

Multiseeded Segmentation Using Fuzzy Connectedness

Gabor T. Herman, *Fellow, IEEE*, and Bruno M. Carvalho, *Student Member, IEEE*

Abstract—Fuzzy connectedness has been effectively used to segment out an object in a badly corrupted image. We generalize the approach by providing a definition which is shown to always determine a simultaneous segmentation of multiple objects. For any set of seed points, the segmentation is uniquely determined by the definition. An algorithm for finding this segmentation is presented and its output is illustrated. The algorithm is fast as compared to other segmentation algorithms in current use. We also report on an evaluation of the accuracy and robustness of the algorithm based on experiments in which several users were repeatedly asked to identify the seed points for the algorithm in a number of images.

Index Terms—Segmentation, fuzzy connectedness, feature extraction, algorithms, clustering.

1 INTRODUCTION

SEGMENTATION is the process of recognizing objects in an image. If what distinguishes these are not the exact values assigned to the pixels, but rather some textural property (as it is the case for images containing random noise and/or shading), then fuzzy connectedness can be usefully employed to achieve segmentation (see [1], [2], [3], [4], [5], [6] and their references). Fuzzy connectedness was explicitly introduced by Rosenfeld [7], but it had been foreshadowed earlier (for example, by the “Minimum Method” in [8]). Our approach is based on that advocated in [9], but is generalized to arbitrary digital spaces [10].

In this very general approach, we deal with an arbitrary set V . Because of the nature of the applications that we have in mind, we refer to elements of V as *spels*, which is short for *spatial elements* [10]. The spels can be pixels of an image (as in [1], [2], [3], [4], [5], [7], [9]), but they can also be dots in the plane (as in [11], [12]), or any variety of other things. The theory and algorithm presented here will be independent of the specifics of the application area. They are, in particular, applicable to data clustering [13] in general and so their range of usefulness goes far beyond just image segmentation and includes such distant areas of endeavor as psychology [8] and statistics [14].

The basic concept that we are generalizing in this paper is that of *fuzzy connectedness*. To every ordered pair (c, d) of spels, we assign a real number not less than 0 and not greater than 1, which we define as the fuzzy connectedness of c to d . This indeed provides us with an example of a fuzzy set (as it is normally defined in the literature [15]):

The fuzzy set in question is “the set of connected pairs” and the *grade of membership* of (c, d) in this set is the fuzzy connectedness of c to d . In the approach used below (and in earlier, already cited, papers), fuzzy connectedness is defined in the following general manner.

We call a sequence of spels a *chain*, its *links* are the ordered pairs of consecutive spels in the sequence. The strength of a link is also a fuzzy concept (i.e., for every ordered pair (c, d) of spels, we assign a real number not less than 0 and not greater than 1, which we define as the *strength of the link* from c to d). For example, if the set of spels V is a finite set of dots in the plane, we may define the strength of the link from one dot to another as the reciprocal of the distance between them (we need to make the unit of distance such that all distinct dots are at least one unit from each other). As we will see in the body of the paper, for the purpose of fuzzy segmentation of images, the strength of any link of one pixel to another can often be automatically defined based on statistical properties of the links within regions identified by the user as belonging to the object of interest. The *strength of a chain* is the strength of its weakest link. The *fuzzy connectedness* of c to d is then defined as the strength of the strongest chain from c to d .

In the literature which concerned itself with identifying a single fuzzy object containing a given seed pixel [2], [3], [4], [9], the following approach has been taken: The grade of membership in the object of an arbitrary pixel is its fuzzy connectedness to the seed pixel. In this paper, we generalize this approach: Each of the different objects in the image has its own definition of strength for the links and its own set of seed pixels. Each of the objects is then defined as the collection of those pixels which are connected entirely within the object to one of its own seed pixels in a stronger way than to any of the other seed pixels. This intuitive notion will be made precise. (An essential feature of our approach is that it does *not* simply calculate, for every pixel, the grade of membership to each of the individual objects of that pixel and then assigns the pixel to the object for which its grade of membership is maximal. The reason for this is that if a spel is separated from the seed points of Object 1 by

• G.T. Herman is with the Center for Computer Science and Applied Mathematics, Temple University, Philadelphia, PA 19122. E-mail: gabor@euclid.temple.edu.

• B.M. Carvalho is with the Center for Computer Science and Applied Mathematics, Temple University, Philadelphia, PA 19122 and also the Department of Computer & Information Science, University of Pennsylvania, Philadelphia, PA 19104-6389. E-mail: bruno@csam.temple.edu.

Manuscript received 2 Sept. 1999; revised 5 July 2000; accepted 6 Oct. 2000. Recommended for acceptance by D. Jacobs.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 110537.

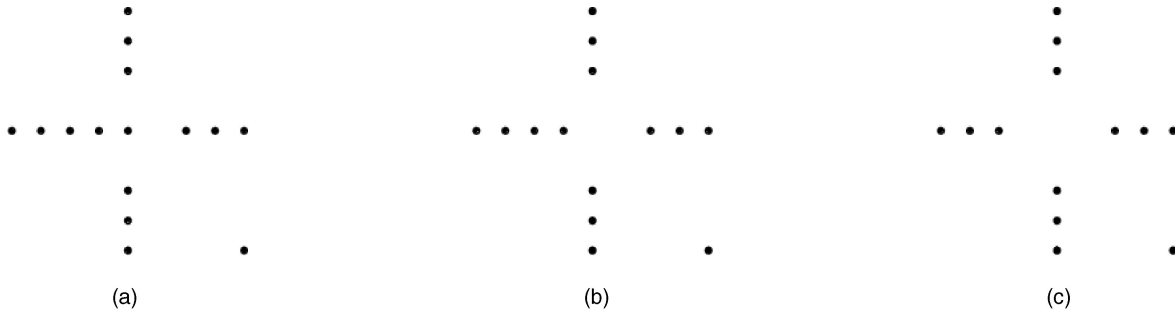


Fig. 1. Three examples of a set of spels V . In each case, the spels are dots in the plane and $V = T \cup L \cup R \cup B \cup \{o\}$, where T contains the top three dots, L contains the (a) five, (b) four, or (c) three horizontally-centered dots on the left, R contains the three horizontally-centered dots on the right, B contains the three vertically-centered dots on the bottom, and o is the dot on the bottom-right.

spels belonging to Object 2, then it should not be assigned to Object 1. The gestalt that we are trying to capture here is a segmentation in which the chains which determine “belonging to an object” must lie entirely in that object.)

An intuitive picture of our algorithm is the following: There are M competing armies (one corresponding to each object). Initially, they each have full strength and they occupy their respective seed spels. All armies try to increase their respective territories, but the moving from one spel to another reduces the strength of the soldiers to be the minimum of their strength on the previous spel and the affinity (for that army or object) between the spels. At any given time, a spel will be occupied by the soldiers of the armies which were not weaker than any other soldiers who reached that spel by that time. Eventually, a steady state is reached in which the objects occupied by each of the armies have a desirable joint property discussed in the next section.

A potentially time-consuming step in finding such objects is the calculation of the multiple fuzzy connectedness of all the pixels to the seed pixels. We devised a greedy (and, hence, reasonably efficient [16]) algorithm which provides the desired segmentation. We demonstrate its performance on five mathematically-defined and on six physically-obtained (real) images. The output of the process is a segmentation into fuzzy sets in the classical sense ([15], p. 39) that, for each spel, we also produce a “grade of membership” in the object(s) to which it belongs.

The approach we are advocating here turns out to be semiautomatic in its actual applications. We rely on the user of our method to be able to identify seed spels which definitely belong to the various objects into which we desire to segment the images and we suggest (as other advocates of segmentation based on fuzzy connectedness have done before us) that the user-selected seed spels can be used for automatic calculation of the definitions of the strengths of links of each of the objects. Since this implies that the output of our algorithm is user-dependent, we report on experiments (in which five users segmented five images, each five times) which validate the accuracy and robustness of our approach.

2 THEORY AND ALGORITHM

The discussion in this section is quite general, it is applicable to arbitrary digital spaces [10] and not to just those defined by traditional rectangular tessellations.

Some of the terms are defined in a slightly novel way. In Sections 3 and 4, we apply these ideas to images defined over both the hexagonal and the standard square tessellations of the plane. In contrast to this, in this section, we illustrate the definitions on sets of spels which are dots in the plane, see Fig. 1.

For a positive integer M , an M -semisegmentation of a set V (of spels, short for spatial element) is a function σ which maps each $c \in V$ into an $(M+1)$ -dimensional vector $\sigma^c = (\sigma_0^c, \sigma_1^c, \dots, \sigma_M^c)$, such that $\sigma_0^c \in [0, 1]$ (i.e., it is nonnegative but not greater than 1) and for at least one m in the range $1 \leq m \leq M$ $\sigma_m^c = \sigma_0^c$, and for all other m it is either 0 or σ_0^c . We say that σ is an M -segmentation if, for every spel c , σ_0^c is positive.

We illustrate these definitions for $M = 2$ using the sets V defined in Fig. 1. In all three cases, Figs. 1a, 1b, and 1c, we let $\sigma^c = (1, 1, 0)$ for $c \in T$ and $\sigma^c = (1, 0, 1)$ for $c \in L$. In other words, this means that the top three dots belong to the first object and the horizontally-centered dots on the left belong to the second object, with the maximal grade of membership in both cases. The assignment of the other dots to the two objects (and the associated grades of membership) depend, for now intuitively but later more precisely, on the fuzzy connectedness ideas outlined in the previous section. For $c \in R$, we let $\sigma^c = (1/2, 0, 1/2)$ in Fig. 1a and let $\sigma^c = (1/\sqrt{8}, 1/\sqrt{8}, 0)$ in Figs. 1b and 1c. This reflects the fact that the cluster R of dots is nearer to the cluster L of dots in Fig. 1a, but it is nearer to the cluster T of dots in the other two cases. For $c \in B$, we let $\sigma^c = (1/2, 0, 1/2)$ in Fig. 1a, $\sigma^c = (1/\sqrt{5}, 0, 1/\sqrt{5})$ in Fig. 1b, and $\sigma^c = (1/\sqrt{8}, 1/\sqrt{8}, 1/\sqrt{8})$ in Fig. 1c. The last of these assignments reflect the fact that the cluster B of dots is equidistant from the clusters L and R of dots in Fig. 1c, and so it is logical that dots in B have the same grade of membership in both objects. Finally, we may consider that the dot o is simply too far from both objects and let $\sigma^o = (0, 0, 0)$, in which case, σ is a 2-semisegmentation but not a 2-segmentation. Alternatively, we may let $\sigma^o = (1/4, 0, 1/4)$ in Fig. 1a and let $\sigma^o = (1/4, 1/4, 1/4)$ in Figs. 1b and 1c, in which case, σ is a 2-segmentation.

A fuzzy spel affinity on V is a function $\psi : V^2 \rightarrow [0, 1]$. We think of (c, d) as a link and of $\psi(c, d)$ as its ψ -strength. (In the previous literature, it was also assumed that ψ is symmetric: for all $(c, d) \in V^2$, $\psi(c, d) = \psi(d, c)$; we do not need this restriction.) We define a chain in $U (\subseteq V)$ from $c^{(0)}$ to $c^{(K)}$ to be a sequence $\langle c^{(0)}, \dots, c^{(K)} \rangle$ of spels in U

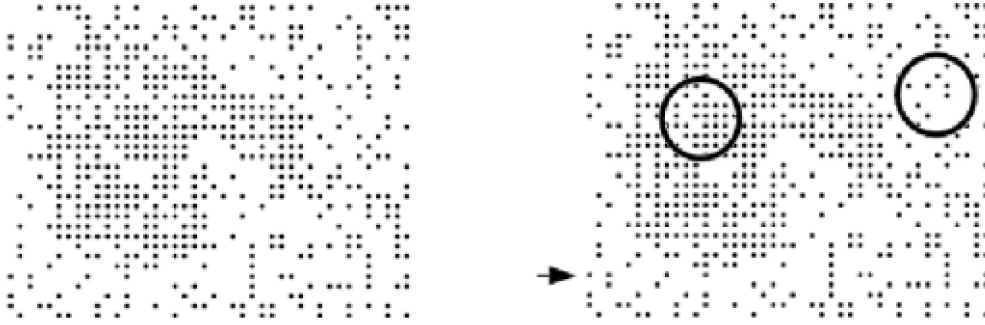


Fig. 2. A region of the plane is distinguished from its background by randomly retaining three out of every four points of a square grid within the region and randomly retaining only one out of every four points in the background (on the left). While the border between the region and its background is far from obvious, areas can be identified which are clearly in the region and in the background, respectively (see the circles on the right). The statistical properties of the distribution of the dots within these regions can then be used to (fuzzily) assign dots to the region and/or to the background (see text).

and the ψ -strength of this chain as the ψ -strength of its weakest link $(c^{(k-1)}, c^{(k)})$, $1 \leq k \leq K$. (In case $K = 0$, the ψ -strength is defined to be 1.) We say that U is ψ -connected if, for every pair of distinct spels in U , there is a chain in U of positive ψ -strength from the first spel of the pair to the second.

Again, we illustrate these ideas on the sets V of Fig. 1. Assuming that the unit of length is such that the distance between the nearest distinct points in V is 1, we can define a fuzzy spel affinity on V as any of the following:

$$\psi(c, d) = \begin{cases} 0, & \text{if } c = d, \\ 1/\|c - d\|, & \text{otherwise,} \end{cases} \quad (1)$$

where $\|c - d\|$ is the Euclidean distance between the dots c and d ,

$$\bar{\psi}(c, d) = \begin{cases} \psi(c, d), & \text{if } \|c - d\| \leq 3, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

and

$$\tilde{\psi}(c, d) = \begin{cases} 1/3, & \text{if } \|c - d\| \leq 4, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Under these definitions, any subset U of V is ψ -connected. (Indeed, for any pair (c, d) of distinct dots in U , (c, d) is a chain in U of positive ψ -strength from c to d .) On the other hand, while $L \cup R$ is $\bar{\psi}$ -connected in Fig. 1a, it is not $\bar{\psi}$ -connected in Fig. 1c. In all cases, $T \cup L \cup R \cup B$ is $\bar{\psi}$ -connected, but no subset of V which contains o and an element distinct from o is $\bar{\psi}$ -connected. The set V itself is $\bar{\psi}$ -connected, but the subset $T \cup L \cup \{o\}$ is not.

If there are multiple objects to be segmented, it is reasonable that each should have its own fuzzy spel affinity. For images, this idea has been introduced in [17] and will be further illustrated in the next section. In this section, we illustrate the idea on clusters of dots in the plane. Consider Fig. 2, which is similar in its nature to Fig. 1h of [11]. It leads to the following.

As illustrated in Fig. 2, it is far from obvious for a human being to decide exactly which dots belong to the region and which to the background, but it is reasonably easy to specify sets of dots R and B which are clearly in the region and in the background, respectively. Letting V be the set of all

dots, we define for every $c \in V$, $i(c)$ to be the Euclidean distance of c to that other dot in V , which is nearest to it. (Intuitively, $i(c)$ measures the isolation of c from the rest of the dots in V . A property which can be used to distinguish the region from its background in Fig. 2 is the isolation of the dots in it.) A subset S of V containing at least a few dots gives rise to a fuzzy spel affinity on V , defined as follows: Let r_S and s_S the mean and the standard deviation of the set of values $\{i(c) \mid c \in S\}$. Then,

$$\psi_S(c, d) = \begin{cases} 0, & \text{if } i(c) \neq \|d - c\|, \\ e^{-\frac{(r_S - i(c))^2}{2s_S^2}}, & \text{otherwise.} \end{cases} \quad (4)$$

(Note that this definition is an example in which the fuzzy spel affinity is not necessarily symmetric. This can arise because d being the nearest distinct dot in V to c does not imply that c is the nearest distinct dot in V to d .) We see that $\psi_S(c, d)$ has a nonzero value only if d is a nearest distinct dot in V to c and its value will be high if $\|d - c\|$ is typical of the distances from dots in S to their nearest distinct neighbors. Hence, ψ_R and ψ_B appear to be appropriate ways of defining the fuzzy spel affinity for the region and background, respectively, in Fig. 2.

In general, an M -fuzzy graph is a pair (V, Ψ) , where V is a nonempty finite set and $\psi = (\psi_1, \dots, \psi_M)$ with ψ_m (for $1 \leq m \leq M$) a fuzzy spel affinity such that V is $(\min_{1 \leq m \leq M} \psi_m)$ -connected. (This is defined by

$$(\min_{1 \leq m \leq M} \psi_m)(c, d) = \min_{1 \leq m \leq M} \psi_m(c, d).)$$

From this definition and the previous discussion, it follows that $(V, (\psi_1, \psi_2))$ is a 2-fuzzy graph as long as V is any of the three sets of dots in Fig. 1 and ψ_1 and ψ_2 are either the ψ of (1) or the $\tilde{\psi}$ of (3). On the other hand, if either ψ_1 or ψ_2 is the $\bar{\psi}$ of (2), then $(V, (\psi_1, \psi_2))$ is not a 2-fuzzy graph, since in such a case V is not $(\min_{1 \leq m \leq 2} \psi_m)$ -connected. Similarly, the $(V, (\psi_R, \psi_B))$ associated with Fig. 2 is not a 2-fuzzy graph. For example, if d is the dot indicated by the arrow (on the right of the figure), then for all $c \in V$ we have that $\psi_R(c, d) = \psi_B(c, d) = 0$.

For an M -semisegmentation σ of V and for $1 \leq m \leq M$, the chain $\langle c^{(0)}, \dots, c^{(K)} \rangle$ is said to be a σm -chain if $\sigma_m^{c^{(k)}} > 0$, for $0 \leq k \leq K$. Further, for $U \subseteq V$, $W \subseteq V$ and $c \in V$, we use $\mu_{\sigma, m, U, W}(c)$ to denote the maximal ψ_m -strength of a

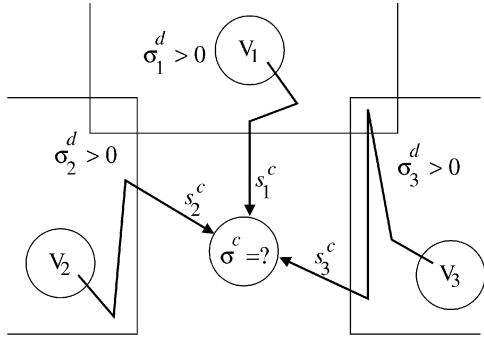


Fig. 3. Illustration of the desirability of the M -segmentation whose existence (and uniqueness) is guaranteed by the Theorem. Let c be an arbitrary spel and suppose that σ^d is known for all other spels d . Then, (for $1 \leq n \leq M$) the s_n^c of (6) is the maximal ψ_n -strength of a chain $\langle d^{(0)}, \dots, d^{(L)}, c \rangle$ from a seed spel in V_n to c such that $\sigma_n^{d^{(l)}} > 0$ (i.e., $d^{(l)}$ belongs to the n th object), for $0 \leq l \leq L$. (s_n^c is defined to be 0 if there is no such chain.) Intuitively, the n th object can “claim” that c belongs to it if and only if s_n^c is maximal. This is indeed how things get sorted out in (7): σ_m^c has a positive value only for such objects.

σm -chain in U from a spel in W to c . (This is 0 if there is no such chain.)

We now illustrate these ideas with V consisting of the dots of Fig. 1c, σ defined by

$$\sigma^c = \begin{cases} (1, 1, 0), & \text{if } c \in T, \\ (1, 0, 1), & \text{if } c \in L, \\ (1/\sqrt{8}, 1/\sqrt{8}, 0), & \text{if } c \in R, \\ (1/\sqrt{8}, 1/\sqrt{8}, 1/\sqrt{8}), & \text{if } c \in B, \end{cases} \quad (5)$$

$\sigma^o = (1/4, 1/4, 1/4)$ and $\Psi = (\psi_1, \psi_2)$ with $\psi_1 = \tilde{\psi}$ of (3) and $\psi_2 = \psi$ of (1). Since there is no $\sigma 1$ -chain containing an element of L and there is no $\sigma 2$ -chain containing an element of T , we have that, for any $U \subseteq V$, $\mu_{\sigma,1,U,L}(o) = \mu_{\sigma,2,U,T}(o) = 0$. It is easy to check that

$$\mu_{\sigma,1,T \cup B \cup \{o\},T}(o) = \mu_{\sigma,1,T \cup R \cup \{o\},T}(o) = 1/3.$$

On the other hand, due to the difference between $\psi_1 = \tilde{\psi}$ and $\psi_2 = \psi$, $\mu_{\sigma,2,L \cup B \cup \{o\},L}(o) = 1/4$. More significantly, since there is no $\sigma 2$ -chain containing an element of R , we have that $\mu_{\sigma,2,L \cup R \cup \{o\},L}(o) = 1/\sqrt{52}$.

Theorem. If (V, Ψ) is an M -fuzzy graph and, for $1 \leq m \leq M$, V_m is a subset (of seed spels) of V such that at least one of these subsets is nonempty, then there exists a unique M -semisegmentation (which is, in fact, an M -segmentation) σ of V with the following property: For every $c \in V$, if for $1 \leq n \leq M$

$$s_n^c = \begin{cases} 1, & \text{if } c \in V_n, \\ \max_{d \in V} (\min(\mu_{\sigma,n,V,V_n}(d), \psi_n(d, c))), & \text{otherwise,} \end{cases} \quad (6)$$

then for $1 \leq m \leq M$

$$\sigma_m^c = \begin{cases} s_m^c, & \text{if } s_m^c \geq s_n^c \text{ for } 1 \leq n \leq M, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Before discussing the validity of this theorem, let us discuss in less mathematical terms what it says. The property stated in the theorem is a reasonable one (see Fig. 3). Furthermore, it is a local one in the following sense:

For a fixed spel c , we can work out the values of the s_n^c using (6) and what we request that, at that spel c , (7) be satisfied. What the theorem says is that there is one, and only one, M -semisegmentation which satisfies this reasonable local property simultaneously everywhere, and that this M -semisegmentation is in fact an M -segmentation.

The three 2-segmentations that we discussed earlier in association with Fig. 1 are, in fact, examples of what is guaranteed by this theorem in general. In all three cases, the 2-fuzzy graph is $(V, (\psi, \psi))$, where V is different in the three cases (as shown in Fig. 1a, 1b, and 1c, respectively), ψ is defined by (1), $V_1 = T$ and $V_2 = L$.

At a first glance, it appears that the statement of the theorem is unnecessarily complicated in the sense that one maybe able to replace $\mu_{\sigma,n,V,V_n}(d)$ in (6) by σ_n^d . However, this altered version of the theorem would not be true, as demonstrated by the following simple example. Let $V = \{u, v, w\}$ and let $V_1 = \{v\}$. Consider the 1-fuzzy graph $(V, \{\psi_1\})$, where

$$\psi_1(c, d) = \begin{cases} 0.1, & \text{if } v \in \{c, d\}, \\ 1.0, & \text{if } v \notin \{c, d\}. \end{cases} \quad (8)$$

Consider the 1-segmentation $\sigma^v = (1, 1)$ and $\sigma^u = \sigma^w = (r, r)$. The only value of r for which this 1-segmentation satisfies the property in the theorem as stated is $r = 0.1$. However, were we to replace μ_{σ,n,V,V_n} in (6) by σ_n^d , then the property would be satisfied whenever $0.1 \leq r \leq 1.0$ and, so, the uniqueness claimed in the theorem would no longer be true.

In the previous literature, certain uniqueness results (such as Theorem 5.2.4 of [10]) have been derived from a proposition (Theorem 5.2.1 of [10] or Proposition 2.3 of [9]; less explicitly indicated earlier in [8]), which in the notation introduced in this paper says the following.

Proposition. Let $(V, \{\psi\})$ be a 1-fuzzy graph with a symmetric ψ , let σ be a 1-segmentation of V and let $0 \leq t \leq 1$. Then, the binary relation $K_{\psi,t}$ on V defined by

$$(c, d) \in K_{\psi,t} \Leftrightarrow \mu_{\sigma,1,V,\{d\}}(c) \geq t \quad (9)$$

is an equivalence relation.

It turns out that this proposition does not generalize to the case discussed in this paper in which we do not require ψ to be symmetric. This can be seen by the simple example in which $V = \{c, d\}$, $\psi(c, d) = 1.0$, and $\psi(d, c) = 0.1$. In this case, $(c, d) \notin K_{\psi,1}$ but $(d, c) \in K_{\psi,1}$. So, $K_{\psi,1}$ is not symmetric and so is not an equivalence relation and, so, the proposition is not valid in our more general context.

Nevertheless, the following argument validates the uniqueness claim stated in the theorem. Suppose that there are two different M -semisegmentations σ and τ of V having the stated property. We choose a spel c , such that $\sigma^c \neq \tau^c$, but for all $d \in V$ such that $\max(\sigma_0^d, \tau_0^d) > \max(\sigma_m^c, \tau_m^c)$, $\sigma^d = \tau^d$. Without loss of generality, we assume that $\sigma_0^c \geq \tau_0^c$, from which it follows that, for some $m \in \{1, \dots, M\}$, $\sigma_m^c > \tau_m^c (\geq 0)$ and, so, by (7), $\sigma_m^c = s_m^c$ and $c \notin V_m$. This implies that there exists a σm -chain $\langle d^{(0)}, \dots, d^{(L)} \rangle$ in V of ψ_m -strength not less than $\sigma_m^c (> 0)$ such that $d^{(0)} \in V_m$ and $\psi_m(d^{(L)}, c) \geq \sigma_m^c$. Next, we show that $\langle d^{(0)}, \dots, d^{(L)} \rangle$ is a τm -chain.

We need to show that, for $0 \leq l \leq L$, $\tau_m^{d^{(l)}} > 0$. This is true for 0, since $d^{(0)} \in V_m$. Now, assume that it is true for $l-1$ ($1 \leq l \leq L$). Since $\langle d^{(0)}, \dots, d^{(l-1)} \rangle$ is a τm -chain in V of ψ_m -strength at least $\sigma_m^c (> 0)$ from an element of V_m , we have $\mu_{\tau m, V, V_m}(d^{(l-1)}) \geq \sigma_m^c$. Since we also know that $\psi_m(d^{(l-1)}, d^{(l)}) \geq \sigma_m^c$, we get $t_m^{d^{(l)}} \geq \sigma_m^c$ (where t is defined for τ as s is defined for σ in (6)). The only way $\tau_m^{d^{(l)}}$ could be 0, is if there were an $n \in \{1, \dots, M\}$ such that

$$\begin{aligned} \max(\sigma_0^{d^{(l)}}, \tau_0^{d^{(l)}}) &\geq \tau_0^{d^{(l)}} = \tau_n^{d^{(l)}} = t_n^{d^{(l)}} \\ &> t_m^{d^{(l)}} \geq \sigma_m^c = \sigma_0^c = \max(\sigma_0^c, \tau_0^c). \end{aligned}$$

By the choice of c , this would imply that $\sigma^{d^{(l)}} = \tau^{d^{(l)}}$, which cannot be since $\sigma_m^{d^{(l)}} \neq 0$.

From all this, it follows that $\tau_0^c \geq t_m^c \geq \sigma_m^c = \sigma_0^c \geq \tau_0^c$, implying that all the inequalities are, in fact, equalities. But then $\sigma_m^c = t_m^c = \tau_m^c$, contradicting $\sigma_m^c > \tau_m^c$ and, thereby, validating uniqueness.

Next, we show that any M -semisegmentation having the stated property is, in fact, an M -segmentation. We observe that it is a consequence of (7) that, for any spel c , $\sigma_0^c = \max_{1 \leq m \leq M} \sigma_m^c$. Let $\langle c^{(0)}, \dots, c^{(K)} \rangle$ be a chain of positive ($\min_{1 \leq m \leq M} \psi_m$)-strength from a seed spel to an arbitrary spel c . We now show inductively that, for $0 \leq k \leq K$, $\sigma_0^{c^{(k)}} > 0$. Clearly, this is so for $k=0$. Now, suppose that it is so for $k-1$. Choose an m ($1 \leq m \leq M$) such that $\sigma_0^{c^{(k-1)}} = \sigma_m^{c^{(k-1)}} = \sigma_m^{c^{(k-1)}}$. Then, there is a σm -chain of positive ψ_m -strength from a spel in V_m to $c^{(k-1)}$. Since $\psi_m(c^{(k-1)}, c^{(k)}) > 0$, $\sigma_0^{c^{(k)}} \geq \sigma_m^{c^{(k)}} > 0$.

We do not give a mathematical proof of the existence of the desired M -semisegmentation; instead, we carefully describe an algorithm which produces it. In designing the algorithm, we aim at making it efficient: As is illustrated in the next section, our implementation of it allowed us to find 3-segmentations of images with over 10,000 spels in less than a fifth of a second.

Before getting into a detailed discussion of the algorithm, it is useful to point out that it does not simply run in parallel the fuzzy connectedness algorithm for the segmentation of a single object for each of the objects and then assigns the pixels at the end to the object for which the grade of membership is maximal. (Such an algorithm is discussed in [17].) We consider it more appropriate that the assignment of a spel to an object will be blocked if all chains to that spel from the seed spels of that object are broken by higher claims of other objects. We illustrate this in Fig. 1a, using the 2-fuzzy graph $(V, \{\tilde{\psi}, \psi\})$; see (3) and (1) with $V_1 = T$ and $V_2 = L$. Clearly, there is a chain in V from an element of T to o whose $\tilde{\psi}$ -strength is $1/3$, but the ψ -strength of any chain in V from an element of L to o cannot exceed $1/4$. Nevertheless, the unique 2-segmentation of our theorem assigns o to the second object. This is because all chains in V from an element of T to o contain at least one spel c such that there is a chain from L to c whose ψ -strength is greater than the $\tilde{\psi}$ -strength of any chain from T to c . In particular, for any $c \in R \cup B$, there is a chain from an element of L to c whose ψ -strength is $1/2$, but there is no chain from an element of T to o whose $\tilde{\psi}$ -strength is more than $1/3$. Our algorithm has to be subtle enough to recognize such breaking of chains for one object by a competing object.

As the algorithm proceeds, it maintains (and repeatedly changes) an M -semisegmentation σ . The claim is that at the time when the algorithm terminates, σ satisfies the property of the theorem. The omitted mathematical details of the proof of this claim resemble those needed to prove the correctness of the simpler Dijkstra's algorithm [16].

The algorithm makes use of a priority queue H of spels c , with associated keys σ_0^c [16]. Such a priority queue has the property that the key of the spel at its head is maximal (its value is denoted by $\text{Maximum-Key}(H)$, which is defined to be 0 if H is empty). As the algorithm proceeds, each spel is inserted into H exactly once (using the operation $H \leftarrow H \cup \{c\}$) and is eventually removed from H (using the operation $\text{Remove-Max}(H)$, which removes the spel c from the head of the priority queue). At the time when a spel c is removed from H , the vector σ^c has its final value. Spels are removed from H in a nonincreasing order of the final value of σ_0^c . We use the variable l to store the current value of $\text{Maximum-Key}(H)$.

The process is initialized (Steps 1-9 below) by setting σ_m^c to 0, for each spel c and $0 \leq m \leq M$. Then, for every seed spel $c \in V_m$, c is put into H and both σ_0^c and σ_m^c are set to 1. Following this, l is also set to 1.

At the end of the initialization, the following conditions are satisfied:

1. σ is an M -semisegmentation of V .
2. A spel c is in H if and only if $0 < \sigma_0^c \leq l$.
3. $l = \text{Maximum-Key}(H)$.
4. For $1 \leq m \leq M$, $V_m = \{c \in H \mid \sigma_m^c = l\}$.

The initialization is followed by the main loop of the algorithm. At the beginning of each execution of this loop, Conditions 1 to 4 above are satisfied. The main loop is repeatedly performed for decreasing values of l until l becomes 0, at which time the algorithm terminates (Step 10). There are two parts to the main loop, each of which has a very different function.

The first part of the main loop (Steps 11-23 below) is the essential part of the algorithm. It is in here, where we update our best guess so far of the final values of the σ_m^c . A current value is replaced by a larger one if it is found that there is a σm -chain from a seed spel in the initial V_m to c of ψ_m -strength greater than the old value (the previously maximal ψ_m -strength of the known σm -chains of this kind) and it is replaced by 0 if it is found that (for an $n \neq m$) there is a σn -chain from a seed spel in the initial V_n to c of ψ_n -strength greater than the old value of σ_m^c .

The purpose of the second part of the main loop (Steps 24-28 below) is to restore the satisfaction of Conditions 3 and 4 above for a new (smaller) value of l .

We now give a detailed specification of the algorithm using the conventions adopted in [16].

1. **for** $c \in V$
2. **do for** $m \leftarrow 0$ **to** M
3. **do** $\sigma_m^c \leftarrow 0$
4. $H \leftarrow \emptyset$
5. **for** $m \leftarrow 1$ **to** M
6. **do for** $c \in V_m$
7. **do if** $\sigma_0^c = 0$ **then** $H \leftarrow H \cup \{c\}$
8. $\sigma_0^c \leftarrow \sigma_m^c \leftarrow 1$

```

9.  $l \leftarrow 1$ 
10. while  $l > 0$ 
11.   for  $m \leftarrow 1$  to  $M$ 
12.     do while  $V_m \neq \emptyset$ 
13.       do remove a spel  $d$  from  $V_m$ 
14.        $C \leftarrow \{c \in V \mid \sigma_m^c < \min(l, \psi_m(d, c)) \text{ and } \sigma_0^c \leq l\}$ 
15.       while  $C \neq \emptyset$ 
16.         do remove a spel  $c$  from  $C$ 
17.          $t \leftarrow \min(l, \psi_m(d, c))$ 
18.         if  $l = t$  and  $\sigma_m^c < l$  then
19.            $V_m \leftarrow V_m \cup \{c\}$ 
20.         if  $\sigma_0^c < t$  then
21.           if  $\sigma_0^c = 0$  then  $H \leftarrow H \cup \{c\}$ 
22.           for  $n \leftarrow 1$  to  $M$ 
23.             do  $\sigma_n^c \leftarrow 0$ 
24.           if  $\sigma_0^c \leq t$  then  $\sigma_0^c \leftarrow \sigma_m^c \leftarrow t$ 
25.       while Maximum-Key( $H$ ) =  $l$ 
26.       Remove-Max( $H$ )
27.        $l \leftarrow$  Maximum-Key( $H$ )
28.   for  $m \leftarrow 1$  to  $M$ 
29.      $V_m \leftarrow \{c \in H \mid \sigma_m^c = l\}$ 

```

To help with the understanding of why this algorithm performs as desired, we comment that just prior to entering its main loop (Steps 10-28), there are four kinds of spels. There are those spels d which have previously been put into and have subsequently been removed from H , for these spels, not only does the vector σ^d has its final value, but also we have already put into H (and possibly even have already removed from H) every spel c such that $\psi_m(d, c) > 0$, for at least one m . (For spels of this first kind, $\sigma_0^d > l$.) Second, there are the spels d which are in at least one of the V_m , for these spels the vector σ^d has its final value, but we may not have yet put into H every spel c such that $\psi_m(d, c) > 0$ for at least one m . (For spels of this second kind, $\sigma_0^d = \sigma_m^d = l$.) This will get done in the next execution of Steps 12-20, while Steps 21-23 will insure that σ^c get updated appropriately. Consequently, the spels c which are in H but not in any of the V_m are those for which there is a σm -chain (for the current σ) from a seed spel in the initial V_m to c ; for the rest of the spels (those which have not as yet been put into H) there is no σm -chain (for the current σ) from a seed spel in the initial V_m to c . (For spels c of the third and fourth kinds, $0 < \sigma_0^c < l$ and $0 = \sigma_0^c$, respectively.)

One tricky aspect of the algorithm is that, as a result of Step 23, a spel of the third kind may become of the second kind and a spel of the fourth kind may become of the third (or even of the second) kind during the execution of the main loop. That the description of the four kinds of spels remains as given in the previous paragraph is insured by Steps 18 and 20. (Step 20 also insures that Condition 2 stated above the algorithm remains satisfied. To see this, observe that Step 14 guarantees that if c is put into C , then $0 < \min(l, \psi_m(d, c))$ and, consequently, the t defined in Step 17 and used in Step 23 is also positive. That Condition 1 stated above the algorithm remains satisfied is obvious from Steps 19-23.)

We complete this section with a brief discussion of our implementation of the algorithm. As suggested in [16], we

use a heap to implement the priority queue H . This provides us with efficient implementations of the operations of insertion into ($H \leftarrow H \cup c$) and removal from (Remove-Max(H)) the priority queue, as well as of Step 28. In applications it is typically the case that, for every spel d , there is at most a fixed number of spels c such that $\sum_{m=1}^M \psi_m(d, c) > 0$ and a list of all such c is inexpensive to produce. In such a case, the cost of executing Step 14 becomes proportional to a constant (four or six in the examples of the next section) independent of the size of V . Using L to denote this constant, the computational complexity of the algorithm is $O(N(\log N + ML))$, where N is the number of elements in V .

3 APPLICATION TO MULTISEEDED IMAGE SEGMENTATION

One of the beauties of the segmentation approach above is that in many applications, the appropriate fuzzy spel affinities can be automatically created by a computer program, based on some minimal information supplied by a user [4], [9], [10]. We demonstrate this in Fig. 4.

On the left are images defined on a V consisting of regular hexagons which are inside a large hexagon (with 60 spels on each side, a total of 10,621 spels). In all the examples of Fig. 4 $M = 3$. We now describe one of the possible ways of specifying ψ_m and V_m ($1 \leq m \leq 3$). The user points (and clicks) at some spels in the image and V_m is formed by these spels and the six neighbors of each. We define g_m to be the mean and h_m to be the standard deviation of the gray values of the spels of V_m and we define a_m to be the mean and b_m to be the standard deviation of the absolute differences between gray values of pairs of adjacent spels in V_m . Then, $\psi_m(c, d)$ is defined to be 0 if c and d are not adjacent and to be $[\rho_{g_m, h_m}(g) + \rho_{a_m, b_m}(a)]/2$ if they are, where g is the mean and a is the absolute difference of the gray values of c and d and the function $\rho_{r,s}(x)$ is obtained from a Gaussian distribution with mean r and standard deviation s by multiplying it by a constant so that the peak value becomes 1.

The right column of Fig. 4 shows the resulting maps of the σ_m , for $1 \leq m \leq 3$. The seed sets V_m consist of the brightest spels. (For the first image, we selected the seed spels so that $V_1 = V_2$, for the second image, we selected the seed spels so that $V_2 = V_3$, and for the rest of the images, the three sets of seed spels are pairwise disjoint, which happens to result—due to the large number of gray levels used in the images to be segmented—in the three objects being pairwise disjoint as well). The time taken to calculate these three maps using our algorithm on a 450 MHz Pentium III personal computer were between 180 ms and 190 ms for each of the seven images (average = 185.71 ms). Since these images contain 10,621 pixels, we conclude that the execution time is less than 20 microseconds per pixel. The same was true for all the other image segmentations that we tried, some of which are reported in what follows.

To permit comparisons with other algorithms, we also applied our algorithm to a range of real images which appeared in the recent image segmentation literature. Since in all these images V consists of squares inside a rectangular

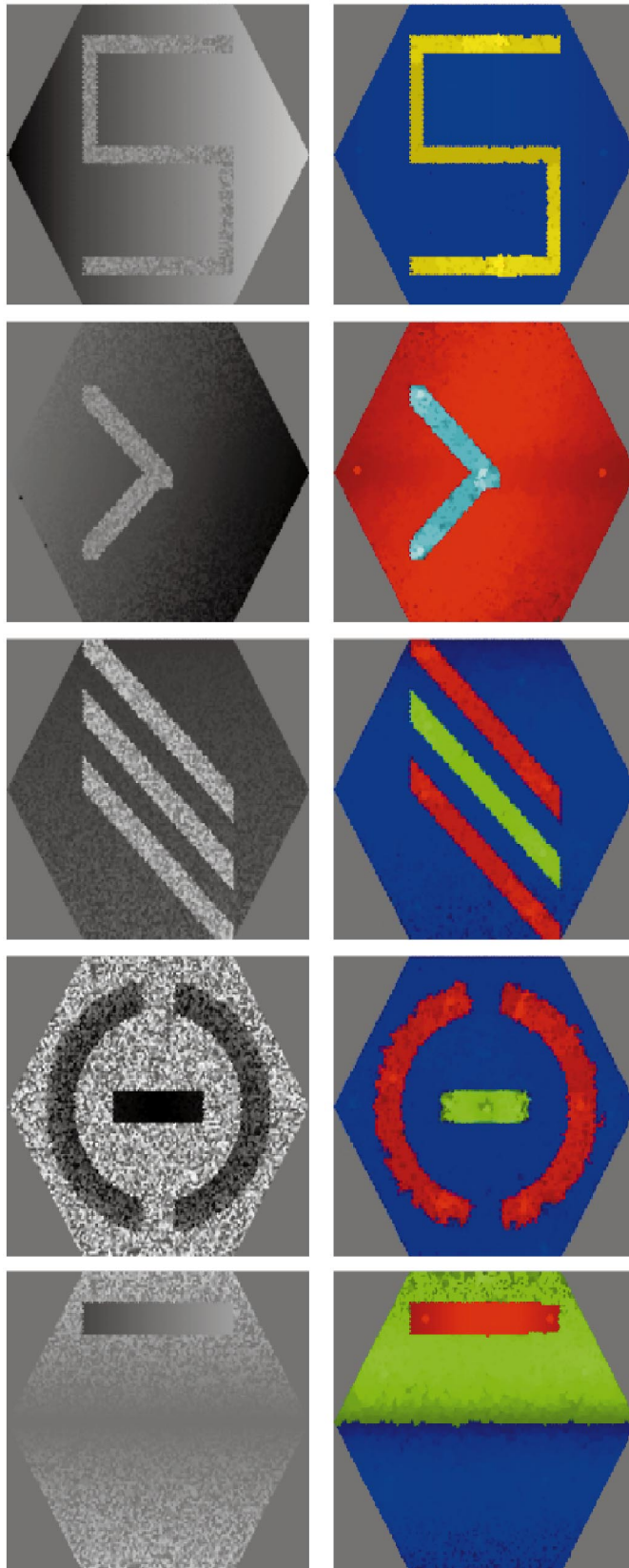


Fig. 4. The first five of the images on the left were mathematically defined, so that they include various levels of noise and shading. The last two images on the left were obtained using magnetic resonance imaging (MRI) of heads of patients. On the right, we show the corresponding three maps (obtained by the method described in the text) by assigning the color $(r, g, b) = 255 \times (\sigma_1^c, \sigma_2^c, \sigma_3^c)$ to the spel c . Note that not only the hue, but also the brightness of the color is important: The less brightly colored red areas for the last two images correspond to the ventricular cavities in the brain, correctly reflecting a low grade of membership of these spels in the object which was defined by seed spels which are all in brain tissue.

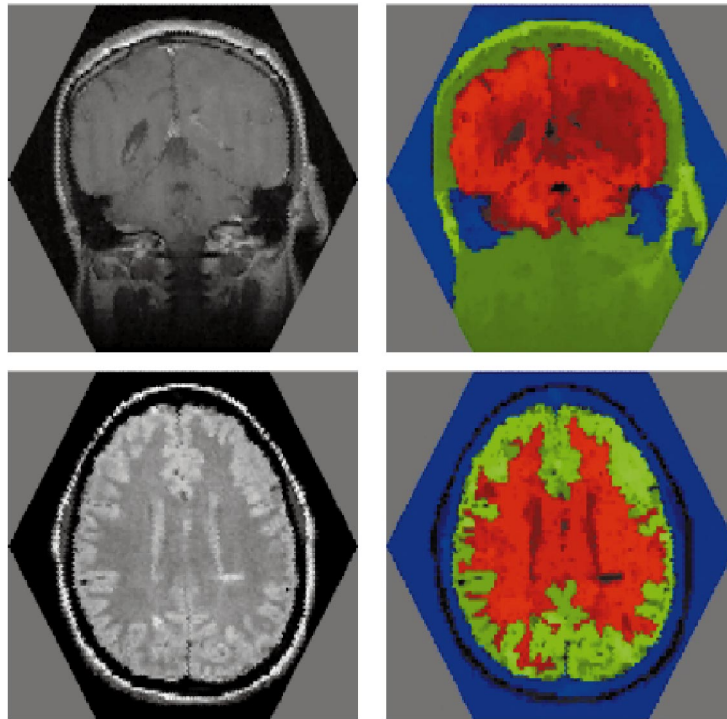


Fig. 4. (continued).

region, the automatic calculation of the fuzzy spel affinities that is described near the beginning of this section (for regular hexagons inside a hexagonal region) has to be adapted. In our adaptation, we still defined adjacency to mean that the squares have exactly one edge in common (the so-called 4-adjacency [10]), but the sets of seeds are formed by the squares at which the user points together with the eight other squares which share an edge or a vertex with that square. Except for this adaptation, the previous specification is verbatim what we used for the experiments which we now describe.

In [18], Pollak et al. demonstrate their proposed technique to segment an SAR image of trees and grass

(their Fig. 1, our Fig. 5a). They point out that “the accurate segmentation of such imagery is quite challenging and in particular cannot be accomplished using standard edge detection algorithms.” They validate this claim by demonstrating how the algorithm of [19] fails on this image. As illustrated in Fig. 5b, our technique produces a satisfactory segmentation. In this image, the computer time needed by our algorithm was 0.6 seconds (on the same 450 MHz Pentium III personal computer which we used for all our experiments), while according to a personal communication from Pollak, the method of [18] “took about 50 seconds to reach the 2-region segmentation for this 201-by-201 image on Sparc 20, with the code written in C.”

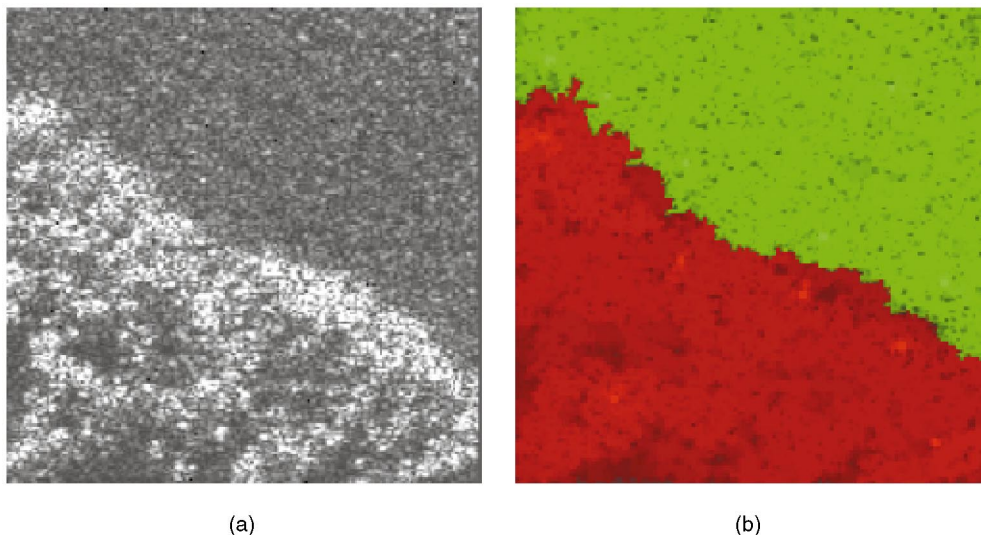
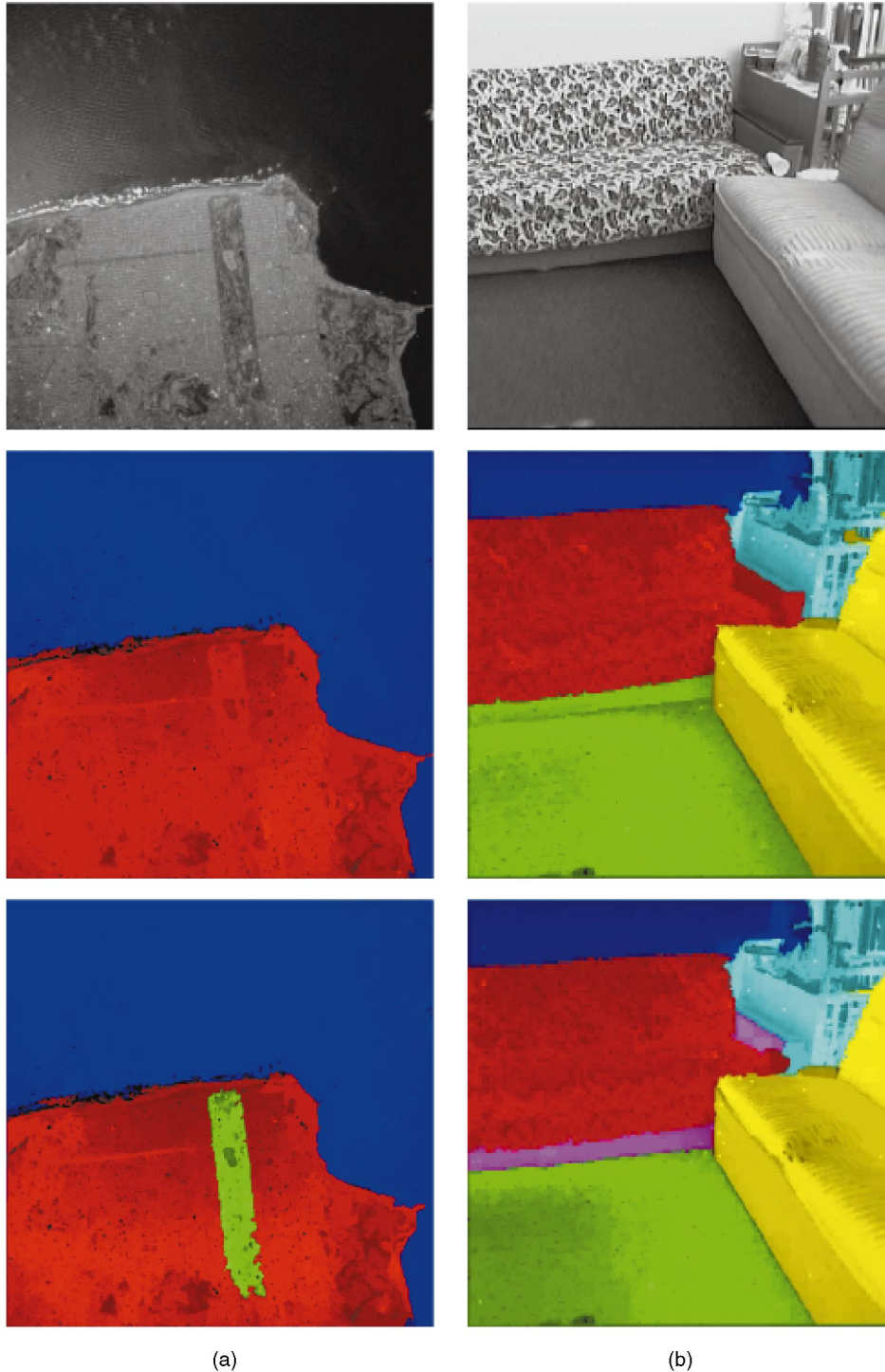


Fig. 5. An SAR image of (a) trees and grass and (b) its 2-segmentation.



(a)

(b)

Fig. 6. An aerial image of (a) San Francisco and (b) an indoor image of an office. The original images (from [20]) are on the top, the other images are multiseeded segmentations produced by our approach (see text).

In Fig. 6, we report on the results of applying our approach to two physically-obtained images from [20]: One (on the top-left) is an aerial image of San Francisco and the other (on the top-right) is an indoor image of an office. The 2-segmentation of San Francisco is into land and sea, the 3-segmentation separates out Golden Gate Park from the rest of the land. For the segmentation of the office scene (as well as the segmentation in Fig. 7), $3 < M < 7$ and, so, we used a different colored hue for each object with gray for

spels which belong to multiple objects; the lightness of the color indicates the grade of membership. In the 5-segmentation shown in the middle, the bottom of the sofa on the left is merged into the same object as the carpet on the floor. Indeed, this also happens with the unsupervised texture segmentation reported in [20]. However, while separating these two regions with an unsupervised texture segmenter seems an unreasonable task (the regions are very similar texturally, it is only our general knowledge

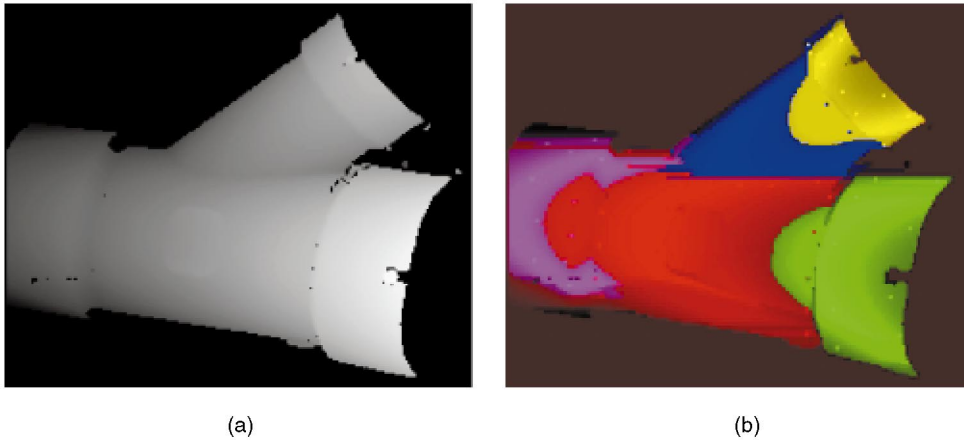


Fig. 7. (a) A range image and (b) its segmentation.

of the world which tells us that one should belong to the sofa), our method allows us to introduce an extra object. In the 6-segmentation (on the bottom-right), the bottom and lower-part of the arm of the sofa are detected as a separate object.

It is stated in [20] that the times needed for the segmentations reported “are in the range of less than five seconds” (on a Sun UltraSparc). Our times to obtain the segmentations shown in Fig. 6 are all between 3.8 and 4.2 seconds. However, there is a basic difference in the resolutions of the segmentations. Since the segmentation method used in [20] is texture based, the original 512×512 images are subdivided into 64×64 “sites” using a square window of size 8×8 per site. In the final segmentations of [20], all pixels in a particular window are assigned to the same object. As opposed to this, in our segmentations, any pixel can be assigned to any object. Another way of putting this is that we could also make our spels to be the 8×8 windows of [20] and thereby reduce the size of the V to be a 64th of what it is currently. This should result in a two order of magnitude speedup in the performance of our segmentation algorithm (at the cost of a loss of resolution in the segmentations to the level used in [20]).

Our final image from the recent literature is on the left of Fig. 7. It is a range image from [13]. As can be seen from the segmentation on the right of Fig. 7, this is an example in which the simple approach proposed in this section for calculating the fuzzy spel affinities from the gray values of the pixels fails: The fuzzy segmentation that is produced (in less than 2 seconds) is not the correct one (based on higher-order knowledge of the object under consideration). The “final segmentation” shown in Fig. 27d of [13], while not perfect, is in much better correspondence with human intuition. However, for the purpose of segmentation, the method of [13] starts with assigning to each pixel a six-dimensional feature vector (consisting of the estimated spatial coordinates of the surface point shown at that pixel, together with the coordinates of the estimated unit normal to the surface at that point). The method described in this section for calculating the fuzzy spel affinities based on Gaussian distributions determined by the gray values of the seed pixels is easily expandable to a method for calculating the fuzzy spel affinities based on (multivariate) Gaussian distributions determined by the feature vectors of the seed

pixels. The algorithm given in the previous section is then applicable, without any changes, to achieve segmentation using fuzzy spel affinities defined in this more sophisticated way. In brief, while the simple approach to calculating the fuzzy spel affinities proposed in this section is not guaranteed to lead to a desired segmentation in all cases, the general algorithm of the last section has a more universal appeal and its practical usefulness is only limited by our ability to find the definitions of the fuzzy spel affinities which are appropriate to the application at hand.

4 ACCURACY AND ROBUSTNESS

Since in the method discussed in the last section, the affinities are based on manually selected seeds, it follows that (in spite of the uniqueness property expressed in our theorem) the practical performance of the algorithm as implemented need to be experimentally evaluated, both for accuracy and for robustness. In this section, we report on such experiments and their results. In all these experiments, we used the first five images on the left of Fig. 4. Since all these images were based on mathematically defined shapes (onto which we imposed gray values, followed by noise and shading), the “correct” segmentation of the images were known to us (but not to the users who were the subjects of our experiments).

An interactive program was created whose first panel had the following instructions:

- You will be asked to select four seeds for each object in the image by clicking at some point on the image and then clicking the button [Set Spel]. The program will not let you set a spel that does not belong to the object in question, issuing a warning.
- The images can have two or three objects and each one will have its associated window explaining which regions correspond to the objects to be detected and their correspondent numbers. After selecting the seeds you should click the button [Run MFuzzy Alg].
- After each segmentation is completed, click on the button [Next Image] to go to the next image. After the fifth image is segmented, you can click either [Quit] or [Next Image] to exit the program.

TABLE 1
Point Accuracy of Segmentations

	Image 1	Image 2	Image 3	Image 4	Image 5	Average
User 1	91.44 \pm 14.95	99.99 \pm 0.02	99.87 \pm 0.11	97.08 \pm 0.63	91.77 \pm 4.66	96.03 \pm 7.46
User 2	99.53 \pm 0.53	100.00 \pm 0.01	99.18 \pm 0.84	96.68 \pm 0.49	91.74 \pm 4.95	97.42 \pm 3.75
User 3	99.60 \pm 0.41	100.00 \pm 0.00	99.67 \pm 0.35	95.10 \pm 2.82	88.58 \pm 4.27	96.59 \pm 4.95
User 4	99.70 \pm 0.22	100.00 \pm 0.00	99.16 \pm 1.01	96.83 \pm 0.65	92.52 \pm 4.02	97.64 \pm 3.33
User 5	99.69 \pm 0.10	100.00 \pm 0.00	99.03 \pm 0.96	96.27 \pm 0.73	95.42 \pm 4.40	98.08 \pm 2.67
Average	97.99 \pm 6.96	100.00 \pm 0.01	99.38 \pm 0.76	96.39 \pm 1.45	92.01 \pm 4.65	97.15 \pm 4.72

- You must close this window to continue, but you can move the specific instructions windows and keep them while selecting the seeds. Please follow the instructions on those windows.
- Thank you for your help.

After these, the five images were presented (always in a newly-calculated random order), each accompanied by a set of instructions. For example, the instructions for the first and fifth image from the left of Fig. 4 were the following:

- The object number 1 is the *s* shaped object and the varying background is the object number 2.
- The object number 1 is the rectangular area of varying intensity, object number 2 is the bottom part of the varying background and object number 3 is the upper part of the varying background.

Five volunteer users, who were not familiar with the images, were asked to work five times each with this interactive program. Since, in each session with the program, all five images were processed, this resulted in a total of 125 segmentations. We now report on the various ways we analyzed these 125 outputs.

We considered two reasonable ways of measuring the accuracy of the segmentations: In one, we simply consider if the spel is assigned to the correct object, in the other, we

take into consideration the grade of membership as well. The *point accuracy* of a segmentation is defined as the number of spels correctly identified divided by the total number of spels multiplied by 100. The *membership accuracy* of a segmentation is defined as the sum of the grades of membership of all the spels which are correctly identified divided by the total sum of the grades of membership of all spels in the segmentation multiplied by 100.

In Tables 1 and 2, we report on our results. The entries in these tables (and in later ones) are to be interpreted as follows: If the entry in the cell labeled User *i* and Image *j* is $x \pm y$ that means that x is the mean and y is the standard deviation of the accuracies of the five segmentations of Image *j* by User *i* which were obtained during the five separate sessions. The Average entries for a particular user report on the mean and standard deviation of the accuracies of the 25 segmentations performed by that user, while the Average entries for a particular image report on the mean and standard deviation of the accuracies of the 25 segmentations of that image. Finally, the entry in the cell labeled Average/Average reports the mean and standard deviation of the accuracies of all the 125 segmentations.

The first thing to note about these tables is that they are rather similar. This is reassuring, since the definitions of both of the accuracies were somewhat ad hoc and, so, the

TABLE 2
Membership Accuracy of the Segmentations

	Image 1	Image 2	Image 3	Image 4	Image 5	Average
User 1	92.62 \pm 12.11	99.99 \pm 0.01	99.91 \pm 0.07	97.27 \pm 0.52	94.45 \pm 3.44	96.85 \pm 5.95
User 2	99.60 \pm 0.35	100.00 \pm 0.01	99.50 \pm 0.50	96.92 \pm 0.54	93.88 \pm 3.07	97.98 \pm 2.70
User 3	99.59 \pm 0.39	100.00 \pm 0.00	99.76 \pm 0.27	95.69 \pm 2.23	90.88 \pm 5.11	97.18 \pm 4.27
User 4	99.70 \pm 0.19	100.00 \pm 0.00	99.45 \pm 0.66	97.14 \pm 0.50	94.59 \pm 3.19	98.18 \pm 2.50
User 5	99.71 \pm 0.06	100.00 \pm 0.00	99.39 \pm 0.60	96.68 \pm 0.63	95.84 \pm 4.32	98.32 \pm 2.51
Average	98.25 \pm 5.72	100.00 \pm 0.01	99.60 \pm 0.48	96.74 \pm 1.17	93.93 \pm 3.94	97.70 \pm 3.82

TABLE 3
Intrauser Consistency of the Segmentations

	Image 1	Image 2	Image 3	Image 4	Image 5	Average
User 1	86.32 \pm 13.70	99.98 \pm 0.02	99.89 \pm 0.07	97.55 \pm 0.83	90.43 \pm 4.96	94.84 \pm 8.37
User 2	99.52 \pm 0.47	99.99 \pm 0.01	99.27 \pm 0.48	97.52 \pm 1.04	90.07 \pm 6.46	97.28 \pm 4.68
User 3	99.48 \pm 0.56	100.00 \pm 0.00	99.69 \pm 0.26	94.07 \pm 3.96	86.76 \pm 7.50	96.00 \pm 6.32
User 4	99.72 \pm 0.24	100.00 \pm 0.00	99.16 \pm 0.66	97.72 \pm 0.95	94.69 \pm 2.67	98.26 \pm 2.33
User 5	99.85 \pm 0.06	100.00 \pm 0.00	99.20 \pm 0.53	97.96 \pm 0.63	93.15 \pm 4.93	98.03 \pm 3.35
Average	96.98 \pm 7.97	99.99 \pm 0.01	99.44 \pm 0.53	96.97 \pm 2.37	91.02 \pm 6.00	96.88 \pm 5.56

fact that they yield similar results indicates that the reported figures of merit are not oversensitive to the precise nature of the definition of accuracy.

The second thing to note is that there is an outlier in these tables, namely, the entry for User 1 and Image 1. The reason for this is that in one of the sessions, User 1 placed all the seeds for the background in Image 1 more or less vertically aligned, resulting in a fuzzy spel affinity for the background which did not properly take into account the horizontal shading across the background. This resulted in much of the background being assigned to the object containing the seeds which were placed into the s-shape. In any practical application, this error would have been immediately obvious to the user (since the segmentation appears within a second of entering the seeds); however, in our automated experiment, there was no option given to the users to correct any errors in seed placement. This single uncorrected error makes all our average results worse than they would be in an actual application.

Nevertheless, the results are still quite impressive. The average *error* (defined as “100 less point accuracy”) over all segmentations is less than 3 percent, the average errors for the first four images are all less than 4 percent, while for the fifth one it is less than 8 percent. (The difficulty that the users had with the fifth image is the following: In the “correct” segmentation, the upper and lower halves of the image were separated at exactly half way. Since the users were not informed of this fact, there was no reason why they would have placed the seeds symmetrically above and below the horizontal central line and this resulted in the two halves being separated in the segmentations somewhat off center.) This compares quite favorably with the state-of-the-art: In [20], Hoffmann et al. report that a “mean segmentation error rate as low as 6.0 percent was obtained.”

In order to measure the robustness of our procedure, we need to define the similarity of two segmentations. The *point similarity* of two segmentations is defined as the number of spels which are assigned to the same object in the two segmentations divided by the total number of spels multiplied by 100. The *membership similarity* of two segmentations is defined as the sum of the grades of memberships (in both segmentations) of all the spels which are assigned to the same object in the two segmentations divided by the total sum of the grades of membership (in both segmentations) of all the spels multiplied by 100. (Note that, for both these measures of similarity, identical segmentations will be given the value 100 and segmentations in which every spel is assigned to a different object in the two segmentations will be given the value 0.) Because of the likeness of Tables 1 and 2, in what follows we do not report on results based on both of these definitions of similarity, but rather select the one which appears to us the more appropriate.

In Table 3, we report on the consistency of the users with themselves. Since User i segmented Image j five times, there are 10 possible ways of pairing these segmentations. In the cell labeled User i and Image j of Table 3, we report on the mean and standard deviation of the membership similarity for these 10 pairs of segmentations. Again, the results are quite satisfactory.

In Table 4, we report on the consistency between users. In order to do this, for each user and each image, we selected the *most typical segmentation* by that user of that image. This is defined as that segmentation for which the sum of membership similarities between it and the other four segmentations by that user of that image is maximal. This way, we obtained five segmentations for each Image j (the most typical ones for each of the five users) and again

TABLE 4
Interuser Consistency of the Segmentations

Image 1	Image 2	Image 3	Image 4	Image 5	Average
99.00 \pm 1.24	100.00 \pm 0.00	99.81 \pm 0.10	98.17 \pm 0.61	96.59 \pm 1.59	98.71 \pm 1.55

TABLE 5
Consistency of the Segmentations under Object Merging

	Image 3	Image 5	Average
User 1	99.90 \pm 0.13	98.44 \pm 2.80	99.17 \pm 2.02
User 2	99.96 \pm 0.04	98.03 \pm 2.98	98.99 \pm 2.23
User 3	99.91 \pm 0.11	98.39 \pm 2.00	99.15 \pm 1.56
User 4	99.84 \pm 0.23	99.90 \pm 0.07	99.87 \pm 0.16
User 5	99.91 \pm 0.16	99.03 \pm 0.89	99.47 \pm 0.76
Average	99.90 \pm 0.14	98.76 \pm 2.01	99.33 \pm 1.52

there are 10 possible pairings of these five segmentations. In the cell labeled Image j of Table 4, we report on the mean and standard deviation of the membership similarity for these 10 pairs of segmentations. Note that the interuser consistency is even better than the intrauser consistency, mainly due to the selection of most typical segmentation for each user we eliminate the influence of the relatively bad segmentations.

Finally, we did some calculations of the sensitivity of our approach to M (the predetermined number of objects in the image). In the third image, the distinction between the red and green objects and in the fifth image between the green and blue objects is artificial; the nature of the regions assigned to these objects is the same (see Fig. 4). The question arises: If we merge these two objects into one, do we get a similar 2-segmentation as what would be obtained by merging the seed points associated with the two objects into a single set of seed points and then applying our algorithm? (This is clearly a desirable robustness property of our approach.) Table 5 reports on the analysis of doing just this for the total of 50 readings by our five users on Images 3 and 5.

5 DISCUSSION

In this paper, we provided an algorithm for the segmentation of multiple objects which combines 1) very general applicability, 2) ease of use, and 3) rapid response. In fact, items 2 and 3 are interrelated: The rapid response provides the user with immediate feedback which makes possible interactive correction of the information provided to the program.

What is the basic principle here? The user is presented with some image, such as the dots of Fig. 2 or the gray level pictures of Fig. 4. Typically, for a human endowed with a general understanding of images (or, better yet, for an expert knowledgeable in the field of endeavor which gave rise to the images), it will be easy to identify regions or points in the images which definitely belong to different objects which are to be distinguished from each other. The algorithm proposed here produces a rapid feedback showing the consequences of the user-provided identification. At this moment, the user may accept the resulting

segmentation or may redefine the identifications provided to the program in view of the (presumably undesirable) nature of what the program has returned. (For example, one may reasonably argue that there should be additional pixels assigned to the red region on the lower-right part of the last image of Fig. 4. Such a situation can be remedied by introducing additional seed spells.) In fields where we do not have, statistically-speaking, a complete knowledge of the distribution of the images which we are likely to encounter (and this is more the rule than the exception), such interactive approach is likely to be as reliable (and often less misleading) than approaches using some idealized model of the world (based on incomplete knowledge).

A basic potential weakness of our approach is that the outcome may be more dependent on the actions of the user than on the image being segmented. The experiments that we have reported in the previous section indicate that this is not the case. A real fault of the approach, as opposed to unsupervised segmentation approaches (such as reported in [20]), is that since it makes essential use of human input, it cannot be used in autonomous robotics.

The main result of the paper (referred to as the Theorem) and the accompanying algorithm is very general. In one sense, it cannot possibly fail: It will do exactly what is claimed. On the other hand, in some specific instances, it may produce undesirable results (see Fig. 7). However, one can never say that such a thing has been a consequence of the general approach: The possibility always exists that the user made an inappropriate choice of the affinities (or, at a higher level, in the general way of calculating the affinities from the user provided information). For example, the affinities calculated based on (4) are clearly inappropriate for segmenting the dot clusters of Fig. 2, since they do not even provide us with a 2-fuzzy graph (and, so, the Theorem is not applicable). However, the situation can easily be remedied, by redefining the affinities so that they are no longer zero-valued for all but the very nearest neighbors. Such an adjustment of the details of a general approach to specific applications is the norm in the field of segmentation; see how this is done for detecting gestalt clusters in [11] and it is the essence of Bayesian optimization approaches based on Markov Random Fields (see, e.g., [21]) which mandate the selection of a Gibbs distribution as

a prior specific to the field of application. It is hard to compare two such general approaches from the point of view of accuracy since, in both cases, there are so many user dependent details of an actual implementation that any single set of choices cannot be generalized to statements about the overall approach. There is one thing that we can say with confidence (based on much computational experience) about our approach, as opposed to approaches involving stochastic optimization (such as [21]): Our approach is faster. Although we do not have a similar personal experience with other approaches (such as those advocated in [22]), our reading of the reported results in the literature imply to us that our proposed algorithm is very fast (in computational time) as compared to most of the previous art.

The proposed approach belongs to the general class of graph-theoretic clustering methods. There is a large literature on these; in addition to the already cited papers, see [23], [24]. However, the literature prior to [17] assumed that a single labeled graph is to be used. In our paper, we have demonstrated that the extension of such approaches to what we have defined as an M -fuzzy graph is useful; combined also with the newly-introduced notion of M -semisegmentation, it provides both a strong theory and a promising new algorithm for clustering.

It has been argued in [8] that a good clustering algorithm should satisfy three particular properties. We now state these properties, as they have been succinctly rephrased in [11], and discuss how our algorithm relates to them.

1. "Input data should consist solely of a point set and a matrix of similarities." The point set here corresponds to our V and, due to our generalization, the single matrix of similarities is replaced by M fuzzy affinities. However, in addition, we do need some seed spels. What's good and bad about this additional input requirement has been discussed earlier in the section. We also point out that in our implementation, the "matrix of similarities" are not given explicitly, rather, we calculate its entries (the fuzzy affinities) if and when they are needed (in Step 14 of the algorithm). The practical computational complexity of our algorithm is less than the $O(MN^2)$ required to compute M "matrices of similarities."
2. "The method should be such that a clear, explicit, and intuitive description of what the clustering accomplishes is possible." Such a description is embedded in our theorem and the discussion which follows it (including Fig. 3).
3. "The method should be invariant under monotone transformations of similarity measure." We now make precise in our context, the strong sense in which this property is satisfied by our algorithm. Let ϕ be a strictly monotone mapping of $[0, 1]$ onto $[0, 1]$. If $(V, (\psi_1, \dots, \psi_M))$ is an M -fuzzy graph, then so is $(V, (\phi\psi_1, \dots, \phi\psi_M))$. (By definition, $\phi\psi_m(c, d) = \phi(\psi_m(c, d))$.) Furthermore, if for given seed spels σ and τ are the M -segmentations whose

existence and uniqueness are guaranteed by the theorem for

$$(V, (\psi_1, \dots, \psi_M))$$

and

$$(V, (\phi\psi_1, \dots, \phi\psi_M)),$$

respectively, then $\tau_m^c = \phi(\sigma_m^c)$ for all $c \in V$ and $0 \leq m \leq M$. We forego providing the details of the quite easy proof.

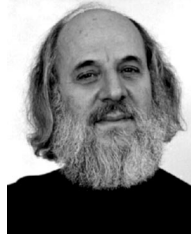
ACKNOWLEDGMENTS

This research has been supported by NIH Grant HL28432 (GTH and BMC), US National Science Foundation Grant DMS96122077 (GTH), and CAPES-BRASILIA-BRAZIL (BMC).

REFERENCES

- [1] S.G. Dellepiane, F. Fontana, and G.L. Vernazza, "Nonlinear Image Labeling for Multivalued Segmentation," *IEEE Trans. Image Processing*, vol. 5, pp. 429-446, 1996.
- [2] 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.
- [3] H.A. Moghaddam and J.F. Lerallut, "Volume Visualization of the Heart Using MRI 4D Cardiac Images," *J. Computer Information Technology*, vol. 6, pp. 215-228, 1998.
- [4] B.M. Carvalho, C.J. Gau, G.T. Herman, and T.Y. Kong, "Algorithms for Fuzzy Segmentation," *Pattern Analysis Applications*, vol. 2, pp. 73-81, 1999.
- [5] P.K. Saha, J.K. Udupa, and D. Odhner, "Scale-Based Fuzzy Connected Image Segmentation: Theory, Algorithms, and Validation," *Computer Vision Image Understanding*, vol. 77, pp. 145-174, 2000.
- [6] B.L. Rice and J.K. Udupa, "Clutter-Free Volume Rendering for Magnetic Resonance Angiography Using Fuzzy Connectedness," *Int'l J. Image Systems Technology*, vol. 11, pp. 62-70, 2000.
- [7] A. Rosenfeld, "Fuzzy Digital Topology," *Information and Control*, vol. 40, pp. 76-87, 1979.
- [8] S.C. Johnson, "Hierarchical Clustering Schemes," *Psychometrika*, vol. 32, pp. 241-254, 1967.
- [9] J.K. Udupa and S. Samarasekera, "Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation," *Graphic Models Image Processing*, vol. 58, pp. 246-261, 1996.
- [10] G.T. Herman, *Geometry of Digital Spaces*. Boston: Birkhauser, 1998.
- [11] C.T. Zahn, "Graph-Theoretic Methods for Detecting and Describing Gestalt Clusters," *IEEE Trans. Computer*, vol. 1, pp. 68-86, 1971.
- [12] N. Ahuja, "Dot Pattern Processing Using Voronoi Neighborhoods," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 3, pp. 336-343, 1982.
- [13] A.K. Jain, M.N. Murty, and P.J. Flynn, "Data Clustering: A Review," *ACM Computer Surveys*, vol. 31, pp. 264-323, 1999.
- [14] J.C. Gower and G.J.S. Ross, "Minimum Spanning Trees and Single Linkage Cluster Analysis," *Application Statistics*, vol. 18, pp. 54-64, 1969.
- [15] S.K. Pal and D.K.D. Majumder, *Fuzzy Mathematical Approach to Pattern Recognition*. New Delhi, India: Wiley Eastern Ltd., 1986.
- [16] T.H. Cormen, C.E. Leiserson, and R.L. Rivest, *Introduction to Algorithms*. Cambridge, Mass.: MIT Press, 1990.
- [17] J.K. Udupa, P.K. Saha, and R.A. Lotufo, "Fuzzy Connected Object Definition in Images with Respect to Co-Objects," *Proc. SPIE Image Processing*, K.M. Hanson, ed., vol. 3661, pp. 236-245 1999.
- [18] I. Pollak, A.S. Willsky, and H. Krim, "Image Segmentation and Edge Enhancement with Stabilized Inverse Diffusion Equations," *IEEE Trans. Image Processing*, vol. 9, pp. 256-266, 2000.
- [19] G. Koepfler, C. Lopez, and J.-M. Morel, "A Multiscale Algorithm for Image Segmentation by Variational Method," *SIAM J. Numerical Analysis*, vol. 31, pp. 282-299, 1994.

- [20] T. Hofmann, J. Puzicha, and J.M. Buhmann, "Unsupervised Texture Segmentation in a Deterministic Annealing Framework," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, pp. 803-818, 1998.
- [21] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.
- [22] S.C. Zhu and A. Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, pp. 884-900, 1996.
- [23] L.G. Shapiro and R.M. Haralick, "Decomposition of Two-Dimensional Shape by Graph-Theoretic Clustering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 1, pp. 10-20, 1979.
- [24] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Proc. Computer Vision Pattern Recognition*, pp. 731-737, 1994.



Gabor T. Herman received the PhD degree in 1968. From 1969 to 1981, he was with the Department of Computer Science, State University of New York (SUNY) at Buffalo, where he directed the Medical Image Processing Group. From 1981 to 2000, he was a professor in the Medical Imaging Section of the Department of Radiology at the University of Pennsylvania, during which time he was the editor-in-chief of the *IEEE Transactions on Medical Imaging*. Currently, he is the director of the Center for Computer Science and Applied Mathematics at Temple University. His books include *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography* (Academic, 1980), *3D Imaging in Medicine* (CRC, 1991 and 2000), *Geometry of Digital Spaces* (Birkhauser, 1998), and *Discrete Tomography: Foundations, Algorithms and Applications* (Birkhauser, 1999). He is a fellow of the IEEE.



Bruno M. Carvalho received the BSc degree in computer science from the Federal University of Rio Grande do Norte, Brazil, in 1992, the MSc degree in computer science from the Federal University of Pernambuco in 1995, and the MSc degree in engineering from the University of Pennsylvania in 1999. Currently, he is a PhD student in the Department of Computer and Information Sciences at the University of Pennsylvania. His main interests are medical imaging, computer graphics, and computer vision. He is a student member of the IEEE.