Published in final edited form as:
Comput Vis Image Underst. 2011 June 1; 115(6): 721–734. doi:10.1016/j.cviu.2011.01.003.

A framework for comparing different image segmentation methods and its use in studying equivalences between level set and fuzzy connectedness frameworks

Krzysztof Chris Ciesielski[*,a,b] and **Jayaram K. Udupa**[b]

[a]Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310

[b]Department of Radiology, MIPG, University of Pennsylvania, Blockley Hall – 4th Floor, 423 Guardian Drive, Philadelphia, PA 19104-6021

Abstract

In the current vast image segmentation literature, there seems to be considerable redundancy among algorithms, while there is a serious lack of methods that would allow their theoretical comparison to establish their similarity, equivalence, or distinctness. In this paper, we make an attempt to fill this gap. To accomplish this goal, we argue that: (1) every digital segmentation algorithm $\mathcal{A}$ should have a well defined continuous counterpart $\mathcal{M}_{\mathcal{A}}$, referred to as its model, which constitutes an asymptotic of $\mathcal{A}$ when image resolution goes to infinity; (2) the equality of two such models $\mathcal{M}_{\mathcal{A}}$ and $\mathcal{M}_{\mathcal{A}'}$ establishes a theoretical (asymptotic) equivalence of their digital counterparts $\mathcal{A}$ and $\mathcal{A}'$. Such a comparison is of full theoretical value only when, for each involved algorithm $\mathcal{A}$, its model $\mathcal{M}_{\mathcal{A}}$ is *proved* to be an asymptotic of $\mathcal{A}$. So far, such proofs do not appear anywhere in the literature, even in the case of algorithms introduced as digitizations of continuous models, like level set segmentation algorithms.

The main goal of this article is to explore a line of investigation for formally pairing the digital segmentation algorithms with their asymptotic models, justifying such relations with mathematical proofs, and using the results to compare the segmentation algorithms in this general theoretical framework. As a first step towards this general goal, we *prove* here that the gradient based thresholding model $\mathcal{M}_{\nabla}$ is the asymptotic for the fuzzy connectedness Udupa and Samarasekera segmentation algorithm used with gradient based affinity $\mathcal{A}_{\nabla}$. We also argue that, in a sense, $\mathcal{M}_{\nabla}$ is the asymptotic for the original front propagation level set algorithm of Malladi, Sethian, and Vemuri, thus establishing a theoretical equivalence between these two specific algorithms. Experimental evidence of this last equivalence is also provided.

Keywords

segmentation; delineation; algorithm equivalence; convergence; level sets; fuzzy connectedness; medical images

[*]Corresponding author, partially supported by NSF grant DMS-0623906. KCies@math.wvu.edu; web: http://www.math.wvu.edu/~kcies.

1 Introduction

Most scientific contributions to any practical application-motivated model theory fall under one of two related, but slightly different, archetypes: (1) application-oriented *engineering-standard*, which is focused on the work's practical usability (like numerical stability of an implemented algorithm or a degree of similarity of its output to a real phenomenon it models), but is less preoccupied with a formal theoretical analysis of the contribution; and (2) theoretically-oriented *mathematical-standard*, which is focussed on a logically impeccable theoretical analysis of one or more models, but is less preoccupied with (often even oblivious to) a practical usability and/or value of the models. In the application-motivated theories that model a given phenomenon (including the image segmentation theory, in which we are interested in), usually many engineering-standard contributions are made before a mathematical-standard work starts to appear. This seems to be caused by the initial user-induced demand for different model-based usable products and by the fact that the mathematical-standard work especially thrives at the later stage, when there is enough engineering-standard contributions to analyze. It is our belief, that the image segmentation theory has long reached the stage in its development where mathematical-standard analysis of the existing models should start to appear.

In this work, we present new tools for theoretical analysis and comparison of segmentation algorithms with the hope that they will help to initiate a greater influx of mathematical-standard contributions to image segmentation theory. To stress the value of the approach we have taken, we needed to point out different mathematical-standard weaknesses in the current image segmentation literature. This, however, does not mean that we do not appreciate the great scientific value of this literature. In particular, we have chosen to cite only the papers whose contribution we highly value and which are relevant in one way or another to this paper.

Image segmentation—the process of partitioning the image domain into meaningful object regions—is perhaps the most challenging and critical problem in image processing and analysis. Its central position in image processing comes from the fact that the delineation of objects is usually the first step in other higher level processing tasks, like image interpretation, diagnosis, analysis, visualization, virtual object manipulation, and often even registration. Image segmentation may be thought of as consisting of two related processes: recognition and delineation. *Recognition* is the high-level process of determining roughly the whereabouts of an object of interest in the image. *Delineation* is the low-level process of determining the precise spatial extent and point-by-point composition (material membership percentage) of the object in the image. The topic of this paper concerns image delineation.

General segmentation frameworks are usually broadly classified into three groups: boundary-based [15,22,23,28,32,34,36,37], region-based [6,10,45,49–52], and hybrid [13,26]. However, for the purpose of this paper, the more important classification comes from the mode of algorithm introduction, that is, either as *discretizations* of continuous models (e.g., functional optimization methods, usually implemented via level sets [32,37,45], including active contour [28,30]), or as purely *discrete* algorithms (e.g., graph-cut [8–10], fuzzy connectedness [17,41,52], and watershed[1] [54·38·46]). The first group of algorithms comes always with two theoretical constructs: the continuous model $\mathcal{M}$ and its discretization algorithm $\mathcal{A}$. On the other hand, for any algorithm $\mathcal{A}$ introduced in a purely

[1]Of course, the ideas for watershed were always continuous. However, the continuous version of watershed models were formally introduced after discrete watershed algorithms (see [38]) and there still seems to be no formal connections between the two in the sense defined in this paper. In fact, the approximation of the continuous watershed model with the discrete watershed algorithms is not a straightforward matter, as the results from [9] indicate.

discrete fashion, the existence and format of its continuous counterpart $\mathcal{M}$ is, in general, not clear. Our main tool for algorithm comparison will require a full description of the pairs $\langle \mathcal{M}, \mathcal{A} \rangle$ and a good understanding of the relation between $\mathcal{M}$ and $\mathcal{A}$. Therefore, in what follows, we will argue that all reasonable segmentation algorithms, including those introduced in a purely discrete fashion, should have their continuous counterparts. We will also formalize the intuitive relation between the continuous segmentation models $\mathcal{M}$ and the related algorithms $\mathcal{A}$. It should be stressed that, even for the algorithms introduced as discretizations of continuous models, there are very few published *proofs* connecting in any mathematically meaningful way the models with the algorithms. (Important exceptions are the papers [7] and [1,2,14], but see comments in the footnote #2.) Therefore, even in this class of algorithms, there is currently no mathematically sound theory of segmentation algorithms.[2] Thus, the most fundamental and theoretical goal of this paper is to initiate the study that will connect *in a provable mathematical way* the theory of continuous image segmentation models with the theory of discrete segmentation algorithms.

A practical motivation for the development of a general theoretical study of image segmentation methodologies as alluded to above is to address several gaps that currently exist in our knowledge in this subject, which are denoted (G1)–(G3) in the following: (G1) Are all different families of segmentation methods/models really fundamentally distinct or are there similarities, or even theoretical equivalences, among them? Although there are some rare attempts to compare the methods at a theoretical level (see e.g. [3,20,25,35]), this is largely an uncharted territory. (G2) Segmentation research has two clearly distinct components: the practical, focused on describing efficient segmentation algorithms that can be practically implemented; and theoretical, concerning development and use of sophisticated tools of infinite (i.e., not discrete) mathematics for the purpose of describing segmentation models of idealized images. One of the peculiarities of the current state of segmentation research is that these two tracks are hardly connected in any formal way. True, the papers that start with a description of a segmentation model of idealized images usually transcribe such a model into a digital image segmentation procedure. However, all such translations are done only at an imprecise intuitive[3] level, without a formal, mathematical argument. In fact, there is even no evidence of the use of any definition formally connecting idealized images (infinite objects) with their digital representations (which are finite). (G3) Another element clearly missing from current segmentation research is a set of properties that any digital segmentation algorithm must or should have. For example, it seems desirable to have the output of any reasonable segmentation algorithm to be reasonably stable if it is fed with the digital approximations of the same idealized image with better and better resolution. It would be also desirable for the segmentation output to remain reasonably unchanged when applied to the same resolution digital representations of the same idealized image that was rotated and/or shifted. (This latter aspect becomes important when we keep

[2]Intuitively, we need to prove that the segmentation algorithms constitute "appropriate numerical schemas for their continuous models." The convergence and stability of numerical systems is well studied in numerical analysis, especially when applied to solve PDE (where it is known as finite-element approximation). It seems that these results can be simply applied to the methods like level sets delineation methods, where the object is defined as $S = \{z \in \Omega, \Psi z, t) \leq 0\}$ (Ω being an open, bounded, and connected subset of $\mathbb{R}^n$) for some function $\Psi{:}\Omega \times [0, \infty) \to \mathbb{R}$ constituting a solution of PDE, see e.g. Section 4. The problem is that even if the approximations Ψ_n (finite, or defined on the entire domain) of Ψ converge to Ψ, the approximate objects $S_n = \{z: \Psi_n(z, t) \leq 0\}$ may be very far from S, see e.g. Example 16. The text [4] is a typical example of overlooking this crucial difference — although it contains an entire chapter on the numerical behavior and convergence of solutions for PDEs relevant for image processing, there is no discussion of how such convergence translates (or not) to the actual image processing output. A similar problem (of lack of full interpretation of proved results) is also visible in a very interesting paper [7], which contains a rare example of a formal proof of convergence in the spirit of this article: the authors prove [7, thm. 2] that for a family $\{u^\varepsilon\}_{\varepsilon>0}$ of numerical approximations of image segmentation, where ε is related to image resolution, *there is a sequence* $\langle \varepsilon_j \rangle$ converging to 0, for which the sequence $\langle u^{\varepsilon_j} \rangle$ of segmentations converges to a single (idealized) image segmentation. However, unless *every* such sequence converges, the delineations $\{u^\varepsilon\}_{\varepsilon>0}$ are actually numerically unstable, when $\varepsilon \to 0$.

[3]The general framework of finite-element approximations remain imprecise, without formal convergence results in a format similar to that given in [7].

in mind that, in many areas such as medical imaging, there is no guarantee that the same object with subtle and fine features will be digitized in the same manner in repeated scans/digitizations, although some empirical evaluations of segmentation algorithms have assessed the variability in repeated scans.) So far, there is little research done systematically addressing points (G1)–(G3), especially for the algorithms that were not motivated by idealized image segmentation models. This paper is a first attempt to fill some of these gaps via the general theory proposed in Section 2.

In the Fuzzy Connectedness, FC, framework [52], a fuzzy topological construct, called fuzzy connectedness, characterizes how the spatial elements (abbreviated as *spels*) of an image hang together to form an object. This construct is arrived at roughly as follows. A function called *affinity* is defined on the image domain; the strength of affinity between any two spels depends on how close the spels are spatially and how similar their intensity-based properties are in the image. Affinity is intended to be a local relation. A global fuzzy relation called *fuzzy connectedness* is induced on the image domain by affinity as follows. For any two spels c and d in the image domain, all possible paths connecting c and d are considered. Each path is assigned a strength of fuzzy connectedness which is simply the minimum of the affinities of consecutive spels along the path. The level of fuzzy connectedness between c and d is considered to be the maximum of the strengths of all paths between c and d. For segmentation purposes, FC is utilized in several ways as described below. See [50] for a review of the different FC definitions and how they are employed in segmentation and applications.

In *absolute FC* (abbreviated AFC) [52], the support of a segmented object is considered to be the maximal set of spels, containing one or more seed spels, within which the level of FC is at or above a specific threshold. To obviate the need for a threshold, *relative FC* (or RFC) [41] was developed by letting all objects in the image to compete simultaneously via FC to claim membership of spels in their sets. To avoid treating the core aspects of an object (that are very strongly connected to its seeds) and the peripheral subtle aspects (that may be less strongly connected to the seeds) in the same footing, an iterative refinement strategy is devised in *iterative RFC* (or IRFC) [17,51]. RFC and IRFC can be viewed as graph cut optimizations for appropriate cost functions [20]. The FC family of methods developed to date consists of various combinations of absolute, relative, and iterative FC. In this paper we will study (in Section 4.1) only the AFC algorithm, considered with a gradient based affinity. Note that gradient based affinity is a generalized affinity notion, in a format introduced and examined in [18,19]. The other forms of FC algorithms will be examined within the general framework of Section 2 in our future work.

The level set method refers to the specific model of an evolving front (surface or curve) in a time dependent manner and to the numerical algorithm tracking such propagating fronts. The model and the associated narrow band propagation algorithm were introduced in 1985 by Sethian [44] which made their way into image segmentation research in 1995 with paper [32]. The popularity of the level set method in segmentation tasks led to a multitude of research papers, as exemplified by the books [39,40,45]. Although the level set method in image segmentation is nowadays more often used indirectly to solve the PDE optimizing the segmentation cost functions (see e.g., [16,33,53]), the original segmentation algorithms are still studied [47,48]. Therefore, in this paper, for the purpose of using the theoretical framework for comparing methods, we have chosen the level set method with front propagation, because of its popularity, and FC, because of our familiarity with it. Our choice of the particular algorithms with each of these classes is not for their popularity or power (factually, the level set algorithms are far more popular than the FC algorithms), but rather to make our point that the algorithms introduced via very different mathematical tools can be asymptotically equivalent.

Our general theoretical framework and its variations are described in Sections 2 and 3. The application of the theory to the analysis of a particular model of FC [52] and to a comparison of its algorithms with the level set delineation algorithm of [32] is presented in Section 4. (An attempt at expressing this level set algorithm [32] without PDE can also be found in [47,48].) Although in this paper we focus only on the algorithms of [52] and [32] for a theoretical comparison, the general framework can be utilized to compare any methods in the literature. In Section 5, we present some practical segmentation examples to illustrate the equivalence proved in Section 4 and state our conclusions.

2 A general image segmentation framework

2.1 Stage set up: What is an image?

We will start off by formalizing the notions of an idealized image and its digitization. This formalization is intuitive and rather standard in the imaging literature. However, most of the imaging papers concentrate only on one of these two kinds of images, leaving unanswered or hazy the fundamental question as to what the relation between them is.

Definition 1 An ($\mathfrak{n}$-*dimensional*) *idealized image* (notice gothic $\mathfrak{n}$) is any function F from a bounded connected subset Ω of the $\mathfrak{n}$-dimensional Euclidean space $\mathbb{R}^{\mathfrak{n}}$ into $\mathbb{R}^{\ell}$. In what follows, we will always assume that Ω is an open subset of $\mathbb{R}^{\mathfrak{n}}$, and often it will be just an $\mathfrak{n}$-dimensional cube $\Omega=(a,b)^{\mathfrak{n}}$.

In general, we do not assume any nice properties for function F. However, it will be often necessary to assume that F is piecewise (uniformly) continuous or that it has (uniformly) continuous derivatives.[4] Note, that the intensity value $F(x)$ of F at $x \in \Omega$ is assumed to be a vector in $\mathbb{R}^{\ell}$, although the scalar case of $\ell = 1$ is included. We will always assume that $\mathfrak{n} \geq 2$, although we will allow it to be larger than 3, as a time sequence of three-dimensional images, for example, can be interpreted as a four-dimensional image.

Definition 2 An ($\mathfrak{n}$-*dimensional*) *digital image* is any function f from a finite subset C of $\mathbb{R}^{\mathfrak{n}}$ into $\mathbb{R}^{\ell}$.

In this definition, we slightly depart from the standard assumption that the coordinates of C are the integer numbers, that is, that $C \subset \mathbb{Z}^{\mathfrak{n}}$. Our generality will help us to lay our theory, while it creates no real implementation difficulty, especially in the most important case when C is a subset of a rectangular grid $\{hk : k \in \mathbb{Z}\}^{\mathfrak{n}}$, where $h > 0$ is fixed.

The relation between these two types of images can be expressed as follows.

Definition 3 A digital image $f : C \to \mathbb{R}^{\ell}$ is a *digitization of an idealized image* $F : \Omega \to \mathbb{R}^{\ell}$ provided f is the restriction $F \restriction C$ of F to C, that is, $C \subset \Omega$ and $f(c) = F(c)$ for every $c \in C$.

In what follows, symbol $\|x\|$ will stand for the Euclidean norm of the vector $x = \langle x_1, \ldots, x_{\mathfrak{n}} \rangle \in \mathbb{R}^{\mathfrak{n}}$, that is, $\| x \| = \sqrt{x_1^2 + \cdots + x_{\mathfrak{n}}^2}$. The distance from an $x \in \mathbb{R}^{\mathfrak{n}}$ to a $T \subset \mathbb{R}^{\mathfrak{n}}$ is defined as $\mathrm{dist}(T, x) = \inf\{\|x - t\| : t \in T\}$.

Remark 4 Possibly the digitization $f : C \to \mathbb{R}^{\ell}$ of a true image F should be defined more generally by allowing $f(c)$ to be some appropriate average of F around c, e.g.,

[4] Clearly, for practical applications, only images with "meaningful information" are of concern. However, the notion of "meaningful information" seems to be impossible to express in a formal mathematical form and the distinction is irrelevant to our theory, which will work independent of such a subjective factor.

$f(c) = \frac{\int_{\mathbb{R}^n} F(x) \cdot K(x-c)}{\int_{\mathbb{R}^n} K(c)}$ for some kernel K. One might use as K a Gaussian function $g(x) = \exp(-\|x\|^2/\sigma^2)$ with a constant σ associated with the resolution of the digital image. (The idea here is that K corresponds roughly to the point spread function of the imaging device.) Definition 3, which we will use in this paper, falls under this schema when K is the Dirac delta.

Remark 5 In all practical digital image acquisitions, some image distortion due to various artifacts, such as noise, background variation, must be expected. Because of this fact of life, it is quite common in the imaging literature to assume that the acquired image intensity is in a form $f+\mathfrak{N}$, where f is the "clear" digital image like in Remark 4 and $\mathfrak{N}$ is the noise/artifact component. Although this approach is very desirable from the practical application point of view, in what follows, we will ignore noise in our current considerations. The rationale for this is that the main goal of this paper is to analyze the segmentation algorithms from the theoretical point of view and the most fundamental situation in which they should work is the noiseless environment. In other words, including the image noise analysis in what follows would only obscure the clarity of the presented material. Nevertheless, we indicate, in Section 5, a possibility to extend the presented framework to encompass images with the noise component. The detailed analysis of such theory will be considered in our future extension of this work.

2.2 The segmentation algorithms

The following definition treats a delineation algorithm as a "black box:" Given an input (a digital image map and the parameters of the algorithm, which may include also some prior knowledge on the object to be segmented), the only outcome that is considered is the actual output of the algorithm, which is a segmented image. This definition is general enough to encompass essentially all possible delineation algorithms. As such, it is pertinent, but not restricted, to the two algorithms to be discussed in the next section: fuzzy connectedness, FC, of Udupa and Samarasekera [52] (see Section 4.1) as well as level set Malladi-Sethian-Vemuri [32] (see Section 4.2). We do not restrict the algorithms, as we define them, to any particular application domain, as it is not important for our considerations. Nevertheless, for practical purposes, many algorithms are often considered only for some specific applications. Thus, the same (general) algorithm may be "good" in one application domain, while it may give unacceptable results in another domain. This discrepant behavior, however, will have no effect on the theory presented below.

Definition 6 A *(digital) delineation algorithm* $\mathcal{A}$ is any effectively defined mapping $\langle f, \vec{p} \rangle \overset{\mathcal{A}}{\mapsto} S$ which to any digital image $f: C \to \mathbb{R}^\ell$ (possibly restricted to some subclass) and a parameter vector $\vec{p}$ associates a subset S of C interpreted as a segment of the image f indicated by the parameters. We will write $\mathcal{A}(f, \vec{p})$ for the output S of $\mathcal{A}$ applied to $\langle f, \vec{p} \rangle$.

The parameters may include a threshold number $\theta \in \mathbb{R}$ and some subsets of $\mathbb{R}^n$ (like a simple closed curve, as in the case of some level set algorithms) approximating respective subsets of the domain C of f which carry information on the objects we seek. Often, a seed point $s \in C$ is used as a parameter which indicates the segment S, that is, with the goal that $s \in S$. Some algorithms use also another seed point $t \in C$ indicating the background, that is, with the goal that $t \notin S$. In the applications that follow, the algorithms will have one, η, or two, $\langle \eta, \varepsilon \rangle$, real parameters (η will be a threshold, ε — the curvature weight) that will need to be

explicitly indicated, leading to the parameters $\langle \eta, \vec{p} \rangle$ and $\langle \eta, \varepsilon, \vec{p} \rangle$ in place of simple $\vec{p}$, and algorithm representations $\mathcal{A}(f, \eta, \vec{p})$ and $\mathcal{A}(f, \eta, \varepsilon, \vec{p})$ for $\mathcal{A}(f, \vec{p})$.

We will often refer to a delineation algorithm as a *segmentation algorithm*. The segmentation algorithm, in general, can return as an output a finite sequence $\langle S_1, \ldots, S_k \rangle$ of (usually pairwise disjoint) subsets of C, while a delineation algorithm returns only one set $S \subset C$. The theory presented below is considerably easier to express for the delineation algorithms, while with little effort it can be applied also to any general segmentation procedure, since any segmentation algorithm $\mathcal{A}$ that returns a k-element sequence $\mathcal{A}(f, \vec{p}) = \langle \mathcal{A}_1(f, \vec{p}), \ldots, \mathcal{A}_k(f, \vec{p}) \rangle$ can be treated as k separate (but dependent) segmentation algorithms $\mathcal{A}_1, \ldots, \mathcal{A}_k$.

Next, we will formalize what we believe to be the most fundamental property that any reasonable delineation algorithm should possess:

> (*) The better the resolution of the digital approximation of the idealized image, the closer the algorithm's output is to the "real object" in the idealized image.

Intuitively, this property expresses the natural requirement that: (1) the algorithm applied to the "digital image" with the infinitesimal size of the spels returns the "real object," and (2) the output of the algorithm should approximate, in a limit, the object from (1). Notice also that (*) expresses the intuition connecting continuous models with their digitizations and that rejection of (*) would imply rejection of the logical connection between abstract continuous mathematical models and their digital counterparts.[5]

If $F: \Omega \to F: \Omega \to \mathbb{R}^\ell$ is the idealized image we approximate, the property (*) can be expressed in a mathematical language as:

> The limit $L = \lim_{C \to \Omega} L = \lim_{C \to \Omega} \mathcal{A}(F \restriction C, \vec{p})$ over finite sets $C \subset \Omega$ exists and represents the "real object."[6]

This template definition requires the explanation of several terms as described below for completion.

Mode of convergence—First, we need to decide the meaning of convergence of sets C to Ω as well as sets $A = \mathcal{A}(F \restriction C, \vec{p}) \subset C$ to $L \subset \Omega$. We will treat these as convergences in Hausdorff distance ρ, although in our actual applications described in the next sections, we will use a slightly more general notion of convergence. In what follows, the Hausdorff distance will be used only in a simple situation of bounded subsets of $\mathbb{R}^n$ comparable by inclusion, say $B \subseteq U$, in which case, ρ can be expressed as

$$\rho(B, U) = \sup \{\operatorname{dist}(x, B) : x \in U\}.$$

Since any ρ-convergence can be expressed in term of sequences, in what follows, the above limit requirement for $\mathcal{A}$ will be replaced by the equivalent form:

---

[5] In fact, if an algorithm is a digitization of a mathematically-precise continuous model but there is no *mathematical* proof that the model is an asymptotic of the algorithm, then there is no *complete mathematical* theory for the *algorithm*. In such situations, which are currently predominant in the image segmentation literature, the mathematical theory of a continuous model should be treated only as an *inspiration* for the algorithm, but not as a complete mathematical theory of the algorithm.

[6] More generally, we can consider also limits $L = \lim_{f \to F} \mathcal{A}(f, \vec{p})$ over all functions $f: C \to \mathbb{R}^\ell$, $C \subset \Omega$ being finite, for an appropriately defined notion of convergence (most likely, via Hausdorff distance) of f to F.

There is a set $L \subset \Omega$, depending on $\mathcal{A}$ F, and $\overrightarrow{p}$, with the property that $L=\lim_{n\to\infty}\mathcal{A}\left(F \restriction C_n, \overrightarrow{p}\right)$ for any (appropriate) sequence of finite sets C_n converging to Ω.

Note that $\rho(B, U) = 0$ implies only that B is a dense subset of U (i.e., the topological closure $\mathrm{cl}(B)$ of B contains U), but not the equality $B = U$, unless both sets are closed. (Hausdorff distance is a metric only when restricted to the family of compact sets.) Since we will be mainly interested in the situation when the result of the limit is dense in some open set (so, not closed), it will be convenient to have the following definition of the limit, which does not use the notion of the Hausdorff distance.

For a subset A of an underlying space X (we can take $X = \Omega$), a *characteristic* (or *indicator*) function χ_A of A is defined as $\chi_A(x) = 1$ for $x \in A$ and $\chi_A(x) = 0$ for $x \in X / A$. Recall (see e.g. [29, p. 173]) that for a sequence $\langle A_i \rangle_{i=1}^{\infty}$ of subsets of X we define $\limsup_i A_i=\bigcap_{j=1}^{\infty}\bigcup_{i\geq j}A_i$. Note that $A = \limsup_i A_i$ holds precisely when $\chi_A = \limsup_i \chi_{A_i}$. Similarly, we define $\liminf_i A_i=\bigcup_{j=1}^{\infty}\bigcap_{i\geq j}A_i$ and we have $A = \liminf_i A_i$ precisely when $\chi_A = \liminf_i \chi_A$. The limit $L = \lim_i A_i$ exists and is equal to $\limsup_i A_i$ provided $\limsup_i A_i$.

Notice that, for L so defined, sets A_i converge to L in Hausdorff distance,[7] but $\langle A_i \rangle_i$ converges in Hausdorff distance also to any $U \supset L$ contained in the closure of L.

Allowable images and approximating grids—Second, for every algorithm $\mathcal{A}$ it will be necessary to decide for what family $\mathcal{F}$ of idealized images F and for what parameters $\overrightarrow{p}$ the algorithm should have the above limit property. This needs to be answered for each algorithm separately, with an answer provided in the statement of each convergence theorem. Similarly, we will usually restrict our attention to the limits over the sequences $\langle C_n \rangle_n$ from some fixed family $\mathcal{C}$ of sequences. In what follows we will use two kinds of such families. The larger of these will be the family $\mathcal{C}_0$ of all $\subset$-increasing sequences $\langle C_n \rangle_n$ (i.e., with $C_1 \subset C_2 \subset \cdots$) of finite subsets of Ω for which the sequence of Ω-resolution numbers

$$r_\Omega(C_n) \stackrel{\text{def}}{=} \rho(C_n, \Omega) \tag{1}$$

converges to 0. Clearly, requirement $\lim_{n\to\infty} r_\Omega(C_n) = 0$ is equivalent to that for $\langle C_n \rangle_n$ converging to Ω in the Hausdorff distance. We will be also interested in a family $\mathcal{C}_R \subset \mathcal{C}_0$ formed from uniformly spaced rectangular grids in Ω. More precisely, for $h > 0$, let $(h\mathbb{Z})^n=\{hk:k \in \mathbb{Z}\}^n$ be the rectangular grid of points in $\mathbb{R}^n$ with the basic distance h and let $\Omega_h=\Omega \cap (h\mathbb{Z})^n$. Then $\mathcal{C}_R$ is the family of all sequences $\left\langle \Omega_{h/2^i} \right\rangle_{i=1}^{\infty}$ for some $h > 0$. Thus, we are doubling the resolution[8] when passing from $C_i = \Omega_{h/2^i}$ to C_{i+1}. Since Ω is an open set, clearly $\rho(\Omega_{h/2^i}, \Omega) \to_i 0$.

"Real object" and its approximation—Finally, we should make clear that we see no mathematical way of treating the "real object" as may be expected in the intuitive sense, that is, there is no mathematical way to directly define a "real object." Instead, we will identify this notion with the output $\mathcal{M}\left(F, \overrightarrow{p}\right)$ of the (idealized) models, as described below. Moreover, since the set $\mathcal{M}\left(F, \overrightarrow{p}\right)$ will be open, there is no way that $L=\lim_{n\to\infty}\mathcal{A}\left(F \restriction C_n, \overrightarrow{p}\right)$

[7] Meaning that $\pi(A_i, L) \to_i 0$.

[8] *Resolution* in the intuitive sense, which for some sets Ω may slightly differ from the notion of Ω-resolution defined above.

being a subset of a countable set $\bigcup_n C_n$, is actually equal to $\mathcal{M}(F,\vec{p})$. However, if L is a dense subset of $\mathcal{M}(F,\vec{p})$, then the sequence $\langle \mathcal{A}(F \restriction C_n,\vec{p})\rangle_n$ ρ-converges to $\mathcal{M}(F,\vec{p})$ Thus, in what follows, the expression *sequence* $\langle A_n\rangle_n$ *converges to* $U=\mathcal{M}(F,\vec{p})$ will be interpreted as "$L = \lim_n A_n$ is a dense subset of U" and will be often expressed (slightly incorrectly) as $U = \lim_n A_n$.

2.3 Segmentation models and related algorithms

In this section, we will define the notion of a segmentation model for an idealized image as well as a fundamental relation between the segmentation models and the related algorithms.

Definition 7 A *delineation model* $\mathcal{M}$ for a class $\mathcal{F}$ of idealized images is any mapping $\langle F,\vec{p}\rangle \overset{\mathcal{M}}{\mapsto} O$ which for any image $F:\Omega \to \mathbb{R}^\ell$ from $\mathcal{F}$ and any parameters $\vec{p}$ associates a subset O of Ω interpreted as a segment of the image F indicated by the parameters. We will write $\mathcal{M}(F,\vec{p})$ for the output O of $\mathcal{M}$ applied to $\langle F,\vec{p}\rangle$. A *segmentation model* $\mathcal{M}$ that returns k is objects defined as a k-element sequence $\mathcal{M}(f,\vec{p}) = \langle \mathcal{M}_1(f,\vec{p}),\ldots,\mathcal{M}_k(f,\vec{p})\rangle$ of delineation models $\mathcal{M}_i$.

As before, we will sometimes list a part of the general parameter vector $\vec{p}$ explicitly, that is, replacing $\vec{p}$ with either $\langle \eta,\vec{p}\,\rangle$ or $\langle \eta,\varepsilon,\vec{p}\rangle$, leading to the notation $\mathcal{M}(F,\eta,\vec{p})$ and $\mathcal{M}(F,\eta,\varepsilon,\vec{p})$ in place of $\mathcal{M}(F,\vec{p})$.

Notice that in our definition of a delineation model $\mathcal{M}$ we do not require it to be effective (in a sense, that $\mathcal{M}$ does not need to come with an *explicit* procedure for finding $\mathcal{M}(F,\vec{p})$), despite the fact that this might be considered as a departure from a terminology used in many mathematical modeling papers. To justify our choice of this terminology, we note that modeling papers (including segmentation modeling) frequently start with a non-effectively defined map $\mathcal{M}$, often as an optimizer (minimizer or maximizer) of some functional, and then proceed to finding an effectively defined procedure $\widehat{\mathcal{M}}$, often via a solution of a differential equation and/or using variational methods, which, as a function, is equal to $\mathcal{M}$. We will refer to such $\widehat{\mathcal{M}}$ as a *solution to a model* $\mathcal{M}$. (In the literature, often the procedure $\widehat{\mathcal{M}}$ itself is designated as the "model.") Since, treated as functions, $\mathcal{M}=\widehat{\mathcal{M}}$, our definition can be applied to $\widehat{\mathcal{M}}$ as well as to $\mathcal{M}$. If one plans to use the effective version $\widehat{\mathcal{M}}$ of the model to find an algorithm $\mathcal{A}$ that approximates $\mathcal{M}=\widehat{\mathcal{M}}$, then $\widehat{\mathcal{M}}$ is at the center of the investigation and it makes sense to designate $\widehat{\mathcal{M}}$ as "the model" of the process. We think of this modeling schema as a two stage process: $\mathcal{M} \to \widehat{\mathcal{M}} \to \mathcal{A}$. In a large class of delineation methods (including the model discussed in Section 4.2, as well as many optimization models), the value of $\widehat{\mathcal{M}}$ is found via time dependent front propagation, usually approximated numerically with fast marching level set algorithms. (See e.g. [39,40].) In many modeling tasks other than segmentation, like modeling of the wave propagation or flame burning, the time sequence of the consecutive approximations is at least as important as the final position of the front, making the front propagation approach the most desirable. However, in segmentation tasks, we are usually not interested in the intermediate stages of object approximation, and we treat the final position of the front as the only output of the model and the algorithm. Thus, in the investigation of the essential aspects of a delineation task, it is more productive to follow directly from $\mathcal{M}$ to $\mathcal{A}$, that is, forgoing the effective

version $\widehat{\mathcal{M}}$ of the model and follow the schema $\mathcal{M} \to \mathcal{A}$. This is the central idea behind our investigation presented in Section 4.

The following definition is the most fundamental in this paper. It formally relates the segmentation model of an idealized image with the associated segmentation algorithm.

Definition 8 If $\mathcal{M}$ is a delineation model for a class $\mathcal{F}$ of idealized images from Ω into $\mathbb{R}^\ell$ and C is a family of sequences $\langle C_n \rangle_{n=1}^{\infty}$ of finite subsets of Ω that converge to Ω, then we say that a delineation algorithm $\mathcal{A}$ *represents model* $\mathcal{M}$ (in terms of sequences from C and for the class $\mathcal{F}$ of images) provided for every $\langle C_n \rangle_n \in C$, $F \in \mathcal{F}$, and parameter $\vec{p}$ appropriate for F, sequence $\mathcal{A}\left(F \restriction C_n, \vec{p}\right)$ converges to $\mathcal{M}\left(F, \vec{p}\right)$.

Similarly, a *segmentation algorithm* $\mathcal{A} = \langle \mathcal{A}_1, \ldots, \mathcal{A}_k \rangle$ *represents a segmentation model* $\mathcal{M} = \langle \mathcal{M}_1, \ldots, \mathcal{M}_k \rangle$ provided each $\mathcal{A}_i$ appropriately $\mathcal{M}_i$.

Two delineation, or segmentation, algorithms $\mathcal{A}$ and $\mathcal{A}'$ are *asymptotically equivalent* (*in terms of sequences from* C *and for the class* $\mathcal{F}$ *of images*) provided they represent the same model $\mathcal{M}$.

The property "$\mathcal{A}\left(F \restriction C_n, \vec{p}\right)$ converges to $\mathcal{M}\left(F, \vec{p}\right)$" will be sometimes expressed as in form $\lim_{n\to\infty} \mathcal{A}\left(F \restriction C_n, \vec{p}\right) = \mathcal{M}\left(F, \vec{p}\right)$, which we will understand as "the set $\lim_{n\to\infty} \mathcal{A}\left(F \restriction C_n, \vec{p}\right)$ is dense in $\mathcal{M}\left(F, \vec{p}\right)$."

It should be stressed here that asymptotically equivalent algorithms behave identically only in the limit at the infinitely best resolution. So, their outputs may still be slightly different for given digital images. This may be contrasted with strongly-equivalent algorithms (defined and studied in [18,19]), which have identical outputs. Note also that equivalent algorithms (in the sense of any of these definitions) may still have very different computational times and/or memory requirements. Thus, equivalent algorithms should still be compared at some more subtle level: by analyzing their computational requirements, by estimating computational errors, and by running comparative simulations. Nevertheless, the equivalence of two segmentation algorithms is a strong *theoretical* evidence that they perform quite similarly.

3 Some application-driven variations of the general definitions

The convergence of $\left\langle \mathcal{A}\left(F \restriction C_n, \vec{p}\right) \right\rangle_n$ to $\mathcal{M}\left(F, \vec{p}\right)$ is very closely related to the continuity of function $\mathcal{M}\left(\cdot, \vec{p}\right)$. In fact, only under small additional assumptions,[9] this convergence implies the continuity of function . Unfortunately, even in some most basic situations the models tend to be discontinuous—there always tend to be a very nice F and parameters $\vec{p}$ such that arbitrary small changes from F to similarly nice F_0 cause a major leak from $\mathcal{M}\left(F, \vec{p}\right)$ to a considerably bigger set $\mathcal{M}\left(F_0, \vec{p}\right)$. (See also comment pre-ceding Example 16.) This seems to be a serious obstacle to the schema described in the previous section: we

[9]E.g. the similar convergence for every F_0 in some neighborhood of F, which implies the local uniform convergence of $\mathcal{M}\left(\cdot, \vec{p}\right)$ near F.

cannot expect that the property (*) holds in a vicinity of such F and it is very difficult to describe the pairs $\langle F, \vec{p} \rangle$ where the difficulty occurs.

In what follows we have chosen to go around this problem. Instead of using the limits described in the previous section, we will use their generalizations, which constitute an analog of one sided limits for the real functions. The benefit of this approach is that, for the class of functions we will consider, these generalized limits will always exist and converge to the appropriate set $\mathcal{M}(F, \vec{p})$, while in the case when $\mathcal{M}(\cdot, \vec{p})$ is continuous at $\mathcal{M}\langle F, \vec{p} \rangle$, these generalized limits actually coincide with the limits from the previous section. Thus, in the remainder of this section we will define the necessary generalized limits, similar in spirit to the notion of Γ-convergence [11], and prove some useful facts about them.

We will use two generalized limit notions, lim* and lim† both of which fall into the general schema of limits converging with respect to a filter in the index space. (See e.g. [43, p. 169]. For example, for lim* defined below, the convergence is with respect to a filter defined on the set $(0, \infty) \times \mathbb{N}$, where $\mathbb{N} = \{1, 2, 3, \dots, \}$ is the set of natural numbers, consisting of all subsets of $(0, \infty) \times \mathbb{N}$ containing a set of the form $\bigcup_{s=s_0}^{\infty} \left[2^{-s-1}, 2^{-s}\right] \times \{j_s, j_s+1, \dots, \}$, where numbers s_0, $j_s \in \mathbb{N}$ are arbitrar evaluation of both of these limits will require calculation of several limits in a hierarchical manner. The main limit notion that we will use for this purpose is defined for the families of sets $\{A_i(\eta) : \eta \in \mathbb{R}, \& \, i = 1, 2, 3, \dots\}$ by a formula

$$\lim_{i,\theta} A_i(\theta) \stackrel{\text{def}}{=} \lim_{\eta \to 0+} \left(\lim_{i \to \infty} \sup A_i(\theta - \eta) \right) = \lim_{\eta \to 0+} \left(\bigcap_{j=1}^{\infty} \bigcup_{i \le j} A_i(\theta - \eta) \right), \tag{2}$$

where we define $B = \lim_{\eta \to 0+} B(\eta)$ if and only if $\chi_B = \lim_{\eta \to 0+} \chi_{B(\eta)}$. However, in Subsection 4.2, the family of sets A_i will have two real parameters, that is, we will deal with sets $\{A_i(\eta, \varepsilon) : \eta, \varepsilon \in \mathbb{R} \, \& \, i = 1, 2, 3, \dots\}$. This will require a modification of lim* to lim† defined as:

$$\lim_{i,\eta,\varepsilon}{}^{\dagger} A_i(\eta, \varepsilon) = \lim_{\varepsilon \to 0+} \left(\lim_{i,\eta} * \, A_i(\eta, \varepsilon) \right).$$

Notice that if $A_i(\eta)$ does not depend on ε, then the two limit notions coincide, that is,

$$\lim_{i,\eta,\varepsilon}{}^{\dagger} A_i(\eta) = \lim_{i,\eta} * \, A_i(\eta). \tag{3}$$

The limit notion from (2) will be applied to the segmentation algorithms as follows. We will assume that the algorithm uses one scalar parameter $\theta \in \mathbb{R}$ and some other parameters $\vec{p}$. Thus, the output of the algorithm can be expressed as $\mathcal{A}(f, \theta, \vec{p})$. Then, for a fixed parameter $\vec{p}$ and fixed sequence $\langle F \restriction C_i \rangle_i$ of the digital approximations of an idealized image $F : \Omega \to \mathbb{R}^\ell$, we will define $A_i(\eta) = \mathcal{A}(F \restriction C_i, \theta - \eta, \vec{p})$ and, to prove that $\mathcal{A}$ represents $\mathcal{M}$, we will require that the limit $\lim^*_{i,\eta} A_i(\eta)$ exists and is dense in $\mathcal{M}(F, \theta, \vec{p})$. For the algorithms which we consider in this paper, these sets $A_i(\eta)$ will satisfy the assumptions of the following fact.

Proposition 9 *Let* $\{A_i(\eta):\eta\in\mathbb{R}\ \&\ i=1,2,3,\ldots\}$ be a family of sets such that $A_i(\eta')\subseteq A_i(\eta)$ for every $i\in\{1,2,3,\ldots\}$ and $\eta'>\eta$. Then $\lim^{*}_{i,\eta}A_i(\eta)$ exists and equals $\bigcup_{\eta>0}\left(\bigcap_{j=1}^{\infty}\bigcup_{i\geq j}A_i(\eta)\right)$.

Proof Let $B(\eta)=\bigcap_{j=1}^{\infty}\bigcup_{i\geq j}A_i(\eta)$. Then, $B(\eta')\subseteq B(\eta)$ for every $\eta'>\eta$. So, the limit $\lim_{\eta\to 0+}B(\eta)$ exists and is equal to $\bigcup_{\eta>0}B(\eta)$.

The limit notion $\lim^{\dagger}$ will be used when the algorithm $\mathcal{A}^{\varepsilon}\left(f,\theta,\vec{p}\right)$ depends, in addition, on a parameter $\varepsilon>0$, in which case we will use it for $A_i(\eta,\varepsilon)=\mathcal{A}^{\varepsilon}\left(F\restriction C_i,\theta-\eta,\vec{p}\right)$. To prove that $\mathcal{A}=\lim_{\varepsilon\to 0+}\mathcal{A}^{\varepsilon}$ represents $\mathcal{M}$, we will require that the limit $\lim^{\dagger}_{i,\eta,\varepsilon}A_i(\eta,\varepsilon)=\lim^{\dagger}_{i,\eta,\varepsilon}\mathcal{A}^{\varepsilon}\left(F\restriction C_i,\theta-\eta,\vec{p}\right)$ exists and is dense in $\mathcal{M}\left(F,\theta,\vec{p}\right)$, see page 13.

4 Gradient Based Edge-Threshold Delineation Model $\mathcal{M}_{\nabla}$

In this section, we will analyze a gradient based thresholding model $\mathcal{M}_{\nabla}$ and give a detailed proof that it is represented by the absolute fuzzy connectedness algorithm of Udupa and Samarasekera [52] used with a gradient (homogeneity) based affinity. We will also present an argument that the front propagation level set algorithm of Malladi, Sethian, and Vemuri from [32] represents $\mathcal{M}_{\nabla}$ as well, thus establishing an asymptotic equivalence between these two algorithms.

In the model $\mathcal{M}_{\nabla}$, the edge (i.e., boundary) of the object $P\subset\Omega$ of interest is identified as the set of points x at which the image intensity, given by F, changes rapidly. Mathematically, this means that at the edge points the gradient magnitude $|\nabla F(x)|$ of F is large. Of course, this has a meaning only when the function $|\nabla F(x)|$ is well defined, that is, when F is differentiable. (A possible meaning of $|\nabla F|$ for non-differentiable F, and its implication to the presented discussion, is outlined in Section 5.) Thus, for this model, we will assume that F is of the class C^1, that is, that F has continuous first order partial derivatives. Also, "large gradient" will be interpreted here "as greater than or equal to some threshold number θ." Thus, the object of interest will be a connected component of the set $\Omega(\theta)=\{x\in\Omega:|\nabla F(x)|<\theta\}$. The component will be indicated by some connected set $S\subset\Omega(\theta)$ of seeds, usually a single point or a simple closed curve. This component is the result of applying the model $\mathcal{M}_{\nabla}$ and parameters $\langle\theta,S\rangle$ to F, that is, it is equal to $\mathcal{M}_{\nabla}(F,\theta,S)$. Note that the continuity of the gradient implies that this set is open. We will usually denote $\mathcal{M}_{\nabla}(F,\theta,S)$ as $P^{F}_{S\theta}$ or just $P_{S\theta}$ when F is clear from the context.

Next, we describe a characterization of $P_{S\theta}$ that can be naturally translated into a numerical algorithm of its approximation. For this, we need the following definitions. A *path* p in Ω is any continuous injection from an interval $[a,b]$ into Ω. We say that a path p is from $S\subseteq\Omega$ to $x\in\Omega$ provided $p(a)\in S$ and $p(b)=x$. In this model, a strength[10] $\mu(p)$ of a path p: $[a,b]\to\Omega$, which depends on F, is defined as

[10]More formally we should use here the term $\preceq$-strength, where $\preceq$ denotes the reverse standard inequality (i.e., $\leq$), since the path strength usually denotes its weakest link. However, we will not use the prefix $\preceq$- in this context, despite the fact that this leads to some language awkwardness. (Compare also (5) and the following comment.) The theory of affinity functions defined in terms of abstract linear order relations is described in [18,19].

$$\mu(p) = \sup_{t\in[a,b]} |\nabla F(p(t))|.$$

Notice that the compactness of $[a, b]$ and the continuity of $\nabla F(x)$ implies that $\mu(p) = |\nabla F(p(t_0))|$ for some $t_0 \in [a, b]$.

Theorem 10 *For every C^1 image $F:\Omega \to \mathbb{R}^\ell$, $\theta \in \mathbb{R}$ and a connected set $S \subset \Omega(\theta)$, the object $P_{S\theta}$ is equal to the set of all $x \in \Omega$ for which there exists a path p from S to x with $\mu(p) < \theta$.*

Proof Let $x \in P_{S\theta}$. To find an appropriate path, notice that $P_{S\theta}$ is path connected, since it is a connected open subset of $\mathbb{R}^n$. Thus, for every $s \in S$ there exists a path p: $[0, 1] \to P_{S\theta}$ from s to x. So, for every $t_0 \in [0, 1]$ we have $|\nabla F(p(t_0))| < \theta$, since $p(t_0) \in P_{S\theta} \subseteq \Omega(\theta)$. In particular, $\mu(p) < \theta$.

Conversely, let p: $[a, b] \to \Omega$ be a path from S to x with $\mu(p) < \theta$. We need to show that $x \in P_{S\theta}$. Indeed, the range range(p) of p (defined as range(p) = $\{p(t): t \in [a, b]\}$) is connected (as a continuous image of a connected set) and it intersects S. So, $S \cup \text{range}(p) \subset \Omega(\theta)$ is connected, and it must be a subset of $P_{S\theta}$, since $P_{S\theta}$ is the largest connected subset of $\Omega(\theta)$ containing S. So, $x \in \text{range}(p) \subset P_{S\theta}$.

The next, robustness, theorem tells us that the form of the object $P_{S\theta}$ essentially does not depend on the choice of the seed set S. This can be viewed as an *invariance property* of the model $\mathcal{M}_\nabla$, that is, a property that the model's outcome remains unchanged under some changes of the input. Interestingly, this is also a general property of the fuzzy connectedness segmentation algorithms, as shown in [41] (compare also [17, cor 2.7]), and its analog holds true also for the algorithm $\mathcal{A}_\nabla$ described below.

Theorem 11 *For every C^1 image $F:\Omega \to \mathbb{R}^\ell$, $\theta \in \mathbb{R}$, and nonempty connected sets $S, T \subset \Omega(\theta)$,*

$$P_{S\theta} = P_{T\theta} \neq \emptyset \text{ if and only if there is a path } p \text{ from } S \text{ to } T \text{ with } \mu(p) < \theta.$$

In particular, if $T \subset P_{S\theta}$, then $P_{S\theta} = P_{T\theta}$.

Proof "⇒" Let $x \in P_{S\theta} = P_{T\theta}$. Then there are paths $p_1 : [a, b] \to \Omega$ from S to x and $p_2 : [b, c] \to \Omega$ from x to T with $\mu(p_i) < \theta$ for $i = 1, 2$. Then $p = p_1 \cup p_2$ is as desired.

"⇐" The set $\emptyset \neq S \cup \text{range}(p) \cup T\Omega(\theta)$ is connected, and $P_{S\theta}$ and $P_{T\theta}$ intersect it. Therefore, since each of these sets is a component of $\Omega(\theta)$, we must have $P_{S\theta} = P_{T\theta} \supset S \cup \text{range}(p) \cup T$.

Another invariance property of $\mathcal{M}_\nabla$, stated precisely in the next theorem, says that the output of $\mathcal{M}_\nabla$ does not depend on the origin and the orientation of the coordinate system imposed on Ω. This is a desirable property, since, in most cases, this is what we would expect from the model, while in the practical processes of acquisition of digital approximations of the real images/objects, we usually cannot insure precise and consistent alignment of the objects with the acquisition equipment.

Theorem 12 *Model $\mathcal{M}_\nabla$ is invariant on the distance preserving transformations of the image domain in the sense that, for every distance preserving transformation (isometry) i of $\mathbb{R}^n$, differentiable image $F:\Omega \to \mathbb{R}^\ell$, $\theta \in \mathbb{R}$, and a seed set $S \subset \Omega$,*

$$\mathcal{M}_\nabla(F \circ i, \theta, i[S]) = i[\mathcal{M}_\nabla(F, \theta, S)],$$

where F o*i is a composition of i restricted to* $i^{-1}(\Omega)$ *and F. In particular*, $\mathcal{M}_\nabla$ is invariant to translation and rotation of the image.

Proof This follows easily from the fact that the gradient magnitude remains unchanged under isometrical transformation of the function domain: $|\nabla F|(i(x)) = |\nabla(F \circ i)|(x)$.

4.1 First algorithm representing model $\mathcal{M}_\nabla$: gradient based Udupa-Samarasekera AFC algorithm $\mathcal{A}_\nabla$

We start here by describing a general form of the absolute fuzzy connectedness, AFC, algorithm of [52]. It is used to delineate the images identified with the intensity functions f from the finite subset C of $\mathbb{R}^n$ into $\mathbb{R}^\ell$. The elements of C are referred to as *spels*.

We will think of f as a restriction of some idealized image $F:\Omega \to \mathbb{R}^\ell$ to a subset C of Ω. In most practical applications, C is a subset of a rectangular grid $(h\mathbb{Z})^n = \{hk : k \in \mathbb{Z}\}^n$ which, in terms of the idealized image, can be defined as $C = \Omega_h = \Omega \cap (h\mathbb{Z})^n$. (In fact, for algorithmic implementation, it is usually assumed that $h = 1$, that is, that $C \subset \mathbb{Z}^n$. This does not change the essence of the algorithm, since $(h\mathbb{Z})^n$ and $\mathbb{Z}^n$ can be naturally identified. Nevertheless, to describe the relation of the algorithm with the model, we need to adhere to the assumption that $C \subset \Omega$.) The special case $C = \Omega_h$ is also easier to handle in the analysis that follows, so we will give it special attention. We should also stress that, in the algorithm that follows, we will never use the fact that f is a restriction of an F. This fact will be used only to help our intuition and to express the convergence theorem. However, we will use Ω as a parameter of the algorithm, unless $C = \Omega_h$, in which case this parameter will be dropped.

Adjacency relation—The domain C of the digital image f, for which $\Omega \supset C$ is fixed, comes with an *adjacency relation* telling us which pairs $c, d \in C$ of spels are adjacent, that is, close enough to be considered spatially connected to each other. In this paper, we will assume that the adjacency relation is expressed in terms of the Euclidean distance in $\mathbb{R}^n$ as follows: for some constant $\alpha > 0$, the spels c and d are said to be adjacent provided $\|c - d\| \leq \alpha$. We will assume that for every $h \in (0, 1]$

$$\alpha \in \begin{cases} \left[h, n^2 h\right] & \text{if } C = \Omega_h \text{ and } h \leq r_\Omega(C) \\ \left(2r_\Omega(C), n^2 r_\Omega(C)\right] & \text{oherwise,} \end{cases} \tag{4}$$

where $r_\Omega(C)$ is defined as in (1). Thus, in general, the choice of α depends on the domain C of f and the set Ω. However, if $C = \Omega_h$ and h is small enough (for example, if Ω contains a ball of radius nh), then $h \leq r_\Omega(C)$ and $\alpha \in [h, n^2 h]$. In particular, α does not depend on Ω in this fundamental case. This is important, since the algorithm we will construct depends on f and α. Thus, in the case when $C = \Omega_h$ and h is reasonably small (in terms of size of Ω), we can assume that the algorithm depends only on f. In the general case, however, Ω will be also a parameter of the algorithm.

If $C = \Omega_h$ and $\alpha = h$, then we deal with 4-adjacency for $n = 2$, and with 6-adjacency for $n = 3$. If $C = \Omega_h$ and $\alpha = \sqrt{3h}$, then we deal with 8-adjacency for $n = 2$, and with 26-adjacency for $n = 3$.[11] The idea behind the adjacency relation is to capture the blurring effect of the "point spread function" of imaging devices; that is, that the neighborhood size α should relate to the width of the point spread function.

A path in a digital scene—The choice of α as in (4) ensures the following important property, where $B[T,\varepsilon]=\{x\in\mathbb{R}^n:\mathrm{dist}(T,x)\le\varepsilon\}$ is a generalized closed ball in $\mathbb{R}^n$ centered at $T\subset\mathbb{R}^n$ and with radius ε > 0. Recall that a *path p in C* is any sequence $\langle c_1,\ldots,c_k\rangle$ of spels in *C*, where consecutive c_i and c_{i+1} are adjacent; *p* is from $c\in C$ to $d\in C$ if $c_1=c$ and $c_k=d$; it is from $S\subset C$ to $T\subset C$ if $c_1\in S$ and $c_k\in T$.

Lemma 13 *For every path* $\widehat{p}:[a,b]\to\Omega$ *from* $s\in C$ *to* $d\in C$ *and* $\varepsilon\ge 2n\alpha$, *if* $B[\mathrm{range}(\widehat{p}),\varepsilon]\subset\Omega$, *then there exists a path* $p=\langle c_1,\ldots,c_k\rangle$ *in C from s to d which is contained in* $B[\mathrm{range}(\widehat{p}),\varepsilon]$.

Proof First, assume that $\alpha\in(2r_\Omega(C),n^2r_\Omega(C)]$. Let $\widehat{\varepsilon}>r_\Omega(C)$ be such that $2r_\Omega(C)<2\widehat{\varepsilon}<\alpha$. Then, by the definition of number $r_\Omega(C)$, $\mathrm{range}(\widehat{p})\subset\Omega\subset\bigcup_{c\in C}B(c,\widehat{\varepsilon})$. Define $C_0=\{c\in C:B(c,\widehat{\varepsilon})\cap\mathrm{range}(\widehat{p})\neq\emptyset\}$ and let $\mathcal{B}=\{B(c,\widehat{\varepsilon}):c\in C_0\}$. Then $\bigcup\mathcal{B}$ is connected, since it is a union of connected sets, each intersecting a connected set $\mathrm{range}(\widehat{p})\subset\bigcup\mathcal{B}$. Also $\bigcup\mathcal{B}\subset B[\mathrm{range}(\widehat{p}),\varepsilon]$ since each $B(c,\widehat{\varepsilon})\in\mathcal{B}$ intersects $\mathrm{range}(\widehat{p})$ and has diameter $2\widehat{\varepsilon}<\alpha\le\varepsilon$. Let $\mathcal{D}$ be the collection of all balls $B(c,\widehat{\varepsilon})\in\mathcal{B}$ such that there exists a sequence $\langle c_1,\ldots,c_k\rangle$ in C_0 from *s* to *c* for which $\|c_{i+1}-c_i\|<2\widehat{\varepsilon}$ for every $i=1,2,\ldots,k-1$. Notice that $\mathcal{D}=\mathcal{B}$.

Indeed, assume by way of contradiction that $\mathcal{E}=\mathcal{B}\setminus\mathcal{D}\neq\emptyset$. Then the sets $\bigcup\mathcal{D}$ and $\bigcup\mathcal{E}$ are non-empty and open. They cannot be disjoint, since this would constitute a partition of a connected set $\bigcup\mathcal{B}$. So, there are balls $B(c,\widehat{\varepsilon})\in\mathcal{D}$ and $B(\widehat{c},\widehat{\varepsilon})\in\mathcal{E}$ intersecting each other. But this means that $B(\widehat{c},\widehat{\varepsilon})\in\mathcal{D}$, since this is justified by a path $\langle c_1,\ldots,c_k,\widehat{c}\rangle$, where $\langle c_1,\ldots,c_k\rangle$ justifies that $B(c,\widehat{\varepsilon})\in\mathcal{D}$. So, $\mathcal{E}=\emptyset$ and indeed $\mathcal{D}=\mathcal{B}$.

Now, since $B(d,\widehat{\varepsilon})\in\mathcal{B}=\mathcal{D}$, we conclude that there exists a sequence $p=\langle c_1,\ldots,c_k\rangle$ in C_0 from *s* to *d* for which $\|c_{i+1}-c_i\|<2\widehat{\varepsilon}<\alpha$ for every $i=1,2,\ldots,k-1$. This is our desired path.

Next assume that $C=\Omega_h=\Omega\cap(h\mathbb{Z})^n$ for an $h\in(0,r_\Omega(C)]$. Put $\widehat{\varepsilon}=h$ and let ρ be a max metric on $\mathbb{R}^n$, that is, $\rho(x,y)=\max_{i=1,\ldots,n}|x_i-y_i|$. Let C_0 be the set of all $c\in C$ for which the ρ-open ball $B_\rho(c,\widehat{\varepsilon})=\{x\in\mathbb{R}^n:\rho(c,x)<\widehat{\varepsilon}\}$ intersects $\mathrm{range}(\widehat{p})$ and let $\widehat{\mathcal{B}}$ be the family of all such balls. Notice that $\mathrm{range}(\widehat{p})\subset\bigcup\widehat{\mathcal{B}}$ since for every $x\in\mathbb{R}^n$ there exists a $c\in(h\mathbb{Z})^n$ such that $\rho(c,x)<h$ and if $x\in\mathrm{range}(\widehat{p})$, then such a *c* belongs to *C*, since in such a case we have $c\in B[\mathrm{range}(\widehat{p}),\varepsilon]\cap(h\mathbb{Z})^n\subset\Omega\cap(h\mathbb{Z})^n=C$. So, as above, $\bigcup\widehat{\mathcal{B}}$ is connected. Moreover, $\bigcup\widehat{\mathcal{B}}\subset B[\mathrm{range}(\widehat{p}),\varepsilon]$ since each $B_\rho(c,\widehat{\varepsilon})\in\widehat{\mathcal{B}}$ intersects $\mathrm{range}(\widehat{p})$ and has diameter $2\sqrt{n}\widehat{\varepsilon}<2n\alpha\le\varepsilon$. Let $\widehat{\mathcal{D}}$ be the collection of all ρ-balls $B_\rho(c,\widehat{\varepsilon})\in\widehat{\mathcal{B}}$ such that there exists a sequence $\langle d_1,\ldots,d_m\rangle$ in C_0 from *s* to *c* such that $\rho(d_{i+1},d_i)<2\widehat{\varepsilon}$ for every $i=1,2,\ldots,m-1$. The argument as above shows that $\widehat{\mathcal{D}}=\widehat{\mathcal{B}}$.

[11]For *n* = 2 we get 8-adjacency for any $\alpha\in\left[\sqrt{2}h,2h\right)$ (so, also for $\alpha=\sqrt{3}h$), while for *n* = 3 the 26-adjacency results for any $\alpha\in\left[\sqrt{3}h,2h\right)$.

Now, since $B(d,\widehat{\varepsilon}) \in \widehat{\mathcal{B}}=\widehat{\mathcal{D}}$, there exists a sequence $\langle d_1, \ldots, d_m\rangle$ in C_0 from s to d such that $\rho(d_{i+1}, d_i) < 2\widehat{\varepsilon}$ for every $i = 1, 2, \ldots, m-1$. Fix an $i = 1, 2, \ldots, m-1$ and note that $\rho(d_{i+1}, d_i) \le h$. Since the closed ρ-ball $B_\rho[d_i, h] = \{x \in \mathbb{R}^n : \rho(d_i, x) \le h\}$ is contained in $B[\text{range}(\widehat{p}), \varepsilon]$ and it contains d_{i+1}, there exists a path p_i in $B[\text{range}(\widehat{p}), \varepsilon] \cap (h\mathbb{Z})^n$ from d_i to d_{i+1} where consecutive spels are of distance $h \le \alpha$. Then the path p formed as a consecutive sequence of all paths $p_1, \ldots, p_{m-1}$ is as desired.

Affinity function—Recall that any FC algorithm starts with an *affinity function*—a symmetric function κ defined on $C \times C$ for which the value $\kappa(c, d)$ represents a strength of local connectedness of the spels $c, d \in C$. We will use here an approach similar to that from the paper [18,19] and consider for affinity any symmetric function κ from $C \times C$ into any linearly ordered set $\langle L, \preceq \rangle$; however, in general, we will not assume that κ is *reflexive* (which, in [18,19], is expressed as a property that $\kappa(a,b) \preceq \kappa(c,c)$ for every $a, b, c \in C$). We drop the assumption of reflexivity of κ since only in this setting we can find an FC-type of algorithm representing $\mathcal{A}_\nabla$. Although this change will restrict our ability to cite any prior results concerning the FC theory results, this will be of no consequence to us, since we will not use any such result. In this particular subsection we will assume that $\langle L, \preceq \rangle = \langle [0, \infty], \ge \rangle$. Thus, the strongest connectedness (in the sense of $\preceq$) will be given by the value 0, and the weakest connectedness by ∞. Note that in the literature usually only standard affinities are considered, that is, those with the range $\langle L, \preceq \rangle = \langle [0, 1], \le \rangle$ and such that $\kappa(c, c) = 1$ for every $c \in C$. However, any reflexive affinity κ as above can be translated into a standard affinity by a formula $\kappa_\sigma(c, d) = (g_\sigma \circ \kappa)(c, d) = g_\sigma(\kappa(c, d))$, where $g_\sigma(x) = e^{-x^2/\sigma^2}$ is a Gaussian function for some $\sigma > 0$. In this situation affinities κ and κ_σ are naturally equivalent (lead to strongly equivalent algorithms) in a sense defined precisely in [18,19].

Digital path strength and AFC object—The affinity function κ represents the main parameter of the FC algorithms and can be defined differently for different applications. In the algorithm $\mathcal{A}_\nabla$, the definition of κ will be based on the gradient approximation of f. In general, any AFC algorithm, including $\mathcal{A}_\nabla$, depends on the definition of κ as follows. The strength of a path $p = \langle c_1, \ldots, c_k \rangle$ in C is defined as the $\preceq$-weakest link in p, that is,

$$\mu(p) = \max_{i=1,\ldots,k-1} \kappa(c_i, c_{i+1}). \tag{5}$$

For $\theta \in \mathbb{R}$ and a seed $s \in C$, we define the AFC object $P_{s\theta}$ as

$\{c \in C$: there is a path p in C form s to c with $\mu(p) < \sigma\}$.

In other words, if we denote our algorithm by a symbol $\mathcal{A}_\nabla$, then $\mathcal{A}_\nabla(f, \theta, s) = P_{s\theta}$. Our goal is to show that, for an appropriately defined function κ, this algorithm represents a segmentation model $\mathcal{M}_\nabla$. Notice that if $\widehat{\mu}(p)$ equals $\min_{i=1,\ldots,k-1} \kappa_\sigma(c_i, c_{i+1})$ and we put $\widehat{\theta} = g_\sigma(\theta)$, then we have $P_{s\theta} = \{c \in C$: there is a path p from s to c with $\widehat{\mu}(p) > \widehat{\theta}\}$. This is essentially the usual definition of an AFC object defined with the use of the standard affinity κ_σ, except that we use here the strict inequality > rather than the more common ≥. This change is essential for the proof of our convergence theorem.

Gradient based path strength—Our definition of κ will be based on the formula $|\nabla f(c)|$ for the approximation of the magnitude of the gradient of F at c. It will have a property that, under appropriate assumptions on F, the limit $\lim_{r\Omega(c)\to 0} |\nabla(F \restriction C)(c)|$ converges uniformly

to $|\nabla F(c)|$ in a sense that: for every $\varepsilon > 0$ and compact set $B \subset \Omega$, there is a $\delta > 0$ such that for every finite $C \subset \Omega$ with $r_\Omega(C) < \delta$ and every $c \in C$

$$\|\nabla F(c)| - |\nabla(F \restriction C)(c)\| < \varepsilon \text{ when } |\nabla(F \restriction C)(c)| \in \mathbb{R} \tag{6}$$

and

$$|\nabla(F \restriction C)(c)| \in \mathbb{R} \text{ when } c \in B \cap C. \tag{7}$$

It is relatively easy to find such a formula for functions f defined on the sets $C = \Omega_h$. However, the general case is a bit technical, so we will postpone the actual definition of $|\nabla f(c)|$ till Appendix. In the mean time, we will assume that $|\nabla f(c)|$ is already defined and that it satisfies (6) and (7). From this, we define gradient based affinity by a formula

$$\kappa(c,d) = \begin{cases} \max\{|\nabla f(c)|, |\nabla f(d)|\} & \text{for adjacent } c, d \\ \infty & \text{otherwise} \end{cases} \tag{8}$$

In particular, for the affinity defined this way, formula (5) for the strength of a path $p = \langle c_1, \ldots, c_k \rangle$ reduces to

$$\mu(p) = \max_{i=1,\ldots,k} |\nabla f(c_i)|. \tag{9}$$

The following theorem shows that the algorithm $\mathcal{A}_\nabla$ indeed represents the segmentation model $\mathcal{M}_\nabla$. Note that the assumption of uniform continuity of $|\nabla F|$ is satisfied if F is a restriction of a C^1 function defined on the closure cl(Ω) of Ω.

Theorem 14 *Let $F{:}\Omega \to \mathbb{R}^\ell$ be an idealized C^1 image, where Ω is a convex bounded open subset of $\mathbb{R}^n$. Assume that $|\nabla F|$ is uniformly continuous on Ω. Then for every $\theta > \theta' > 0$, finite set $C \subset \Omega$, and $s \in C$, there exists a $\delta > 0$ such that for every finite set $D \subset \Omega$ containing C for which $r_\Omega(D) < \delta$, we have*

$$C \cap P^F_{s\theta'} \subseteq \mathcal{A}_\nabla\left(F \restriction D, \theta', s\right) \subseteq P^F_{s\theta}.$$

In particular, if $\langle C_i \subset \Omega : i \in \mathbb{N}\rangle \in C_0$, then we have $\lim^*_{i,\eta} \mathcal{A}^\eta_\nabla(F \restriction C_i, \theta, s) = P^F_{s\theta} \cap \bigcup_i C_i$ for every $s \in C_1$, where $\mathcal{A}^\eta_\nabla(f, \theta, s) = \mathcal{A}_\nabla(f, \theta - \eta, s)$.

Proof. Note that it is enough to find separately $\delta' > 0$ and $\delta'' > 0$ for which, respectively, the first and the second inclusions hold, since then, the number $\delta = \min\{\delta', \delta''\}$ guarantees both inclusions.

We will begin with the proof of the first inclusion. So, take a $c \in C \cap P^F_{s\theta'}$. We will show that there is a $\delta_c > 0$ such that, for all finite $D \subset \Omega$,

$$c \in \mathcal{A}_\nabla\left(F \restriction D, \theta', s\right) \text{ provided } C \subseteq D \text{ and } r_\Omega(D) < \delta_c.$$

This will do, since then $\delta' = \min\left\{\delta_c : c \in C \cap P^F_{s\theta'}\right\}$ is as desired.

By Theorem 10, there is a path $\widehat{p}: [a, b] \to \Omega$ from s to c with $\mu(\widehat{p}) < \theta'$. Then the range range($\widehat{p}$) of $\widehat{p}$ is contained in the set $P^F_{s\theta'}$. Since range($\widehat{p}$) is a compact subset of an open set $P^F_{s\theta'}$, there exists an $\varepsilon > 0$ such that the closed ball $B = B[\text{range}(\widehat{p}), \varepsilon] = \{x \in \mathbb{R}^n : \text{dist}(\text{range}(\widehat{p}), x) \le \varepsilon\}$ is a subset of $P^F_{s\theta'}$. Since $|\nabla F(x)| < \theta'$ for every $x \in P^F_{s\theta'}$, the compactness of $B[\text{range}(\widehat{p}), \varepsilon]$ insures that there exists an $\widehat{\varepsilon} > 0$ such that $|\nabla F(x)| < \theta' - \widehat{\varepsilon}$ for every $x \in B[\text{range}(\widehat{p}), \varepsilon]$. By (6) and (7), there exists a $\delta_c \in (0, \varepsilon/2n^3)$ such that, for every finite $D \subset \Omega$ with $r_\Omega(D) < \delta_c$,

$$\|\nabla F(x)| - |\nabla(F \restriction D)(x)\| < \widehat{\varepsilon} \text{ for every } x \in B \cap D.$$

To see that δ_c is as desired, take a finite subset D of Ω containing C with $r_\Omega(D) < \delta_c < \varepsilon/2n^3$. Then $\varepsilon > 2n^3 r_\Omega(D) \ge 2n\alpha$ since, by (4), $\alpha \le n^2 r_\Omega(D)$. So, by Lemma 13, there exists a path $p = \langle c_1, \ldots, c_k \rangle$ in D from s to c contained in $B = B[\text{range}(\widehat{p}), \varepsilon]$. To see that c belongs to $\mathcal{A}_\nabla(F \restriction D, \theta', s)$, it is enough to show that $\mu(p) = \max_{i=1,\ldots,k} |\nabla(F \restriction D)(c_i)| < \theta'$. But $|\nabla(F \restriction D)(c_i)| \le \|\nabla(F \restriction D)(c_i)| - |\nabla F(c_i)\| + |\nabla F(c_i)| < \widehat{\varepsilon} + (\theta' - \widehat{\varepsilon}) = \theta'$ for every $i = 1, \ldots. k$. This finishes the proof of the first inclusion.

For the second inclusion, put $\varepsilon = (\theta - \theta')/2$. Since $|\nabla F(x)|$ is uniformly continuos on Ω, there is a $\delta_1 > 0$ such that $\|\nabla F(c)| - |\nabla F(x)\| < \varepsilon$ for every $c, x \in \Omega$ with $\|x - c\| \le \delta_1$. Use (6) to find a $\delta'' \in (0, \delta_1/n^2)$ such that for every finite $D \subset \Omega$ with $r_\Omega(D) < \delta''$ and every $c \in D$

$$\|\nabla F(c)| - |\nabla(F \restriction D)(c)\| < \varepsilon \text{ provided } |\nabla(F \restriction D)(c)| \in \mathbb{R}.$$

To see that for such chosen δ'' the second inclusion holds, take set D as above, choose a $c \in \mathcal{A}_\nabla(F \restriction D, \theta', s)$, and let $p = \langle c_1, \ldots, c_k \rangle$ be a path in D from s to c such that $\mu(p) = \max_{i=1,\ldots,k} |\nabla(F \restriction D(c_i)| < \theta'$. Let $\widehat{p}: [0, 1] \to \mathbb{R}^n$ be a path from $c_1 = s$ to $c_k = c$ which is a linear segment between any two consecutive spels in p. Then range($\widehat{p}$) $\subset \Omega$, since is convex and contains all spels c_i. We will show that $\mu(\widehat{p}) = \sup_{t \in [0,1]} |\nabla F(\widehat{p}(t))| < \theta$, which will prove that $c \in P^F_{s\theta}$. Thus, let $x = p(t)$ be on the segment joining c_i and c_{i+1}. Notice that $\|x - c_i\| \le \|c_{i+1} - c_i\| \le \alpha \le n^2 r_\Omega(D) < n^2\delta'' < n^2\delta'' < \delta_1$, so $\|\nabla F(c_i)| - |\nabla F(x)\| < \varepsilon$. But $|\nabla F(c_i)| \le \|\nabla F(c_i)| - |\nabla(F \restriction D)(c_i)\| + |\nabla(F \restriction D)(c_i)| < \varepsilon + \theta'$, since $|\nabla(F \restriction D)(c_i)| \le \mu(p) < \theta'$. Therefore, as $\|x - c_i\| < \delta_1$,

$$|\nabla F(x)| \le \|\nabla F(x) - |\nabla F(c_i)\| + |\nabla F(c_i)| < \varepsilon + (\varepsilon + \theta') = \theta,$$

which finishes the proof of the second inclusion.

Let $C = \bigcup_i C_i$. To show that $\lim^*_{i,\eta} \mathcal{A}_\nabla(F \restriction C_i, \theta - \eta, s) = P^F_{s\theta} \cap C$ for appropriate C_i's and s, notice that, by what we have proved, for every $n > 0$ and i, there exists an $i_0 > i$ such that, for every $j > i_0$, we have

$$C_i \cap P^F_{s,\theta-\eta} \subseteq \mathcal{A}_\nabla\left(F \restriction C_j, \theta-\eta, s\right) \subseteq P^F_{s\theta} \cap C.$$

So, $C_i \cap P^F_{s,\theta-\eta} \subseteq \bigcap_{j=1}^{\infty}\bigcup_{k\geq j}\mathcal{A}_\nabla(F \restriction C_k, \theta-\eta, s) \subseteq P^F_{s\theta} \cap C$ for every $n > 0$ and i. Hence $P^F_{s,\theta-\eta} \cap C \subseteq \bigcap_{j=1}^{\infty}\bigcup_{k\geq j}\mathcal{A}_\nabla(F \restriction C_k, \theta-\eta, s) \subseteq P^F_{s\theta} \cap C$. Thus,

$$\begin{aligned} P^F_{s,\theta} \cap C &= \bigcup_{\eta>0} P^F_{s,\theta-\eta} \cap C \\ &\subseteq \bigcup_{\eta>0}\bigcap_{j=1}^{\infty}\bigcup_{k\geq j}\mathcal{A}_\nabla(F \restriction C_k, \theta-\eta, s) \subseteq P^F_{s\theta} \cap C. \end{aligned}$$

So, equation $\lim^*_{i,\eta}\mathcal{A}_\nabla(F \restriction C_i, \theta-\eta, s) = \bigcup_{\eta>0}\bigcap_{j=1}^{\infty}\bigcup_{k\geq j}\mathcal{A}_\nabla(F \restriction C_k, \theta-\eta, s) = P^F_{s\theta} \cap C$ holds by Proposition 9.

Theorem 14 together with the construction presented in Appendix leads to the following corollary.

Corollary 15 *The gradient based AFC algorithm* $\mathcal{A}_\nabla$ *represents the segmentation model* $\mathcal{M}_\nabla$ *in terms of sequences from* $\mathcal{C}_0$ and for the class of all functions F from convex bounded open subsets of Ω *of* $\mathbb{R}^n$ *into* $\mathbb{R}^\ell$ *which can be extended to a* C^1 *function defined on an open set* Ω *containing* cl(Ω).

The algorithm $\mathcal{A}_\nabla$ is robust in the sense of Theorem 11. On the other hand, $\mathcal{A}_\nabla$ is not invariant on either translation or rotation, in the sense of Theorem 12: if i is an isometry of $\mathbb{R}^n$ and $\widehat{f}$ is a composition of i restricted to $\widehat{C}=i^{-1}(C)$ with the digital image $f:C \to \mathbb{R}^\ell$, then $\mathcal{A}_\nabla\left(\widehat{f}, \theta, i[S]\right)$ need not be equal to $i[\mathcal{A}_\nabla(f, \theta, S)]$. However, if $f = F \restriction C$ for an idealized image as in Corollary 15, then the outputs of the algorithm applied to a transformed image, $\mathcal{A}_\nabla\left(\widehat{f}, \theta, i[S]\right)$, and the transformed output of the algorithm applied to the original image, $i[\mathcal{A}_\nabla(f, \theta, S)]$, converge to the same object, as the resolution of the image improves (i.e., when $r_\Omega(C) \to 0$). This is obviously important and a very desirable property.

Why complicated limit lim*?—In Theorem 14, we proved that algorithm $\mathcal{A}_\nabla$ represents model $\mathcal{M}_\nabla$ with respect to the limit notion lim*. Does this representation result hold for a simple limit? The following example shows that the answer to this question is negative. More precisely, we describe an image F for which the limit $L=\lim_i \mathcal{A}_\nabla(F \restriction \Omega_{2^{-i}}, \theta, s)$ exists but its closure is considerably larger than $\mathcal{M}_\nabla(F, \theta, s)$. It is also possible to construct a C^1 function F for which the limit L does not exist. However, such an example must be more complicated than the one provided below.

Example 16 Let $\Omega = (-1, 1)^2$, $F(x, y) = x - x^3$, $s = \langle .5, 0\rangle$, and $\theta = 1$. Then $\mathcal{M}_\nabla(F, \theta, s) = (0, 1) \times (-1, 1)$, while $L=\lim_i \mathcal{A}_\nabla(F \restriction \Omega_{2^{-i}}, \theta, s)$ exists and is dense in the entire Ω.

Proof. Note that $|\nabla F(x, y)| = 1 - 3x^2$ on Ω. It has a maximum value 1 attained on a line $x = 0$. Then indeed $\mathcal{M}_\nabla(F, \theta, s) = (0, 1) \times (-1, 1)$. Moreover, for every $i \geq 1$, $\mathcal{A}_\nabla(F \restriction \Omega_{2^{-i}}, \theta, s) = \Omega_{2^{-i}}$, since $|\nabla F \restriction \Omega_{2^{-i}}|(C$ is always less than the maximum value $\theta = 1$ of $|\nabla F|$ on Ω. (This follows from (A.1), (A.2), and the Mean Value Theorem, as

$$|\nabla F \restriction \Omega_{2^{-i}}|(c) = \left|\frac{F(x_1,0) - F(x_2,0)}{x_1 - x_2}\right| = \left|1 - \frac{x_1^3 - x_2^3}{x_1 - x_2}\right| < 1 \text{ for some } \langle x_1, 0\rangle, \langle x_2, 0\rangle \in \Omega_2{}^{-i}.)$$

Thus, $L = \cup_i \Omega_2{}^{-i}$ exists and is dense in Ω.

4.2 Second algorithm representing model $\mathcal{M}_\nabla$: Malladi-Sethian-Vemuri level set algorithm $\mathcal{A}_{LS}$

In this subsection, we will *argue* that the level set algorithm $\mathcal{A}_{LS}$, which is essentially the fast marching algorithm described by Malladi, Sethian, and Vemuri in [32] (compare with [45, Chapter 17]), represents the segmentation model $\mathcal{M}_\nabla$ in terms of sequences from $\mathcal{C}_R$ and for the appropriate class of C^1 functions $F{:}\Omega \to \mathbb{R}^\ell$. (The representation is only in terms of rectangular scenes forming $\mathcal{C}_R$, since the algorithm of [32] is described only for the rectangular grids.) Thus, both algorithms $\mathcal{A}_\nabla$ and $\mathcal{A}_{LS}$ are asymptotically equivalent in terms sequences from $\mathcal{C}_R$. We use in the above the term "argue" rather than "proof," since some parts of the argument are left as conjectures, that is, without a full proof.

The level set delineation model $\mathcal{M}_{LS}$ of the idealized image is described in terms different from the model $\mathcal{M}_\nabla$. Thus, we will start with its description. The model $\mathcal{M}_{LS}$ is applied to an ideal image $F{:}\Omega \to \mathbb{R}$, where Ω is an open convex bounded subset of $\mathbb{R}^n$. Basically, to use $\mathcal{M}_{LS}$ we pick a smooth simple closed surface Γ_0 (diffeomorphic with the $(n-1)$-dimensional sphere)[12] inside the region that is to be delineated—it plays the role of a seed—and then we let Γ_0 propagate outward until it reaches the boundary of the region we seek. The propagation is controlled by the speed function v which indicates at every point z on the front (i.e., on the current position of the propagated surface) the speed $v(z)$ at which this point propagates in the direction normal to the front. The set of points in Ω that are eventually inside the propagating front represents the output of $\mathcal{M}_{LS}$. More precisely, if R_t represents the set of points of Ω that are strictly inside the front Γ_t at a time $t \geq 0$, then $\mathcal{M}_{LS}$ equals $\cup_{t \geq 0} R_t$. The front Γ_t at time $t \geq 0$ is represented as the zero level set $\{z \in \Omega: \Psi(z, t) = 0\}$ for some function $\Psi{:}\Omega \times [0, \infty) \to \mathbb{R}$. To make region R_t inside the front easier identify, it is also assumed that Ψ is negative inside Γ_t and non-negative outside Γ_t. In other words, $R_t = \{z \in \Omega: \Psi(z, t) < 0\}$. In [32], the authors define Ψ at time $t = 0$ as a signed distance from Γ_0, that is, $\Psi(z, 0) = \text{dist}(z, \Gamma_0)$ for z outside Γ_0, and $\Psi(z, 0) = -\text{dist}(z, \Gamma_0)$ for z inside Γ_0. Then, they find $\Psi{:}\Omega \times [0, \infty) \to \mathbb{R}$ extending $\psi(\cdot, 0)$ as a solution of a PDE described below.

The boundary of the object is defined as the set of points where the image intensity changes rapidly, that is, when the magnitude of the gradient $|\nabla F|$ is "large." To force the front propagation "…to stop in the vicinity of the desired objects' boundaries…" the propagation speed v is defined in such a way that v goes to zero precisely when $|\nabla F|$ approaches the "large" threshold value $\theta \in (0, \infty]$.[13] Neither "large value" θ nor formula for v is uniquely defined in [32]. Formulas (13) and (16) from [32] suggest that the speed should be reduced to zero at the points $z \in \Omega$ when $|\nabla F|(z)$ is equal to the maximum M of $|\nabla F|$ on Ω, which means that $\theta = M$. (The authors of [32] do not explain why such maximum should exist.) Alternatively, formulas (14), (15), and (17) from [32] suggest that the speed should be a product of a positive factor independent of F and a number of the form $(1 + |\nabla F|)^{-1}$; that is, the propagation speed should go to zero only as $|\nabla F|$ goes to ∞, meaning that $\theta = \infty$. The first from these options suggests that $\mathcal{M}_{LS}(F, \Gamma_0)$ is equal to $\mathcal{M}_\nabla(F, M, \Gamma_0)$. The second makes $\mathcal{M}_{LS}(F, \Gamma_0)$ equal $\mathcal{M}_\nabla(F, \infty, \Gamma_0)$ which, for C^1 function F, is equal to the entire Ω. To stop the algorithm associated with $\mathcal{M}_{LS}(F, \Gamma_0) = \mathcal{M}_\nabla(F, \infty, \Gamma_0)$, the authors introduce the

[12] In the plane, it is a smooth simple closed curve.

[13] The quote comes from the first paragraph of [32, Section III]. A similar statement can be also found in [45, page 220].

maximum number of algorithm iterations (see [32, page 164]). In the work presented below, we reconcile both of these approaches by making the value of $\mathcal{M}_{LS}$ dependent on θ and by reducing the propagation speed v to 0, when $|\nabla F|$ reaches θ. As an example, we take $v(x) = (|\nabla F|(x) - \theta)^2$. Then, we define $\mathcal{M}_{LS}(F, \theta, \Gamma_0)$ as the set of all points of that are eventually inside the propagating curve, that is, $\mathcal{M}_{LS}(F, \theta, \Gamma_0) = \bigcup_{t \geq 0} R_t$.

This general setup allows us to relate models $\mathcal{M}_{\nabla}$ and $\mathcal{M}_{LS}$ as follows.

Lemma 17 $\mathcal{M}_{LS}(F, \theta, \Gamma_0) \subset \mathcal{M}_{\nabla}(F, \theta, \Gamma_0)$ for every C^1 image $F: \Omega \to \mathbb{R}^{\ell}$, $\theta \in \mathbb{R}$, and smooth simple closed surface Γ_0 such that $\Gamma_0 \cup R_0 \subset \mathcal{M}_{\nabla}(F, \theta, \Gamma_0)$.

The assumption $\Gamma_0 \cup R_0$ ensures that every point z of the initiation set $\Gamma_0 \cup R_0$ satisfies the thresholding condition $|\nabla F(z)| < \theta$.

Proof. For $\theta \leq 0$, the assumption $\Gamma_0 \subset \mathcal{M}_{\nabla}(F, \theta, \Gamma_0)$ is false and there is nothing to prove. So, assume that $\theta > 0$. Then $\mathcal{M}_{LS}(F, \theta, \Gamma_0)$ is connected, since it is a union of connected sets $\Gamma_0 \cup R_0$ and the trajectories of points $z \in \Gamma_0$. Thus, it is enough to prove that $\mathcal{M}_{LS}(F, \theta, \Gamma_0)$ is a subset of $\Omega(\theta) = \{x \in \Omega: |\nabla F(x)| < \theta\}$, as $\mathcal{M}_{LS}(F, \theta, \Gamma_0)$ is the largest connected subset of $\Omega(\theta)$ containing Γ_0. To see that $\mathcal{M}_{LS}(F, \theta, \Gamma_0)$ is contained in $\Omega(\theta)$, first note that

$$v(z) \neq 0 \text{ for every } z \in \mathcal{M}_{LS}(F, \theta, \Gamma_0). \tag{10}$$

Indeed, take a $z \in \Omega$ with $v(z) = 0$. Then, by our assumption, $z \notin R_0$. Let $\widehat{t} = \sup\{t \geq 0: z \notin R_t\}$. If $\widehat{t} = \infty$, then $z \notin \mathcal{M}_{LS}(F, \theta, \Gamma_0)$ as desired. If $\widehat{t} < \infty$, then z belongs to Γ_t for every $t \geq \widehat{t}$, since the speed of propagation of front at z is 0. Then, once again, $z \notin \mathcal{M}_{LS}(F, \theta, \Gamma_0)$, finishing the proof of (10).

Now, property (10) implies that $|\nabla F(z)| \neq \theta$ for every $z \in \mathcal{M}_{LS}(F, \theta, \Gamma_0)$. If there was a point $z \in \mathcal{M}_{LS}(F, \theta, \Gamma_0)$ with $|\nabla F(z)| \geq \theta$, then the open sets $\{z \in \mathcal{M}_{LS}(F, \theta, \Gamma_0): |\nabla F(z)| < \theta\}$ and $\{z \in \mathcal{M}_{LS}(F, \theta, \Gamma_0): |\nabla F(z)| > \theta\}$ would be non-empty and they would form a partition of $\mathcal{M}_{LS}(F, \theta, \Gamma_0)$, contradicting its connectedness. Therefore, $\mathcal{M}_{LS}(F, \theta, \Gamma_0) \subset \Omega(\theta)$ and $\mathcal{M}_{LS}(F, \theta, \Gamma_0) \subset \mathcal{M}_{\nabla}(F, \theta, \Gamma_0)$.

The above proof is topological in nature. The other inclusion is also true, but its proof depends on some missing details concerning the definition of $\mathcal{M}_{LS}$. In particular, we need to clarify the meaning of front propagation, as described in [32]. For every point $z \in \Gamma_0$, let $T_z: [0, \infty) \to \Omega$ be a trajectory of z propagated according to the rules described above. Then $\Psi(T_z(t), t) = 0$ for every z and t. So, its derivative $\frac{d}{dt}\Psi(T_z(t), t) = 0$ is also equal to 0. By using chain rule, it is easy to transform this last equation (see [32] or [45]) to $\frac{\partial \Psi}{\partial t}(T_z(t), t) + v(T_z(t)) \cdot |\nabla \Psi|(T_z(t), t) = 0$, where the gradient $\nabla\Psi$ concerns only spatial variables. In particular, any solution of the PDE

$$\frac{\partial \Psi}{\partial t}(x, t) + v(x) \cdot |\nabla \Psi|(x, t) = 0, x \in \Omega, t \geq 0 \tag{11}$$

with the initial condition $\Psi(\cdot, 0) = \Psi_0$ leads to the unique front propagation and the model $\mathcal{M}_{LS}$.

Finding a solution to (11) is not a simple matter since, even in very simple cases, it does not need to have a smooth solution. (See e.g. [45].) However, it always has a weak solution, that is, one that satisfies (11) at all points at which it is differentiable. Although, a weak solution does not need to be unique, the *viscosity solution*, introduced by Crandall and Lions [24], is unique and this is the solution of (11) chosen in [32].

The viscosity solution of (11) is defined in [32] as a limit $\Psi = \lim_{\varepsilon \to 0+} \Psi_\varepsilon$, where Ψ_ε is a solution of

$$\frac{\partial \Psi}{\partial t} + (1 - \varepsilon \mathcal{K})\, v \cdot |\nabla \Psi| = 0, \Psi(\cdot, 0) = \Psi_0, \tag{12}$$

where $\mathcal{K} = \nabla \cdot \frac{\nabla \Psi}{|\nabla \Psi|}$, the divergence of the unit normal vector $\frac{\nabla \Psi}{|\nabla \Psi|}$ to the front, is the curvature of the level surface. The theoretical value of this approach is based on the results summarized in the following proposition, which can be found in [27].

Proposition 18 The PDE (12) has a global smooth solution for smooth Γ_0. Moreover, the solutions Ψ_ε of (12) converge, as ε goes to 0, to the viscosity solution Ψ for (11).

Using Proposition 18 we can prove the equality between models $\mathcal{M}_\nabla$ and $\mathcal{M}_{LS}$.

Theorem 19 $\mathcal{M}_{LS}(F, \theta, \Gamma_0) = \mathcal{M}_\nabla(F, \theta, \Gamma_0)$ for every C^1 image $F: \Omega \to \mathbb{R}^\ell$, $\theta \in \mathbb{R}$, and smooth simple closed surface Γ_0 such that $\Gamma_0 \cup R_0 \subset \mathcal{M}_\nabla(F, \theta, \Gamma_0)$.

Proof. Inclusion $\mathcal{M}_{LS}(F, \theta, \Gamma_0) \subset \mathcal{M}_\nabla(F, \theta, \Gamma_0)$ was proved in Lemma 17. To prove the other inclusion, let Ψ be the viscosity solution of (11). Its existence is guaranteed by Proposition 18. Take a $z \in \mathcal{M}_\nabla(F, \theta, \Gamma_0)$. We need to show that $z \in \mathcal{M}_{LS}(F, \theta, \Gamma_0)$, that is, that there exists a $t \geq 0$ for which $\Psi(z, t) < 0$.

Let Ω_0 be an open region containing $\{z\} \cup \Gamma_0 \cup R_0$ with its closure cl(Ω) contained in Ω. (For example, if P is a path in Ω from z to Γ_0, and $d > 0$ is a distance from $P \cup \Gamma_0 \cup R_0$ to the complement $\mathbb{R}^n / \Omega$ of Ω, then Ω_0 can be defined as the set of all $x \in \Omega$ with the distance from $P \cup \Gamma_0 \cup R_0$ being $< d/2$.) Let $v_0 = \inf\{v(x) : x \in \mathrm{cl}(\Omega)\}$. Since cl(Ω) is a compact subset of Ω and v is positive on Ω, we have $v_0 > 0$.

Let $\Phi: \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ be the viscosity solution of the boundary value problem $\frac{\partial \Phi}{\partial t}(x, t) + v_0 \cdot |\nabla \Phi|(x, t) = 0$ with the boundary condition Φ(·, 0) being the signed distance from Γ_0. (So, Φ(·, 0) coincides with Ψ(·, 0) on Ω.) This equation represents a constant speed front propagation. The existence of Φ is, in particular, guaranteed by Proposition 18.

Notice that Φ satisfies $|\nabla \Phi|(x, t) = 1$ in the viscosity sense (so, for almost all (x, t)). The explicit proof of this fact can be found in [12, section 3], where the author notices that this result is implicitly contained in [31, theorem 4.2]. Hence, from the PDE, we conclude that $\frac{\partial \Phi}{\partial t}(x, t) = -v_0$, again in the viscosity sense. From here we obtain (compare [12, lemma 4.4]) the equation $\Phi(x, t) = \Phi(x, 0) - v_0 t$ for all (x, t). In particular, for every x there is a $t_x > 0$ with $\Phi(x, t_x) < 0$.

Now, since Φ is a viscosity of (11), it is also a viscosity of (11) when restricted to Ω_0, since the definition of a viscosity is local in its nature. (See [27].) Similarly, Φ restricted to Ω_0 is a

viscosity of $\frac{\partial \Phi}{\partial t}(x,t)+v_0\cdot|\nabla\Phi|(x,t)=0$ with the same boundary condition. Moreover, since $v(x) \geq v_0$ on Ω, Φ is a supersolution of (11) on Ω_0. (Compare [12, lemma 3.1].) Therefore,$\psi \leq \Phi$ on $\Omega_0 \times [0,\infty)$.

Finally, since $z \in \Omega_0$, we conclude that $\Psi(z, t_z) \leq \Phi(z, t_z) < 0$, that is, that indeed $z \in \mathcal{M}_\nabla(F,\theta,\Gamma_0)$.

The delineation algorithm $\mathcal{A}_{LS}$ described in [32] depends on $\varepsilon > 0$ and finds its value from a numerical approximation for Ψ_ε. In order to prove formally that $\mathcal{A}_{LS}$ represents $\mathcal{M}^\varepsilon_{LS}$, we should first show that,

- (•) $\mathcal{A}^\varepsilon_{LS}$ represents $\mathcal{M}^\varepsilon_{LS}$ in terms of sequences from $\mathcal{C}_R$ and for the class $\mathcal{F}$ of C^1 images with uniformly continuous gradient

in a sense that for every appropriate $F:\Omega \to \mathbb{R}^\ell$, Γ_0, and the parameters θ, ε, $h > 0$ the limit $\lim^*_{i,\eta}\mathcal{A}^\varepsilon_{LS}\left(F \restriction \Omega_{h/2^i}, \theta-\eta, \Gamma_0\right)$ exists and is a dense subset of the model $\mathcal{M}^\varepsilon_{LS}(F,\theta,\Gamma_0)$ defined as $\{z \in \Omega: \psi_\varepsilon(z, t) < 0 \text{ for some } t \geq 0\}$. This seems to follow from the general theory of solving PDE's by finite approximations. (We mean here the fact that the numerical approximations, calculated by $\mathcal{A}^\varepsilon_{LS}$, convergence to Ψ_ε. This, in general, may not immediately translate to a convergence of the segmented objects, as indicated by Example 16. But, the general setup of our limit structure actually ensures that this will not be the issue in this case.) However, since we cannot point out to any specific theorem that implies (•), we leave it here as a conjecture and state the final corollary as a conditional statement, that depends on (•).

Notice that Proposition 18 and Theorem 19 imply that $\lim_{\varepsilon\to 0+}\mathcal{M}^\varepsilon_{LS}=\mathcal{M}_{LS}$. Therefore, if we define $\mathcal{A}^{\varepsilon,\eta}_{LS}(f,\theta,\Gamma_0)=\mathcal{A}^\varepsilon_{LS}(f,\theta-\eta,\Gamma_0)$, then for appropriate $F\,\Gamma_0$, and θ, we have $\lim^\dagger_{i,\eta,\varepsilon}\mathcal{A}^{\varepsilon,\eta}_{LS}\left(F \restriction \Omega_{h/2^i},\theta,\Gamma_0\right)=\lim_{\varepsilon\to 0+}\mathcal{M}^\varepsilon_{LS}(F,\theta,\Gamma_0)=\mathcal{M}_{LS}(F,\theta,\Gamma_0)$. In particular,

if (•) holds true, then $\mathcal{A}_{LS}$ represents $\mathcal{M}_{LS}$.

This can be rephrased as follows.

Corollary 20 Assume that (•) holds true. Then the algorithms $\mathcal{A}_\nabla$ and $\mathcal{A}_{LS}$ are asymptotically equivalent in terms of sequences from $\mathcal{C}_R$ and in the class $\mathcal{F}$ of all C^1 images $F:\Omega \to \mathbb{R}^\ell$ having uniformly continuous gradient and such that $\Omega \subset \mathbb{R}^n$ is bounded, open, and convex. In this class, both these algorithms represent model $\mathcal{M}_\nabla=\mathcal{M}_{LS}$.

5 Experiments, discussion, and conclusions

Experiments

Having proved their model equivalence, we wanted to examine how this equivalence is manifested in actual image segmentation by using $\mathcal{A}_\nabla$ and $\mathcal{A}_{LS}$. So, we compared algorithms $\mathcal{A}_\nabla$ and $\mathcal{A}_{LS}$ at the experimental level. The goal in this paper is not really a formal empirical evaluation of the segmentations in a comprehensive manner. Therefore, we provide practical qualitative examples illustrating the stronger theoretical results. Figures 1 and 2 demonstrate the results from two experiments made on 2D images from two medical applications.

In the first experiment, we applied the algorithms to a 2D T2-weighted brain MR image (data obtained from the BrainWeb repository [21]), Fig. 1(a), to delineate the white matter object. The image had 20% background non uniformity and 3% noise [21]. The level set

results, displayed in Fig. 1(c), were obtained with a version of the algorithm $\mathcal{A}_{LS}$ implemented in the open source software ITK [55]. This algorithm has four steps: (i) it applies a Gaussian filter to the original image; (ii) it calculates the gradient magnitude of the filtered image; and (iii) it applies to this image f a non-linear filter and transforming it to $\widehat{f}$ by a formula $\widehat{f}(c) = (Max - Min) \cdot \left(1+e^{-(f(c)-\beta)/\alpha}\right) + Min$, where *Min* and *Max* are the minimum and the maximum of the input image f, respectively, and the parameter values that we used were $\alpha = -0.3$ and $\beta = 2$. To this modified image $\widehat{f}$ the curve propagation step is applied. The results of the application of the fuzzy connectedness algorithm $\mathcal{A}_{\nabla}$ to the same image $\widehat{f}$ are presented in Fig. 1(b). To make the comparison fair, we calculated the path connectivity strength from the filtered gradient image $\widehat{f}$ from step (iii) described above, rather than from the original gradient magnitude image f. The subtle differences seen in the delineated objects are due to different approximations involved in the otherwise equivalent algorithms.

In the second experiment, depicted in Fig. 2, we applied the same procedure as in the first experiment, and used algorithms $\mathcal{A}_{\nabla}$ and $\mathcal{A}_{LS}$ on a 2D chest CT image shown in Fig. 2(a) to segment the liver. To make algorithm outputs easier to interpret, we include in Fig. 2(d) the "true" segmentation of the liver from Fig. 2(a), which was manually drawn by an expert. Figs. 2(b) and 2(c) show the liver delineated with algorithms $\mathcal{A}_{\nabla}$ and $\mathcal{A}_{LS}$, respectively. Most of the holes which appear as black spots in the segmentation results of Figs. 2(b) and 2(c) actually represent blood vessels inside the liver. Since the manual segmentation did not exclude these regions (which is really difficult and laborious to accomplish), we filled all topological holes in the results of Figs. 2(b) and 2(c). The results are displayed in Figs. 2(e) and 2(f). A behavior similar to that observed in Fig. 1 can be seen in the results of the second experiment as well. The differences are due only to the different approximations involved in the otherwise equivalent algorithms.

Differentiability issue

In models from the previous section, we assumed that the image intensity function F is differentiable, so that the thresholding function $G(x) = |\nabla F|(x)$ is well defined. However, the function $\bar{G}(x) = \limsup_{z \to x} \frac{\| F(z) - F(x) \|}{\| z - x \|}$ is always well defined and agrees with $G(x)$ whenever $G(x)$ is well defined. Thus, in principle, we can define the model $\mathcal{M}_{\nabla}$ for arbitrary image intensity functions. However, in order to prove that $\mathcal{A}_{\nabla}$ represents $\mathcal{M}_{\nabla}$, some version of property (6) must be satisfied. Similarly, there is no chance for a good behavior of $\mathcal{A}_{LS}$ and $\mathcal{M}_{LS}$ in a general setting. Nevertheless, the representation theorem remains true if we require the limit in $\bar{G}(x)$ to exist, but allow it to be infinite. Moreover, the proof of the representation results can be repeated under weaker assumptions than full continuity of $\bar{G}$.

General segmentation theory

The general theoretical framework formulated in Sections 2 and 3 sets up the stage for theoretical analyses and comparisons of different delineation and segmentation algorithms, independent of the framework in which they were originally described. The comparison in Section 4 is just the first such example. We are currently working on similar analysis concerning other segmentation methods: different forms of level set related algorithms, different versions of fuzzy connectedness algorithms, as well as algorithms that use graph cut and watershed frameworks.

Our future work in this direction will also include the investigation of noisy digital images. Actually, allowing the random noise component will hardly change the theory, if one treats digital images as modeled via the formula from Remark 4, used with a fixed Gaussian kernel K. However, for the arguments to work in such a setting, it is necessary to assume that the noise level remains unchanged even when the resolution of digital image acquisition increases. The most commonly occurring non-random components of the image artifacts, viz., blur and background intensity non-uniformity (a slow-varying image intensity component that modulates the observed image intensity), can be modeled as

$$f(c)=\frac{\int_{\mathbb{R}^n}\beta(x)\,[F(x)\cdot K(x-c)]}{\int_{\mathbb{R}^n}K(x)}+\gamma(x)$$

, where β and γ model the background variation component in a general manner. (In MR imaging, for example, it is known that this component is multiplicative, so that $\gamma(x) = 0$.) We will use this representation of digital images in our future extension of the framework presented in this paper.

A Gradient magnitude $|\nabla f|$

Let $C=\Omega_h=\Omega\cap(h\mathbb{Z})^n$ for some $h > 0$ and $\Omega \subset \mathbb{R}^n$, and let $f:C \to \mathbb{R}^\ell$ be a digital image. For $i = 1, \ldots, n$ let e_i be the unit vector in the direction of the ith variable. For $c \in C$ we define an approximate partial derivative $D_i f(c)$ with respect to the ith variable as ∞ if none of the spels $c \pm he_i$ belongs to C and by a formula

$$D_i f(c)=\max\left\{\left|\frac{f(c)-f(d)}{h}\right| : d=c\pm he_i\in C.\right\} \tag{A.1}$$

otherwise. Then the approximation of the gradient magnitude is defined as

$$|\nabla f(c)|=|\langle D_1 f(c),\ldots,D_n f(c)\rangle|=\sqrt{\sum_{i=1}^{n}|D_i f(c)|^2}. \tag{A.2}$$

Lemma 21 Let F be a function from an open set $\Omega \subset \mathbb{R}^n$ into $\mathbb{R}^\ell$ and assume that F can be extended to a C^1 function $\widehat{F}$ defined on an open set $\widehat{\Omega}$ containing the closure cl(Ω) of Ω.

The formula given by (A.2) and (A.1) satisfies property (6): for every $\varepsilon > 0$ there is a $\delta > 0$ such that for every finite $C \subset \Omega$ with $r_\Omega(C) <$ and every

$$\|\nabla F(c)|-|\nabla(F\restriction C)(c)\|<\varepsilon \text{ provided } |\nabla(F\restriction C)(c)|\in\mathbb{R},$$

and property (7). Moreover, the appropriate formula can be also found in the general case.

Proof. For every $i = 1, \ldots, n$ let $U_i = \{\langle x, y\rangle \in \text{cl}(\Omega)\times\text{cl}(\Omega): x_j = y_j \text{ for all } j \neq i\}$ define a function q_i from U_i into $\mathbb{R}^\ell$ by a formula

$$q_i(x,y)=\begin{cases}\frac{\widehat{F}(x)-\widehat{F}(y)}{x_i-y_i} & \text{if } x_i\neq y_i,\\ D_i\widehat{F}(x) & \text{otherwise.}\end{cases}$$

The existence and continuity of the partial derivative $D_i\widehat{F}$ implies that q_i is continuous on U_i. (The continuity at points $\langle x, x\rangle$ follows from the Mean Value theorem.) In particular,

since U_i is compact, $|q_i|$ is uniformly continuous. So, for every $\widehat{\varepsilon}>0$ there is a $\delta_i > 0$ such that for every $c \in \Omega$

$$\left|\left|\frac{F(c) - F(c+h_0 e_i)}{h_0}\right| - |D_i F(c)|\right| < \widehat{\varepsilon} \tag{A.3}$$

for every real number h_0 with $0 < |h_0| < \delta_i$ for which $c + h_0 e_i \in \Omega$.

Let M_i be the largest value between the numbers $\{|q_i(x, y)|: \langle x, y\rangle \in U_i\}$. It is finite, since U_i is compact. Let $M = \max_{i=1,\ldots,n} M_i$. Since function $g(x_1, \ldots, x_n) = \sqrt{x_1^2 + \ldots + x_n^2}$ is uniformly continuous on $[0, M]^n$, there is an $\widehat{\varepsilon}>0$ such that $|g(x) - g(y)| < \varepsilon$ for every $x, y \in [0, M]^n$ for which $\max_i |x_i - y_i| < \widehat{\varepsilon}$. Let δ_i be as above for this particular $\widehat{\varepsilon}$ and let $\delta = \min_i \delta_i$. Then this δ satisfies (6). To insure (7) it is enough to take δ less than the distance between B and the complement of Ω, since then $h \le r_\Omega(C) < \delta$ and for every $c \in B \cap C$ the numbers $c \pm he_i$ are in C.

The idea behind the definition of $|\nabla f(c)|$ in the general case, as well as the argument required to prove (6) and (7), are similar. However, we need to use the directional derivatives in place of partial derivatives. So, assume that for some $c \in C$ and for every $i = 1, \ldots, n$ we have chosen $c_i \in C$ such that the vectors $c_i - c$ are linearly indipendent. Let $u_i = \frac{c_i - c}{|c_i - c|}$ be the unit vector in the direction of $c_i - c$. Then the directional derivative of F at c in the direction of vector u_i is equal to $D_{u1} F(c) = u_i \cdot \nabla F(c)$. Let A be an $n \times n$ matrix whose rows are formed by the coordinates of u_i's. Then $[D_{u_1} F(c), \ldots, D_{u_n} F(c),]^T = A \cdot \nabla F(c)$, where $\nabla F(c)$ is considered as a vertical matrix and T stands for matrix transposition operation. Notice that A^{-1} exists, since vectors u_i's are linearly independent. Thus, $\nabla F(c) = A^{-1} \cdot D_{\mathbf{u}} F(c)$, where $D_{\mathbf{u}} F(c) = [D_{u_1}] F(c), \ldots, D_{u_n} F(c)]^T$. Let $D_{u_i} f(c) = \frac{f(c_i) - f(c)}{|c_i - c|}$, put $D_{\mathbf{u}} F(c) = [D_{u_1}] F(c), \ldots, D_{u_n} F(c)]^T$, and define

$$\nabla f(c) = A^{-1} \cdot D_{\mathbf{u}} f(c).$$

Then $\big||\nabla F(c)| - |\nabla f(c)|\big| \le |\nabla F(c)| - |\nabla f(c)| = |A^{-1} \cdot (D_{\mathbf{B}}) F(c) - (D_{\mathbf{B}}) f(c))|$. Next, we will show that under some assumption on the choice of vectors u_i, there is a constant K such that

$$|A^{-1} \cdot w| \le K|w| \tag{A.4}$$

for every vector $w \in \mathbb{R}^n$. so $\big||\nabla F(c)| - |\nabla f(c)|\big| \le K|D_{\mathbf{u}} F(c - D_{\mathbf{u}} F(c)|$. By the version of (A.3) for arbitrary directional derivative, there exists a $\delta_0 > 0$ such that $|D_{\mathbf{u}} F(c - D_{\mathbf{u}} F(c)| > \varepsilon/K$ whenever $|c_i - c| < \varepsilon_0$. Thus, to finish the proof, we need to describe the choice of c_i's that insures (A.4) and the inequality $|c_i - c| < \delta_0$ whenever $r_\Omega(C) < \delta$.

So let $r = r\Omega(C)$. For every $c \in C$ and $i = 1, \ldots, n$, we will chose $c_i \in C$ in a ball $B(c + 2nre_i, r)$ if it exists. Otherwise we will put $D_{\mathbf{u}} f(c) = \infty$. Note that by the definition of $r_\Omega(C)$, such c_i exists when $c + 2nre_i \in \Omega$ and that this happens for every $c \in B$ provided the distance ρ from B to the complement of Ω exceeds $2nr_{\Omega(C)}$. Thus, to insure (7) it is enough to choose $\delta < \pi/2n$. Now, this choice of ci insures that for $u_i = \frac{c_i - c}{|c_i - c|} = \langle u_i^1, \ldots, u_i^n \rangle$ we have $|u_i^i|/n > |u_i^j|$ for all $j \ne i$. It is not very difficult to see that this condition implies (A.4).

Indeed, first note that for A defined with such u_i^j's we have

$$|A \cdot v| \geq \frac{1}{2n^2}|v|.$$

To see this, let i be such that $|v_i| \geq |v_j|$ for all j. For simplicity of notation, we assume that $i = 1$. Then, since $|u_1^1| > 1/2$ and $|v| \leq \sqrt{n}|v_1|$,

$$\begin{aligned} |A \cdot v| &= \left|\left\langle \sum_j u_1^j v_j, \ldots, \sum_j u_n^j v_j \right\rangle\right| \\ &\geq \left|\sum_j u_1^j v_j\right| \\ &\geq |u_1^1 v_1| - \left(|u_1^2 v_2| + \cdots + |u_1^n v_n|\right) \\ &\geq |u_1^1 v_1| - (n-1)\left(|u_1^1|/n\right)|v_1| \\ &= |u_1^1||v_1|/n > |v_1|/2n \\ &\geq \tfrac{1}{2n^2}|v|. \end{aligned}$$

Now, putting $v = A^{-1} \cdot w$ in the above inequality we get $|w| \leq \frac{1}{2n^2}|A^{-1} \cdot w|$, that is, (A.4) holds for $K = 2n^2$.

Acknowledgments

Partially supported by NIH grant HL 105212.

References

[1]. Ambrosio L, Tortorelli VM. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. Comm. Pure Appl. Math. 1990; 43:999–1036.

[2]. Ambrosio L, Tortorelli VM. On the approximation of free discontinuity problems. Boll. Un. Mat. Ital. 1992; 6-B:105–123.

[3]. Aubert G, Blanc-Féraud L. Some Remarks on the Equivalence between 2D and 3D Classical Snakes and Geodesic Active Contours. International Journal of Computer Vision. 1999; 34(1): 19–28.

[4]. Aubert, G.; Kornprobst, P. Applied Mathematical Sciences. Vol. 147. Springer; 2006. Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations.

[5]. Barles G, Ley O. Nonlocal first-order Hamilton-Jacobi equations modelling dislocations dynamics. Comm. Partial Differential Equations. 2006; 31(8):1191–1208.

[6]. Beucher, S. The watershed transformation applied to image segmentation. 10th Pfefferkorn Conf. Signal and Image Processing in Microscopy and Microanalysis; 1992. p. 299-314.

[7]. Bourdin B, Chambolle A. Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. Numerische Mathematik. 2000; 85(4):609–646.

[8]. Boykov Y, Funka-Lea G. Graph cuts and efficient N-D image segmentation. International Journal of Computer Vision. 2006; 70(2):109–131.

[9]. Boykov, Y.; Kolmogorov, V. Computing geodesics and minimal surfaces via graph cuts. International Conference on Computer Vision, I; Nice, France. 2003.

[10]. Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Machine Intell. 2001; 23(11):1222–1239.

[11]. Braides, A. Γ-*convergence for Beginners*. Oxford University Press; Oxford: 2002.

[12]. Burger, Martin. Growth fronts of first-order Hamilton-Jacobi equations. 2002. see www.sfb013.uni-linz.ac.at/%7Esfb/reports/2002/ps-files/sfb02-08.ps.gz SFB Report 02-8

[13]. Chakraborty A, Staib L, Duncan J. Deformable boundary finding in medical images by integrating gradient and region information. IEEE Trans. Med. Imag. 1996; 15(6):859–870.

[14]. Chambolle A. Image Segmentation by Variational Methods: Mumford and Shah Functional and the Discrete Approximations. SIAM Journal of Applied Mathematics. 1995; 55(3):827–863.

[15]. Chan T, Shen J. Image processing and analysis: variational, PDE, wavelet, and stochastic methods. Society of Industrial and Applied Mathematics. 2005

[16]. Chan TF, Vese LA. Active contours without edges. IEEE Transaction on Image Processing. 2001; 10:266–277.

[17]. Ciesielski KC, Udupa JK, Saha PK, Zhuge Y. Iterative relative fuzzy connectedness for multiple objects, allowing multiple seeds. Computer Vision and Image Understanding. 2007; 107(3):160–182. [PubMed: 18769655]

[18]. Ciesielski KC, Udupa JK. Affinity functions in fuzzy connectedness based image segmentation I: Equivalence of affinities. Computer Vision and Image Understanding. 2010; 114:146–154.

[19]. Ciesielski KC, Udupa JK. Affinity functions in fuzzy connectedness based image segmentation II: Defining and recognizing truly novel affinities. Computer Vision and Image Understanding. 2010; 114:155–166.

[20]. Ciesielski, KC.; Udupa, JK. Region-Based Segmentation: Fuzzy Connectedness, Graph Cut and Related Algorithms. In: Deserno, Thomas M., editor. Recent Advances in Biomedical Image Processing and Analysis. Springer; 2011. in print

[21]. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, Evans AC. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag. 1998; 17(3):463–468.

[22]. Cootes T, Taylor C, Cooper D. Active shape models-their training and application. Computer Vision and Image Understanding. 1995; 61:38–59.

[23]. Cootes T, Edwards G, Taylor C. Active appearance models. IEEE Trans. Pattern Anal. Machine Intell. 2001; 23(6):681–685.

[24]. Crandall MG, Lions PL. Viscosity solutions of Hamilton-Jacobi equations. Trans. Amer. Math. Soc. 1983; 277:1–42.

[25]. Falcão AX, Stolfi J, Lotufo RA. The image foresting transform: Theory, algorithms, and applications. IEEE Trans. on Pattern Analysis and Machine Intelligence. 2004; 26(1):19–29.

[26]. Imielinska C, Metaxas D, Udupa JK, Jin Y, Chen T. Hybrid segmentation of anatomical data. Proceedings of MICCAI. 2001:1048–1057.

[27]. Giga, Y. Monographs in Mathematics. Vol. vol. 99. Birghäuser Verlag; 2006. Surface Evolution Equations. A Level set Approach.

[28]. Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. Int. J. Comput. Vision. 1987; 1:321–331.

[29]. Kechris, AS. Classical Descriptive Set Theory. Springer-Verlag; 1994.

[30]. Kimmel, R. Fast edge integration. In: Osher, S.; Paragios, N., editors. Geometric Level Set Method in Imaging, Vision and Graphics. Springer; 2006.

[31]. Ley O. Lower-bound gradient estimates for first-order Hamilton-Jacobi equations and applications to the regularity of propagating fronts. Adv. Differential Equations. 2001; 6(5):547–576.

[32]. Malladi R, Sethian JA, Vemuri BC. Shape Modeling with Front Propagation: A Level Set Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1995; 17(2):158–175.

[33]. Mansouri A, Mitiche A, Vázquez C. Multiregional competition: A level set extension of region competition to multiple region image partitioning. Computer Vision and Image Understanding. 2006; 101:137–150.

[34]. McInerney T, Terzopoulos D. Deformable models in medical image analysis: A survey. Medical Image Analysis. 1996; 1(2):91–108. [PubMed: 9873923]

[35]. Miranda PAV, Falcão AX. Links Between Image Segmentation based on Optimum-Path Forest and Minimum Cut in Graph. Journal of Mathematical Imaging and Vision. 2009; 35:128–142.

[36]. Morgenthaler D, Rosenfeld A. Multidimensional edge detection by hypersurface fitting. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1981; 3:482–486.

[37]. Mumford D, Shah J. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics. 1989; 42:577–685.

[38]. Najman L, Schmitt M. Watershed for a Continuous Function. Signal Processing. 1994; 38(1):99–112.

[39]. Osher, S.; Fedkiw, R. Level Set Method and Dynamic Implicit Surfaces. Springer; 2003.

[40]. Osher, S.; Paragios, N., editors. Geometric Level Set Methods in Imaging, Vision, and Graphics. Springer; 2006.

[41]. Saha PK, Udupa JK. Relative fuzzy connectedness among multiple objects: Theory, algorithms, and applications in image segmentation. Computer Vision and Image Understanding. 2001; 82(1):42–56.

[42]. Saha PK, Udupa JK, Odhner D. Scale-Based Fuzzy Connectedness Image Segmentation: Theory, Algorithms, and Validation. Computer Vision and Image Understanding. 2000; 77:145–174.

[43]. Schechter, E. Handbook of Analysis and Its Foundations. Academic Press; 1997.

[44]. Sethian JA. Curvature and evolution of fronts. Comm. in Math. Physics. 1985; 101:487–499.

[45]. Sethian, JA. Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science. Cambridge Univ. Press; 1999. Fast Marching Methods and Level Sets Methods.

[46]. Shafarenko L, Petrou M, Kittler J. Automatic watershed segmentation of randomly textured color images. IEEE Transactions on Image Processing. 1997; 6:1530–1544. [PubMed: 18282911]

[47]. Shi, Y.; Karl, WC. Technical Report ECE-2005-02. ECE Department, Boston University; January. 2005 A fast implementation of the level set method without solving partial Differential equations.

[48]. Shi Y, Karl WC. “A fast level set method without solving PDEs,” Acoustics, Speech, and Signal Processing. Proceedings ICASSP'05. 2005

[49]. Trivedi M, Bezdek J. Low-level segmentation of aerial images with fuzzy clustering. IEEE Trans. Systems, Man, and Cybernetics. 1986; 16(4):589–598.

[50]. Udupa JK, Saha PK. Fuzzy connectedness in image segmentation. Proceedings of the IEEE. 2003; 91(10):1649–1669.

[51]. Udupa JK, Saha PK, Lotufo RA. Relative fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002; 24:1485–1500.

[52]. Udupa JK, Samarasekera S. Fuzzy connectedness and object definition: theory, algorithms, and applications in image segmentation. Graphical Models and Image Processing. 1996; 58(3):246–261.

[53]. Vese LA, Chan TF. A multiphase level set framework for image segmentation using the Mumford and Shah model. J. Comput. Vis. 2002; 50(3):271–293.

[54]. Vincent L, Soile P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1991; 13(6):583–598.

[55]. National Library of Medicine Insight Segmentation and Registration Toolkit, ITK, program. ShapeDetectionLevelSetFilter. available from http://www.itk.org/HTML/ShapeDetectionLevelSet.htm

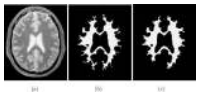

Fig. 1.
The white matter (the region with darker intensities) in a 2D T2-weighted brain MR image (left) delineated with $\mathcal{A}_{\nabla}$ (center) and $\mathcal{A}_{LS}$ (right) algorithms.

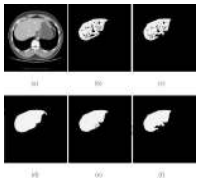

Fig. 2.
(a) A 2D chest CT image. (b) The liver in (a) delineated with algorithm $\mathcal{A}_{\nabla}$. (c) Same as (b), but using algorithm $\mathcal{A}_{LS}$. (d) The "true" segmentation of the liver in (a) delineated manually by an expert. (e)–(f) The same objects as in (b)–(c) from which we removed the 2D holes depicting blood vessels inside the liver.