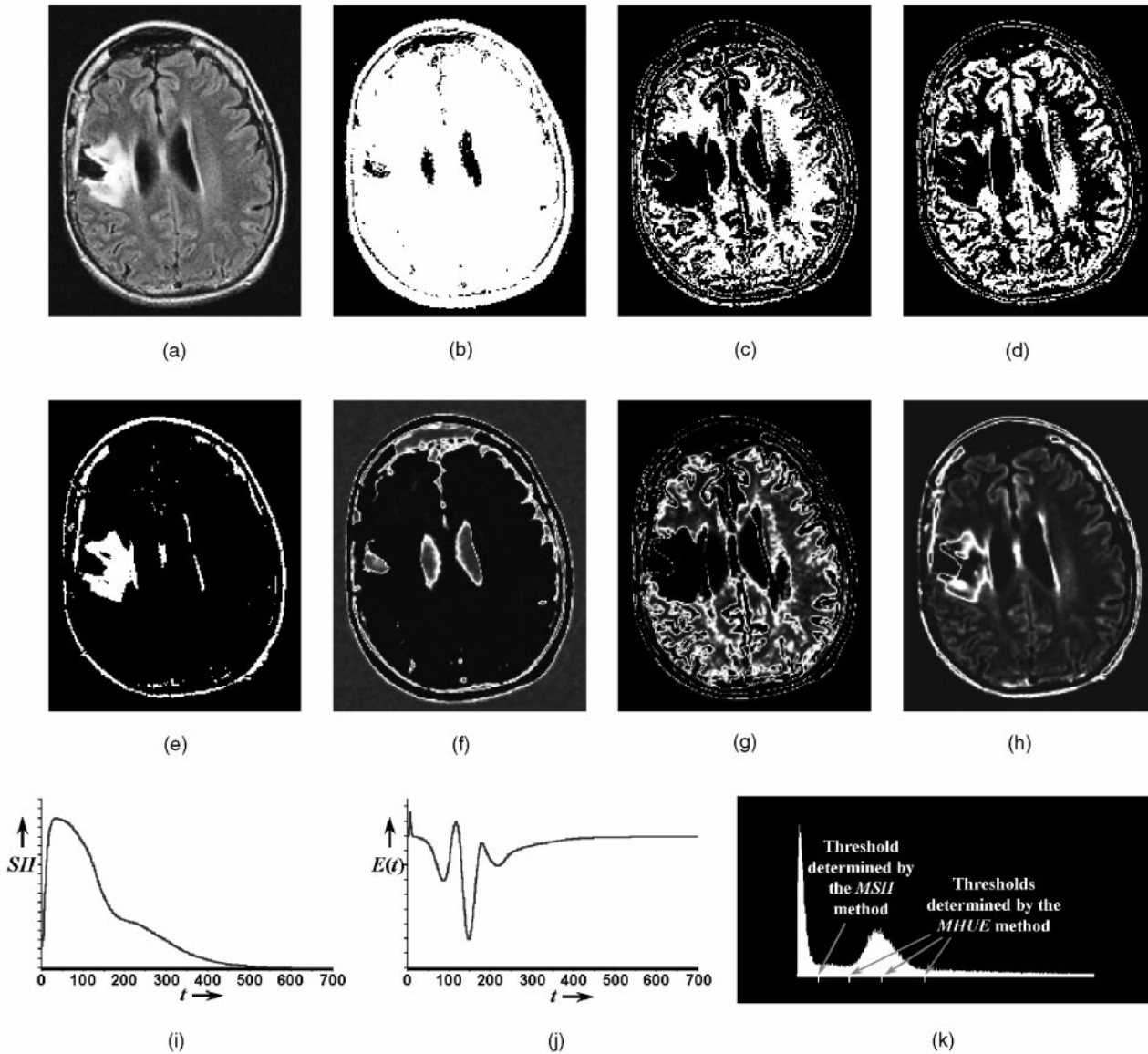


Fig. 7. A comparison between the *MSII* and *MHUE* methods on a 2D slice of a 3D MR scene, acquired as per FLAIR sequence, of a patient's head with brain tumor. (a) The original slice. (b) The binary scene segmented at the threshold selected by the *MSII* method. (c), (d), and (e) Three binary scenes segmented at the thresholds selected by the *MHUE* method. (f), (g), and (h) Three class uncertainty scenes at the thresholds selected by the *MHUE* method. (i) and (j) Plots of *SII* and $E(t)$ as functions of threshold t . (k) Locations of the thresholds on the scene intensity histogram.

computational time for estimating the optimum threshold for the *MSII* and *MHUE* methods for 2D scenes are typically 500-900 milliseconds and 1-2 seconds, respectively, on a 550 MHz Intel Pentium PC running under Linux OS. For a 3D scene, these times are typically 7-9 seconds and 40-50 seconds. It may be noted that *MHUE* is computationally a little more expensive, especially in 3D. This is because of two reasons—1) the *MHUE* method needs to compute the homogeneity map at each spel which is done only once and 2) for every threshold t , estimation of the energy function $E(t)$ needs floating point operations for every spel over the entire image domain. The second part of the algorithm becomes more expensive as the image size and its dimensionality increase. A significant improvement

in computation can be achieved if $E(t)$ is computed as follows:

$$\begin{aligned}
 \text{for any } i \in \Gamma \quad X(i) &= \sum_{\forall c \in C \text{ s.t. } f(c)=i} \mu_{\tau}(c), \quad \text{and} \\
 Y(i) &= \sum_{\forall c \in C \text{ s.t. } f(c)=i} 1 - \mu_{\tau}(c) \\
 E(t) &= \sum_{i \in \Gamma} H_t(i)X(i) + (1 - H_t(i))Y(i).
 \end{aligned} \tag{3.35}$$

It is not difficult to see that $E(t)$ as computed by (2.6) and that computed according to (3.35) are the same. On the other hand, the computation of $E(t)$ is much faster in the latter case because it does not need floating point operations for every spel location, rather, it is performed over the intensity domain whose size is generally much smaller

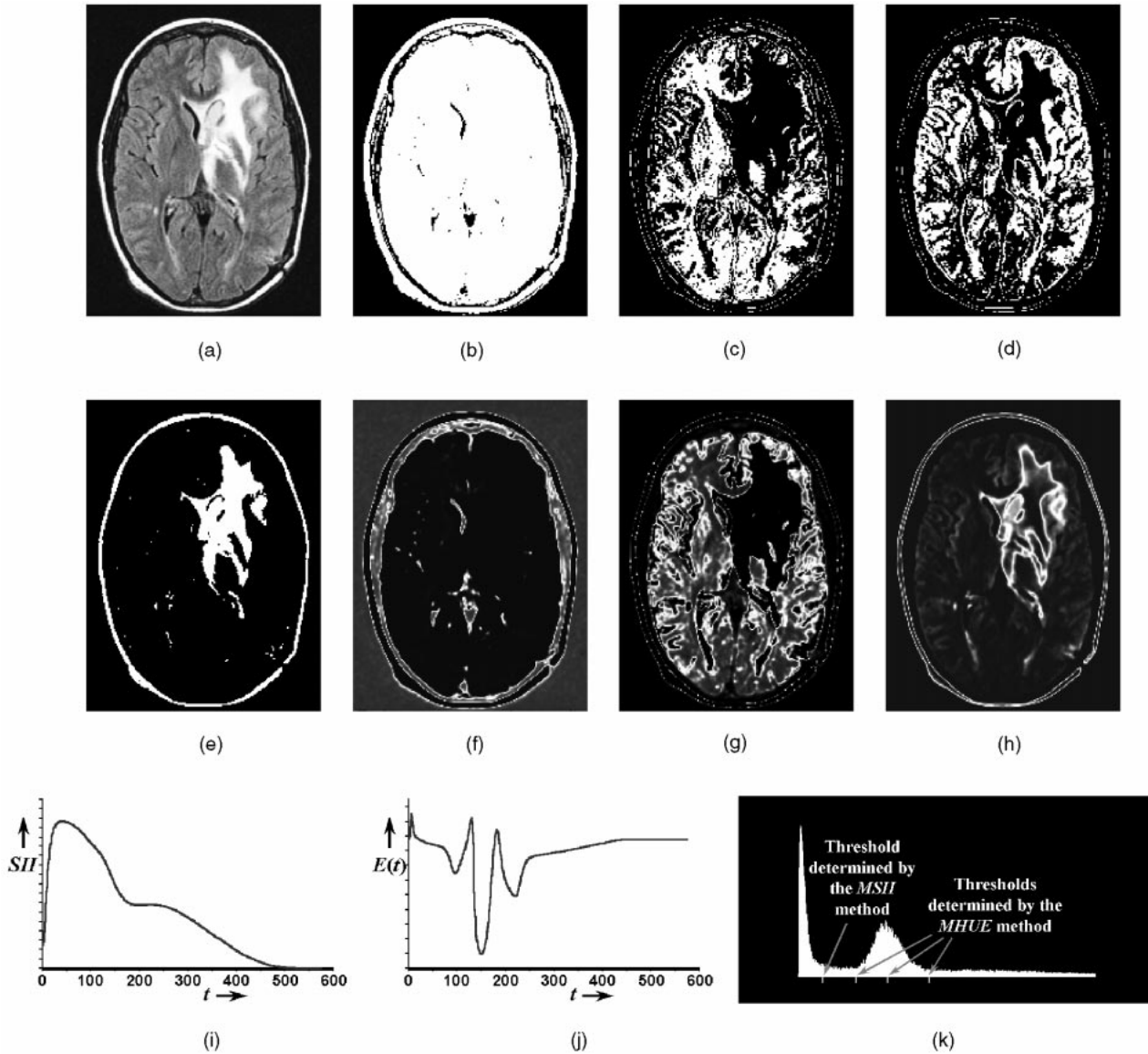


Fig. 8. Same as Fig. 7, but for another patient's head.

compared to the size of the image domain; note that the computation of $X(i)$ and $Y(i)$ is done once and it needs only one scan of the image space.

3.2 Quantitative Comparison

In this section, we describe a phantom experiment to objectively evaluate the performance of the two methods and also to demonstrate the accuracy of the proposed method by comparing its results with those produced by the theoretically best possible thresholds. To do so, we have taken an approach of generating phantom scenes by user-steered segmentation of real medical scenes and, subsequently, applying to this result a contrast measured from real scenes, and various degrees of noise, blurring, and background variation. Specifically, the phantom scenes were generated from 10 3D MR studies of different patients' brains. One transaxial slice, approximately at the same location in the brain, was selected from each of these 10 3D scenes. Each of these 10 2D scenes was segmented carefully to separate the WM region using a user-steered segmentation method [36].

In this boundary-based segmentation approach, optimal boundary segments are determined by an algorithm between user selected points on the boundary of the object. After interactively segmenting the WM region from each of the original 10 2D scenes, a new 2D scene was generated by assigning to spels in the WM region a constant intensity equal to the average of the intensities within the same segmented region in the corresponding original MR 2D scene. A "gray matter background" was provided in each simulated 2D scene by assigning to each spel in the background region a constant intensity equal to the average of the intensities within the (interactively segmented) GM region in the corresponding original MR 2D scene. From this set of 10 simulated 2D scenes, we created a total of 250 2D scenes by blurring (using a 2D Gaussian kernel) each 2D scene to five different degrees of blurring and by adding a zero-mean Gaussian correlated noise component at five different levels. Each of these 250 2D scenes was further modified by adding a fixed slow varying (ramp) background component which

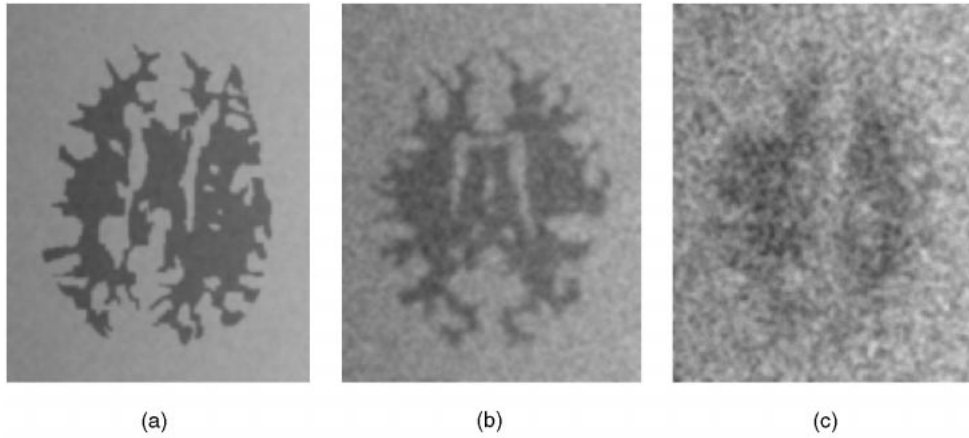


Fig. 9. Three representative samples from the 250 2D phantom scenes created in our experiment. The levels of blurring and noise in these scenes are (a) low, (b) medium, and (c) high.

changed from zero at the first column to 100 at the last column in the 2D scene. This new set of 2D scenes, denoted $\mathcal{I} = \{\mathcal{C}_i \mid \mathcal{C}_i = (C, f_i), 1 \leq i \leq 250\}$, was utilized in our experiments. Three samples from this set are displayed in Figs. 9a, 9b, and 9c.

Let $\mathcal{I}_b = \{\mathcal{C}_{bi} \mid \mathcal{C}_{bi} = (C, f_{bi}), 1 \leq i \leq 250\}$ be the set of binary scenes such that, for $1 \leq i \leq 250$, \mathcal{C}_{bi} represents the original segmentation of the WM region corresponding to the simulated scene \mathcal{C}_i in \mathcal{I} . Scenes in \mathcal{I}_b will be used as true segmentations. For $x \in \{MSII, MHUE\}$, we denote by $\mathcal{I}_x = \{\mathcal{C}_{xi} \mid \mathcal{C}_{xi} = (C, f_{xi}), 1 \leq i \leq 250\}$ the set of binary scenes produced by using method x from the set \mathcal{I} of input scenes. Figs. 10a, 10e, and 10i show the original binary scenes \mathcal{C}_{bi} corresponding to the scenes shown in Figs. 9a, 9b, and 9c, respectively. The corresponding thresholded scenes \mathcal{C}_{xi} are shown in Figs. 10b, 10f, and 10j for $x = MSII$ and in Figs. 10c, 10g, and 10k for $x = MHUE$. Figs. 10d, 10h, and 10l show the three segmentation results each at the best possible threshold at which the area of mismatch with the corresponding initial truth (scenes in \mathcal{I}_b) is the minimum. The set of binary scenes produced in this manner (we shall refer to this as method O), at the best possible thresholds, will be denoted by $\mathcal{I}_O = \{\mathcal{C}_{Oi} \mid \mathcal{C}_{Oi} = (C, f_{Oi}), 1 \leq i \leq 250\}$. We would like to emphasize here that no thresholding method can do better than method O . Figs. 11a, 11b, and 11c display the three class uncertainty scenes at the optimum thresholds selected by $MHUE$ for the three phantom scenes shown in Figs. 9a, 9b, and 9c, respectively.

We define figures of merit FOM_{xi} , for

$$x \in \{MSII, MHUE, O\},$$

to describe the accuracy of segmentation of the three methods as follows: For $x \in \{MSII, MHUE, O\}$ and $1 \leq i \leq 250$,

$$FOM_{xi} = \left(1 - \frac{|\mathcal{C}_{xi} \text{ EOR } \mathcal{C}_{bi}|_1}{|\mathcal{C}|}\right) \times 100, \quad (3.36)$$

where EOR represents the Exclusive OR operation between the two binary scenes and $|\mathcal{C}_{xi} \text{ EOR } \mathcal{C}_{bi}|_1$ denotes the number of 1-valued spels in $\mathcal{C}_{xi} \text{ EOR } \mathcal{C}_{bi}$. For method x , FOM_{xi} represents the degree of match between the original

(true) object region captured in \mathcal{C}_{bi} and the object region in \mathcal{C}_{xi} resulting after thresholding.

Tables 1 and 2 list, in each cell, the mean and the standard deviation of FOM_{xi} for $x \in \{MSII, MHUE\}$, respectively, for each degree of blurring and noise for the 10 scenes. The degree of blurring and noise increases along rows and columns, respectively. The first entry in each cell of Table 3 lists the difference of the mean FOM values for the two methods at the corresponding degree of blurring and noise. A paired Student t test [37] of the 10 pairs of FOM data for each degree of blurring and noise was conducted under the null hypothesis that there is no statistical difference between the two methods. The second entry in each cell of Table 3 shows the p value of the paired t -test at the corresponding degree of blurring and noise. Note from Tables 1 and 2 that, at each degree of noise and blurring, the mean FOM value achieved using the $MSII$ method is less than that achieved using the proposed method indicating the superiority of the $MHUE$ method. As demonstrated in Table 3, these differences are statistically significant ($p < 0.05$) at every combination of blurring and noise except for the two cells representing 1) minimum noise and minimum blurring and 2) second minimum noise and minimum blurring. The high standard deviation of the FOM values resulting from the $MSII$ method seems to be partly responsible for the high p values at these cells.

We also compared the performance of $MHUE$ to method O . Table 4 lists the mean and the standard deviation of FOM_{Oi} for each degree of blurring and noise. Table 5 lists the difference of the mean FOM values for methods $MHUE$ and O . Since the FOM values at the best possible thresholds are minimum by definition, a threshold selection technique can never give a better result than this ideal. Therefore, a paired t -test has no meaning in this context. However, Table 5 gives us an overall assessment of the accuracy of the proposed method. In terms of the percent of region overlap, the new method deviates from the ideal result by not more than 0.2 percent most of the time (average 0.13 percent)! This is vividly demonstrated in Figs. 10k and 10l. It is also clear that any other threshold selection method cannot possibly perform significantly better than the $MHUE$ method, at least on these phantom scenes, since the average room for improvement is only 0.13 percent.

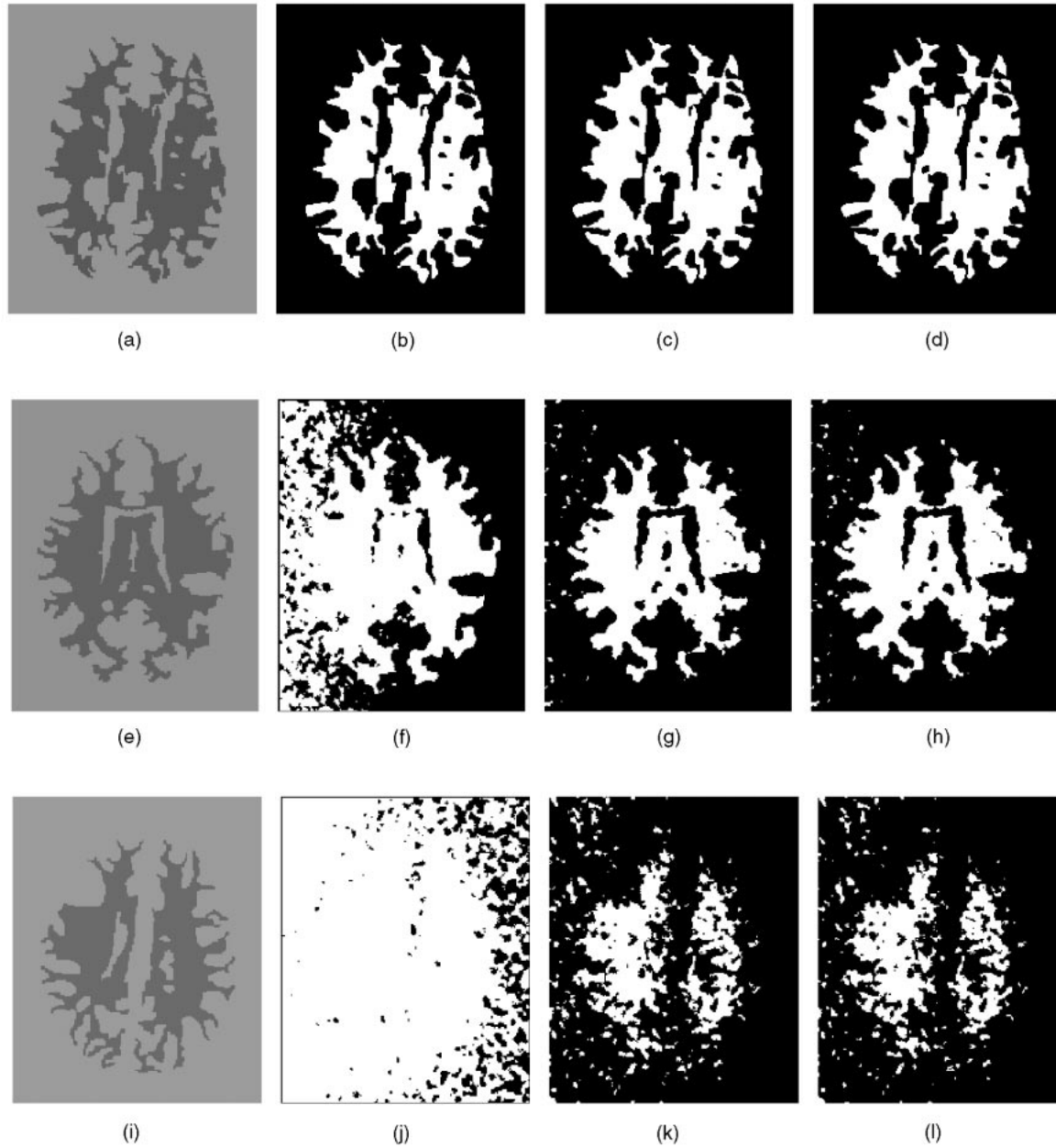


Fig. 10. (a), (e), and (i) Original binary scenes from the set \mathcal{I}_b corresponding to the scenes shown in Figs. 9a, 9b, and 9c, respectively. (b), (f), and (j) Binary scenes corresponding to the scenes of Figs. 9a, 9b, and 9c, respectively, resulting from the *MSII* method. (c), (g), and (k) Analogous to (b), (f), and (j) but resulting from the *MHUE* method. (d), (h), and (l) Analogous to (b), (f), and (j) but resulting from the theoretical best possible thresholds (method *O*) that gave the minimum area of mismatch with the corresponding binary scene in (a), (e), and (i).

4 CONCLUSION

A new thresholding method has been introduced that determines optimum thresholds by considering image morphology. The method effectively combines image homogeneity information with histogram-based class uncertainty information. The fundamental premise of this work is the postulate that, in most real life imaging applications, at the optimum threshold, image elements with the maximum class uncertainty occur in the vicinity of the object boundaries. Guided by this postulate, we have demonstrated how region homogeneity information can be utilized in combination with class uncertainty to dramatically improve the estimation of optimum thresholds.

Qualitative comparisons using both phantoms and clinical medical images indicate significant improvements in the thresholded results obtained by the proposed method over those obtained by a recently published (*MSII*) method [23]. The superiority of the proposed method in multiple object segmentation has also been demonstrated. An objective comparison between the two methods has been made using a set of 250 realistic phantoms generated under a range of conditions of blurring, noise, and background variation. The comparison has clearly indicated the superiority of the proposed method over the *MSII* method. The same quantitative comparison was made between the proposed method and, theoretically, the best possible thresholding strategy utilizing the known true segmentations. This

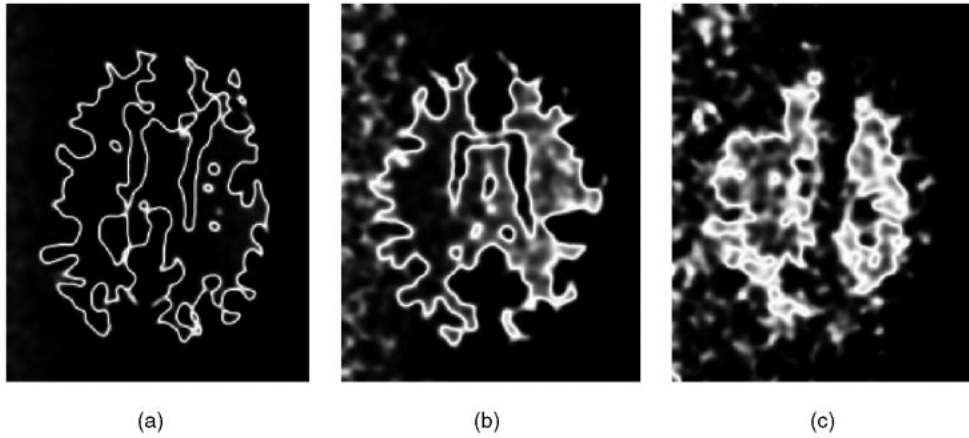


Fig. 11. The class uncertainty scenes corresponding to the scenes in Figs. 9a, 9b, and 9c, respectively, computed at the optimum thresholds estimated by the *MHUE* method.

indicated that the performance of our method was very close to the theoretical best at every combination of the level of blurring and noise that was considered in our experiment. This, we believe, gives a strong evidence of the validity of our central postulate and of the method based on it. Purely histogram-based methods run into difficulty when histograms are modified due to blurring, noise, and background variations, as illustrated in Figs. 10f and 10j. When scene morphology information is utilized, this deficiency in the histogram information can be effectively overcome. This is the essence of this work. Although the results illustrated in this paper are all taken from different medical images, since any acquired image satisfies the conditions of Postulate A, our observations should hold good for all such images. Further, the phantom experiment indicates that any other thresholding method is very unlikely to perform significantly better than the proposed method since the latter is very close in performance to the theoretically best possible (ideal)

thresholding. Although we have not done this investigation in this paper, we believe that, for 3D and higher-dimensional scenes, it is generally better to consider a full dimensional formulation than a slice-by-slice strategy for determining all entities (including scale) associated with the scene homogeneity function μ_r . The higher-dimensional formulations more accurately capture or account for the underlying phenomenon, such as neighborhood, blurring, noise, and change in homogeneity, than the 2D version. In some applications [39], if the background inhomogeneity along the direction orthogonal to the slice planes is high or if the anisotropy of resolution in this direction is high, a slice-by-slice approach may provide a better result. However, such a treatment may also have its side effects such as introducing discontinuities in the third dimension in the geometry of the surface. Although, in this paper, we did not consider scene resolution anisotropy, this is easily incorporated into our framework by taking into account the actual size of the spels

TABLE 1

The Mean (First Entry) and Standard Deviation (Second Entry) of FOM_{xi} Values Are Shown in Each Cell for Different Blurring and Noise Conditions for $x = MSII$

	Blur 1	Blur 2	Blur 3	Blur 4	Blur 5
Noise 1	95.087 10.346	84.139 17.053	77.175 16.906	71.089 16.723	66.316 14.502
Noise 2	95.303 9.060	81.698 16.384	74.858 15.159	72.294 15.747	67.517 13.677
Noise 3	90.952 10.230	83.362 12.862	78.864 12.123	74.906 12.960	68.130 8.793
Noise 4	86.444 6.260	83.914 5.792	81.092 5.198	77.233 4.987	73.985 4.895
Noise 5	82.006 3.954	79.996 3.304	77.869 3.594	75.289 2.895	73.125 3.462

Blurring increases along rows from left to right, while noise increases along columns from top to bottom.

TABLE 2

The Mean (First Entry) and Standard Deviation (Second Entry) of FOM_{xi} Values Are Shown in Each Cell for Different Blurring and Noise Conditions for $x = MHUE$

	Blur 1	Blur 2	Blur 3	Blur 4	Blur 5
Noise 1	98.469 0.343	97.460 0.544	96.507 0.668	94.850 0.902	93.377 1.131
Noise 2	98.111 0.271	97.117 0.482	96.167 0.620	94.468 0.854	92.968 1.054
Noise 3	97.450 0.553	96.341 0.638	95.336 0.731	93.564 0.922	92.029 1.093
Noise 4	94.926 1.014	93.752 1.076	92.672 1.117	90.740 1.194	89.138 1.265
Noise 5	91.691 1.322	90.434 1.353	89.380 1.341	87.503 1.347	85.967 1.414

Blurring increases along rows from left to right, while noise increases along columns from top to bottom.

TABLE 3
The First Entry in Each Cell Shows the Mean of $FOM_{MHUE_i} - FOM_{MSII_i}$ Over the 10 Scenes for Different Degrees of Blurring and Noise

	Blur 1	Blur 2	Blur 3	Blur 4	Blur 5
Noise 1	3.382 0.32415	13.350 0.03500	19.332 0.00509	23.767 0.00134	27.061 0.00022
Noise 2	2.941 0.34957	15.459 0.01497	21.319 0.00145	22.186 0.00140	25.451 0.00022
Noise 3	6.545 0.07144	13.030 0.00932	16.478 0.00158	18.658 0.00108	23.899 0.00001
Noise 4	8.482 0.00090	9.837 0.00015	11.581 0.00001	13.507 0.00000	15.153 0.00000
Noise 5	9.685 0.00000	10.438 0.00000	11.511 0.00000	12.214 0.00000	12.843 0.00000

The second entry of each cell shows the p value of the paired Students' t test. Blurring increases along rows from left to right, while noise increases along columns from top to bottom.

along the different coordinate axes. We believe that for anisotropic scenes, taking into account spel size in this manner will give better results than when isotropicity is forced on them. See [38] for details as to how scale computation can be modified in this fashion. Vector-valued (color, multichannel, multiprotocol) scenes are common in many applications. For best results, we believe that the class uncertainty and homogeneity components of energy should be formulated using a vector variable.

The premise underlying Postulate A, we believe, is a fundamental construction. It has consequences in other image processing operations. We are currently investigating

TABLE 4
The Mean (First Entry) and Standard Deviation (Second Entry) of FOM_{Oi} Values Are Shown in Each Cell for Different Blurring and Noise Conditions

	Blur 1	Blur 2	Blur 3	Blur 4	Blur 5
Noise 1	98.567 0.308	97.545 0.444	96.624 0.572	94.950 0.815	93.530 1.006
Noise 2	98.274 0.369	97.189 0.509	96.246 0.621	94.544 0.855	93.052 1.026
Noise 3	97.536 0.568	96.408 0.669	95.432 0.773	93.638 0.959	92.114 1.124
Noise 4	94.988 0.990	93.826 1.075	92.781 1.111	90.890 1.225	89.316 1.290
Noise 5	91.760 1.275	90.610 1.290	89.594 1.319	87.843 1.344	86.344 1.369

Blurring increases along rows from left to right, while noise increases along columns from top to bottom.

TABLE 5
Mean of $FOM_{Oi} - FOM_{MHUE_i}$ Over the 10 Scenes for Different Degrees of Blurring and Noise

	Blur 1	Blur 2	Blur 3	Blur 4	Blur 5
Noise 1	0.098	0.085	0.117	0.100	0.153
Noise 2	0.163	0.072	0.080	0.077	0.084
Noise 3	0.086	0.067	0.096	0.074	0.085
Noise 4	0.062	0.075	0.108	0.149	0.178
Noise 5	0.069	0.176	0.214	0.339	0.377

Blurring increases along rows from left to right, while noise increases along columns from top to bottom.

its utility in devising fuzzy affinity relations for fuzzy connected object definition and image segmentation. Finally, the purpose of this paper is not a commentary on the effectiveness of thresholding as a segmentation tool compared to other nonthresholding techniques. Rather, if one decides to use thresholding for segmentation, then it demonstrates how to choose this threshold so that the resulting segmentation is as close to optimal as possible.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Warren Gefter, Emily Conant, David Hemmy, and Robert Grossman for the image data utilized in this work. The research reported here is supported by NIH grants NS37172 and AR46902 and a grant from the US Department of Army DAMD 179717271.

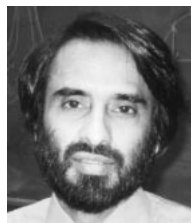
REFERENCES

- [1] Z.H. Cho, J.P. Jones, and M. Sing, *Foundation of Medical Imaging*. New York: John Wiley & Sons, 1993.
- [2] N.R. Pal and S.K. Pal, "A Review on Image Segmentation Techniques," *Pattern Recognition*, vol. 26, pp. 1277-1294, 1993.
- [3] J.K. Udupa and S. Samarasekera, "Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation," *Graphical Models and Image Processing*, vol. 58, no. 3, pp. 246-261, 1996.
- [4] P.K. Sahoo, S. Soltani, A.K.C. Wong, and Y.C. Chen, "A Survey of Thresholding Techniques," *Computer Vision, Graphics, and Image Processing*, vol. 41, pp. 233-260, 1988.
- [5] S.U. Lee, S.Y. Chung, and R.H. Park, "A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation," *Computer Vision, Graphics, and Image Processing*, vol. 52, pp. 171-190, 1990.
- [6] C.A. Glasbey, "An Analysis of Histogram Based Thresholding Algorithms," *Proc. Graphical Models and Image Processing*, vol. 55, pp. 532-537, 1993.
- [7] J.M.S. Prewitt and M.L. Mendelsohn, "The Analysis of Cell Images," *Ann. New York Academy Science*, vol. 128, pp. 1035-1053, 1966.
- [8] W. Doyle, "Operations Useful for Similarity-Invariant Pattern Recognition," *J. ACM*, vol. 9, pp. 259-267, 1962.
- [9] N. Otsu, "A Threshold Selection Method from Gray-Level Histogram," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 8, pp. 62-66, 1978.
- [10] T. Rider and S. Calvard, "Picture Thresholding Using an Iterative Selection Method," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 8, pp. 630-632, 1978.
- [11] H.J. Trussell, "Comments on 'Picture Thresholding Using an Iterative Selection Method,'" *IEEE Trans. Systems, Man, and Cybernetics*, vol. 9, p. 311, 1979.

- [12] W. Tsai, "Moment-Preserving Thresholding: A New Approach," *Computer Vision, Graphics, and Image Processing*, vol. 29, pp. 377-393, 1985.
- [13] T. Pun, "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram," *Signal Processing*, vol. 2, pp. 223-237, 1980.
- [14] A.K.C. Wong and P.K. Sahoo, "A Gray-Level Threshold Selection Method Based on Maximum Entropy Principle," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 19, pp. 866-871, 1989.
- [15] N.R. Pal and S.K. Pal, "Entropy Thresholding," *Signal Processing*, vol. 16, pp. 97-108, 1989.
- [16] N.R. Pal and S.K. Pal, "Entropy: A New Definition and Its Applications," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 21, pp. 1260-1270, 1991.
- [17] J.N. Kapur, P.K. Sahoo, and A.K.C. Wong, "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram," *Computer Vision, Graphics, and Image Processing*, vol. 29, pp. 273-285, 1985.
- [18] A.S. Abutaleb, "Automatic Thresholding of Gray-Level Pictures Using Two-Dimensional Entropy," *Computer Vision, Graphics, and Image Processing*, vol. 47, pp. 22-32, 1989.
- [19] A. Brink, "Maximum Entropy Segmentation Based on the Autocorrelation Function of the Image Histogram," *J. Computer Information Technology*, vol. 2, pp. 77-85, 1994.
- [20] C.H. Li and C.K. Lee, "Minimum Entropy Thresholding," *Pattern Recognition*, vol. 26, pp. 617-625, 1993.
- [21] J. Kittler and J. Illingworth, "Minimum Cross Error Thresholding," *Pattern Recognition*, vol. 19, pp. 41-47, 1986.
- [22] S.M. Dunn, D. Harwood, and L.S. Davis, "Local Estimation of the Uniform Error Threshold," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 1, pp. 742-747, 1984.
- [23] C.K. Leung and F.K. Lam, "Maximum Segmented Image Information Thresholding," *Graphical Models and Image Processing*, vol. 60, pp. 57-76, 1998.
- [24] P.K. Saha and J.K. Udupa, "A New Optimum Thresholding Method Using Region Homogeneity and Class Uncertainty," *Proc. SPIE: Medical Imaging 2000*, vol. 3979, pp. 180-191, 2000.
- [25] F.C. Cheng and J.W. Woods, "Compound Gauss-Markov Random Fields for Image Segmentation," *IEEE Trans. Signal Processing*, vol. 39, pp. 683-697, 1991.
- [26] A. Sarkar, M.K. Biswas, and K.M.S. Sharma, "A Simple Unsupervised MRF Model Based Image Segmentation Approach," *IEEE Trans. Image Processing*, vol. 9, pp. 801-812, 2000.
- [27] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *Int'l J. Computer Vision*, pp. 321-339, 1988.
- [28] A. Kaufmann, *Introduction to the Theory of Fuzzy Subsets, Vol 1*. New York: Academic Press, 1975.
- [29] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana, Ill.: Univ. of Illinois Press, 1964.
- [30] J.K. Udupa, L. Wei, S. Samarasekera, Y. Miki, M.A. van Buchem, and R.I. Grossman, "Multiple Sclerosis Lesion Quantification Using Fuzzy Connectedness Principles," *IEEE Trans. Medical Imaging*, vol. 16, pp. 598-609, 1997.
- [31] P.K. Saha, J.K. Udupa, and D. Odhner, "Scale-Based Fuzzy Connected Image Segmentation: Theory, Algorithms, and Validation," *Computer Vision and Image Understanding*, vol. 77, pp. 145-174, 2000.
- [32] P.K. Saha and J.K. Udupa, "Scale-Based Fuzzy Connectivity: A Novel Image Segmentation Methodology and Its Validation," *Proc. SPIE: Medical Imaging*, vol. 3661, pp. 246-257, 1999.
- [33] S.M. Pizer, D. Eberly, D.S. Fritsch, and B.S. Morse, "Zoom-Invariant Vision of Figural Shape: The Mathematics of Cores," *Computer Vision and Image Understanding*, vol. 69, pp. 55-71, 1998.
- [34] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*. New York: John Wiley & Sons, 1973.
- [35] P.K. Saha, J.K. Udupa, E.F. Conant, and D.P. Chakraborty, "Near-Automatic Segmentation and Quantification of Mammographic Glandular Tissue Density," *Proc. SPIE: Medical Imaging*, vol. 3661, pp. 266-276, 1999.
- [36] A.X. Falcão, J.K. Udupa, S. Samarasekera, and S. Sharma, "User-Steered Image Segmentation Paradigms: Live Wire and Live Lane," *Graphical Models and Image Processing*, vol. 60, pp. 233-260, 1998.
- [37] B. Ronser, *Fundamentals of Biostatistics*. New York: Duxbury Press, 1995.
- [38] P.K. Saha and J.K. Udupa, "Scale-Based Diffusive Filtering of Medical Images," *Proc. SPIE: Medical Imaging*, vol. 3979, pp. 735-746, 2000.
- [39] R. Velthuisen et al., "Unsupervised Tumor Volume Measurement Using Magnetic Resonance Brain Images," *J. Magnetic Resonance Imaging*, vol. 5, pp. 594-606, 1995.



Punam K. Saha received the Bachelor's and Master's degrees in computer science and engineering from Jadavpur University, India, in 1987 and 1989, respectively. In 1997, he received the PhD degree from the Indian Statistical Institute, which he joined as a faculty member in 1993. In 1997, he joined the University of Pennsylvania as a postdoctoral fellow where he is currently working as a Research Associate. His present research interests include biomedical imaging problems and the application of their solutions, digital topology and its application to image processing, and estimation of trabecular bone strength from MR images. He has published a couple of dozen research papers in international journals. He received a Young Scientist award from the Indian Science Congress Association in 1996. He is a member of the IEEE and the International Association for Pattern Recognition and a member of the governing body of the Indian Unit for Pattern Recognition and Artificial Intelligence.



Jayaram K. Udupa received the PhD degree in computer science in 1976 from the Indian Institute of Science, Bangalore, with a gold medal for best research. During the past 25 years, he worked in the areas of biomedical image and signal processing, pattern recognition, biomedical computer graphics, 3D imaging, visualization, and their biomedical applications. He has published more than 100 journal papers and 80 conference papers in these areas, he has also edited two books, given over 100 invited speeches (since 1986), offered consultancy to several industries, and developed widely distributed software packages for medical three-dimensional imaging. He is at present the chief of the medical imaging section and professor of radiological sciences in the Department of Radiology at the University of Pennsylvania, Philadelphia. He is a senior member of the IEEE.

► For further information on this or any computing topic, please visit our Digital Library at <http://computer.org/publications/dlib>.