An Efficient and Provably Correct Algorithm for the Multiscale Estimation of Image Contours by Means of Polygonal Lines

Stefano Casadei and Sanjoy Mitter, *Fellow, IEEE*

Abstract—A large portion of image contours is characterized by local properties such as sharp variations of the image intensity across the contour. The integration of local image descriptors estimated by using these local properties into curvilinear descriptors is a difficult problem from a theoretical viewpoint because of the combinatorially large number of possible curvilinear descriptors. To deal with this difficulty, the notion of compressible graphs is introduced and a contour data model is defined leading to an efficient linear-time algorithm which provably recovers contours with an upper bound on the approximation error.

Index Terms—Curve estimation, edge linking, grouping, image analysis, multiscale edge detection, perceptual organization.

I. INTRODUCTION

MANY computer vision systems require the estimation of image contours to represent the edges between areas of the image with significantly different intensity values (Fig. 1). The standard theory of edge estimation [5], [12], [29] represents edges by means of a collection of points which are characterized as the maxima of the gradient magnitude of the image intensity in the direction of the gradient. Although several methods have been proposed to link these point-like descriptors of edges into curvilinear descriptors [35], [28], [11], [33], a provably correct algorithm for the curvilinear representation of edges has never been proposed. By "provably correct" we mean that the algorithm should generate a list of curves such that every contour in the image satisfying a suitable data model is approximated by a curve in the computed list. In a probabilistic framework, this performance requirement has to be satisfied with high probability.

Provably correct curve-based edge estimation is more difficult than its point-based counterpart since the size of the hypothesis space (i.e., the "volume" of the set of all possible edge descriptors) is exponential in the image size if curve descriptors are used whereas it is only linear if point descriptors are used. Thus developing an efficient (i.e., linear time) provably correct estimation algorithm for curve-based edge representation is a challenging problem. One important assumption that is needed to tackle this problem is that the model which relates image contours to the brightness data be *local*. More specifically, it will be assumed that this model guarantees that every image contour "locally maximizes," in a noise-robust sense, an edginess function $\phi(p, \theta, s)$ which can be computed in linear time from the brightness image by means of a local and spatially homogeneous procedure. Here, p, θ, s denote image location, orientation and scale, respectively. The function $\phi(p, \theta, s)$ can be viewed as a generalization of the intensity gradient. More concretely, an image contour needs to have sufficiently high contrast with respect to the noise amplitude and needs to have a "well-defined" orientation and scale in order to "locally maximize" $\phi(p, \theta, s)$. When these *detectability* conditions occur, the image contour is said to be *supported* by ϕ.

The proposed algorithm for contour estimation is based on a local-to-global strategy in which local (i.e., point-like) contour hypotheses are formulated, locally evaluated, and finally composed into curvilinear descriptors. These computations are carried out by using a graph data-structure, called *edgel-graph*, whose nodes are *edgel-vectors*, that is triples of the form (p, θ, s), (location, orientation and scale). The crucial property to develop a linear time algorithm is ϵ-*compressibility* of the edgel-graph. A graph is ϵ-*compressible*, or compressible with accuracy ϵ, if the Hausdorff distance between any two *regular* paths with the same end-points is less than ϵ.

A more precise description of the proposed algorithm is as follows: 1) Compute an edgel-graph dense enough to contain at least one approximating path near every image contour with high probability. Typically, since the edgel-graph is computed by means of a local procedure, this requirement results in an exponential number of approximating paths for each image contour. 2) For every node of the graph, compute the edginess function $\phi(p, \theta, s)$ and the uncertainty functions $w(p, \theta, s), w_{\mathrm{SC}}(p, \theta, s), w_{\mathrm{or}}(p, \theta, s)$, which, with high probability, provide upper bounds to the contour position, scale and orientation errors (only $w(p, \theta, s)$ and $w_{\mathrm{SC}}(p, \theta, s)$ are used in the current version of the algorithm). These functions are computed by comparing an intensity model of an ideal edge with the intensity data. 3) Reduce the edgel-graph to an ϵ-compressible one by removing certain arcs where $\phi(p, \theta, s)$ is locally minimum. 4) Finally, compute a complete set of maximally long paths in the reduced graph such that any two vertices are connected by exactly one computed path. It can be proven that the computed set of paths approximates every image contour with high probability according to the directed Hausdorff distance.

Fig. 1.   Curve-based representations of image contours computed by the proposed approach (middle) and by more traditional methods (right). The edge representation on the right has been computed by using the Canny's algorithm implementation of the Matlab's image processing toolbox.

This paper, which focusses on the last two parts of the algorithm just described, is organized as follows. Section II reviews related work. Section III contains notations and definitions used throughout the paper. Section IV gives sufficient conditions for compressibility. Section V contains the definition of the algorithm. Section VI introduces the detectability conditions and discusses the performance of the algorithm in a probabilistic setting. Section VII describes some details about the implementation of the algorithm and reports some experimental results. The appendices contain material used to prove the theoretical results.

II. PREVIOUS WORK

The theoretical analysis of the proposed algorithm is based on an error distance defined on a *curve* representation of contours. Previous work modeled a contour as a set of small independent fragments which, essentially, reduces edge detection to a one-dimensional problem. Optimal linear operators for the estimation of the discontinuity point along the gradient have been developed for step edges [5], [12], and more complicated brightness models [29]. Surface fitting methods have also been proposed [16] which are essentially equivalent to linear convolution schemes. Substantial work has been done to assess analytically the one-dimensional estimation performance of these local edge detectors [32], [31], [18]. However, since most of this performance analysis is carried out for point-based models of contours only, the stage of constructing a curve representation from these edge-point fragments is most of the time rather heuristic, with very little theoretical analysis of the overall performance of the algorithm. In the end, performance of the algorithm is usually assessed by means of human judgment [17].

Several other statistical approaches have been proposed for contour estimation and image analysis in general [14], [26], [15], [34]. Most of these methods differ from our statistical approach in that they are based on Bayes' formula. That is, the problem specification must provide a prior density defined on the desired representation and a conditional density of the data given this representation. Estimation consists then in maximizing the *a posteriori* probability of the rep-

resentation given the data. These methods can incorporate global information quite effectively but often result in hard optimization problems. Moreover, these approaches do not usually provide information about the probability distribution of the errors. Most variational and regularization approaches [3], [27], [30], [4] and methods based on criteria such as minimum description length [22] can also be viewed within this statistical framework.

Recently, a statistical approach based on multiscale recursive estimation on trees has been proposed which yields efficient algorithms as well as information about the covariance of the errors [1]. This method has been successfully applied to texture modeling and segmentation.

Wavelets provide an important tool to analyze a multiscale signal [24] and wavelet-based representations can also be used to model nonstationary processes [21].

The importance of multiscale representations for contour estimation has been acknowledged for a long time. Some multiscale algorithms for edge detection proceed in a coarse to fine fashion [25], [2], [30] whereas others are more similar to the approach proposed here in that they emphasize the importance of detecting *all* the relevant scales [23], with priority given to the lowest one [13].

The proposed algorithm exploits the curvilinear nature of contours to augment the information provided by brightness variation. Relaxation labeling has also been used successfully for this purpose [28] as well as "snake" and curve evolution methods [19], [20], [10].

Some of the results in this paper have already been reported and proven within a nonprobabilistic framework, and under the assumption that the scale of the contours is fixed and known [8]. The compressibility condition introduced here is very similar to the stability property discussed in [7] and [8]. The new results presented here generalize the notion of efficient and reliable curve tracking in a graph so that multiple curves can intersect in the image plane, provided that they can be separated by some other slowly varying feature (such as scale). The image contour representation obtained by the proposed algorithm can be used to efficiently hypothesize corners and junctions [6]. It can also be used as an intermediate stage of a more general hierarchical scheme for edge estimation[9].
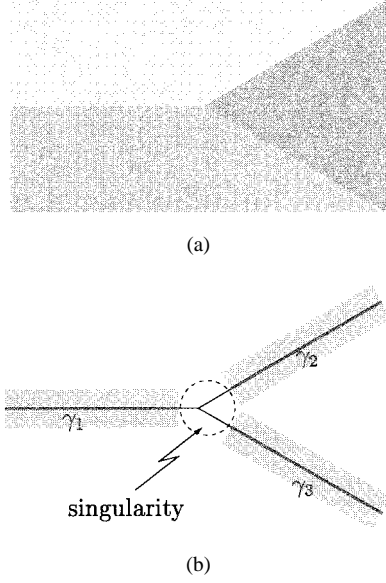
(a)

singularity

(b)

Fig. 2. (a) A Y-junction gives rise to (b) three image contours. The shaded areas indicate the *domains* of the image contours, namely the areas of the image where the data is affected by the presence of the contour (see Section VI-D). Note that the three domains are disjoint.

III. NOTATION, TERMINOLOGY, AND DEFINITIONS

For our purpose, an *image contour* is a curvilinear edge in the image which is sufficiently far away from singularities such as corners and junctions and which has sufficiently strong brightness contrast with respect to the noise amplitude (see Fig. 2).

An image contour is denoted γ and its *trace* (a subset of $\mathbb{R}^2$) is denoted $\sigma(\gamma)$. A set of image contours is denoted Γ. A *flat* contour is an ideal image contour whose trace $\sigma(\gamma)$ is an infinite straight line. The noise-less brightness model of a flat contour is translation invariant along the contour. The orientation of a flat contour γ is denoted $\gamma.\theta$ and the scale of its brightness model is denoted $\gamma.s$.

A contour point hypothesis is represented by a triple $v = (p, \theta, s)$, called an *edgel-vector*, where $p \in \mathbb{R}^2$ is a candidate location in the image plane; $\theta \in [0, 2\pi]$ is a candidate contour orientation; and $s > 0$ is a candidate contour scale. A pair of edgel-vectors $a = (v_1, v_2)$, called an *edgel-arc*, represents a contour fragment hypothesis. A set of edgel-vectors, A, is an *edgel-graph* and its vertices are denoted $V(A)$. A path in an edgel-graph, called an *edgel-path*, is denoted π and represents a curvilinear contour hypothesis.

A. Notation

For the following notation refer to Fig. 3.

$v.p, v.\theta, v.s$	Components of v.
$a.v_1, a.v_2$	Vertices of the arc a.
$\pi.v_0, \cdots, \pi.v_l$	Vertices of the path π.
$\pi.a_1, \cdots, \pi.a_l$	Arcs of the path π.
$V(A)$	Vertices of the edgel-graph A.
$\phi(v) \equiv \phi(p, \theta, s)$	Edginess function.
$w(v) \equiv w(p, \theta, s)$	Position uncertainty.

$w_{\text{SC}}(v) \equiv w_{\text{SC}}(p, \theta, s)$	Scale uncertainty.
$W^{\max}(\pi)$ $= \max\{w(v): v \in V(\pi)\}$	Maximum position uncertainty on a path.
$W^{\max}(A)$ $= \max\{w(v): v \in V(A)\}$	Maximum position uncertainty in the graph.
$\sigma(\gamma) \subset \mathbb{R}^2$	Trace of an image contour.
$\sigma(a) \subset \mathbb{R}^2$	Straight line segment between $a.v_1.p$ and $a.v_2.p$.
$\sigma(\pi) \subset \mathbb{R}^2$	Polygonal line with vertices $\pi.v_0, \cdots, \pi.v_l$.
$\hat{u}(v) = (\cos v.\theta, \sin v.\theta)$	Unit versor along v.
$\hat{u}_\perp(v) = (\sin v.\theta, -\cos v.\theta)$	Unit versor perpendicular to v.
$\hat{u}(a)$	Unit versor parallel to $\sigma(a)$.
$p^-(v) = v.p - w(v) \cdot \hat{u}_\perp(v)$	Edgel location displaced to the left.
$p^+(v) = v.p + w(v) \cdot \hat{u}_\perp(v)$	Edgel location displaced to the right.
$\tau(v) \subset \mathbb{R}^2$	Straight line segment between $p^-(v)$ and $p^+(v)$.
$\tau^i(a)$	$\tau^i(a) = \tau(a.v_i), i = 1, 2$.
$p_i(a)$	$p_i(a) = a.v_i.p; \ i = 1, 2$.
$\beta^-(a) \subset \mathbb{R}^2$	Lateral (left) segment with end points $p^-(a.v_1), p^-(a.v_2)$.
$\beta^+(a) \subset \mathbb{R}^2$	Lateral (right) segment with end points $p^+(a.v_1), p^+(a.v_2)$.
$\beta(a)$	$\beta(a) = \beta^-(a) \cup \beta^+(a)$.
$\beta(\pi)$	$\beta(\pi) = \beta(a_1) \cup \cdots \cup \beta(a_l)$.
$R(a)$	Closed quadrilateral region with vertices $p^-(a.v_1), p^-(a.v_2), p^+(a.v_2), p^+(a.v_1)$.
$R(\pi)$ $= R(\pi.a_1) \cup \cdots \cup R(\pi.a_l)$	Attraction basin of π.
$L_\Delta, \Theta_\Delta, S_\Delta$	Smoothness parameters of a valid edgel-graph (Definition 5 below).
X_0, Θ_0, Δ_0	Accuracy parameters of a covering edgel-graph (Definition 6 below).
$X_1, X_2, \Theta_1, \Theta_2, S_1, S_2$	Parameters used to define the detectability condition (Definitions 10 and 11).
$A|\gamma$	Arcs in A near to γ (see Definition 6).

B. Definitions

Definition 1 (Simple Path): A path π is *simple* if the polygonal line $\sigma(\pi)$ is homeomorphic to a straight line segment (i.e., if it does not self-intersect).

Definition 2 (Regular Path): An edgel-path π is *regular* if it is simple and if $\sigma(\pi) \cap \tau(\pi.v_i) = \{\pi.v_i.p\}, i = 0, \cdots, l$. An edgel-graph is regular if all paths in it are regular.
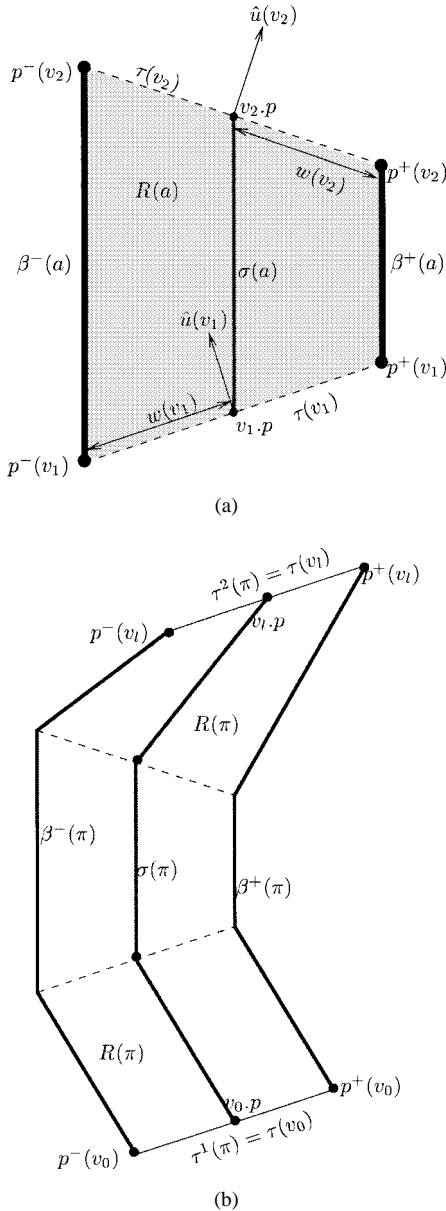
(a)

(b)

Fig. 3.   (a) A valid edgel-arc $a = (v_1, v_2)$. (b) an edgel-path π with vertices $v_1, \cdots, v_l$. The shaded area indicates the *attraction basin* $R(\pi)$.
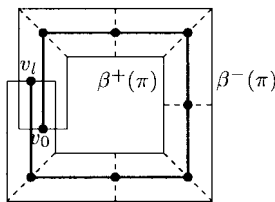

Fig. 4.   An irregular edgel-path.

Fig. 4 shows an irregular edgel-path. Notice that an edgel-cycle (or any edgel-path with repeating vertices) is irregular because $\pi.v_0 = \pi.v_l$ and therefore

$$\sigma(\pi) \cap \tau(\pi.v_0) = \sigma(\pi) \cap \tau(\pi.v_l) \supset \{\pi.v_0.p, \pi.v_l.p\}.$$

Definition 3 (Hausdorff Distance): For any two sets $U_1, U_2 \subset \mathbb{R}^2$, the *directed Hausdorff distance* from U_1 to U_2,

denoted $d(U_1 \to U_2)$ is given by

$$d(U_1 \to U_2) = \sup_{p_1 \in U_1} d(p_1 \to U_2)$$
$$= \sup_{p_1 \in U_1} \inf_{p_2 \in U_2} \|p_1 - p_2\|. \tag{1}$$

The *undirected* Hausdorff distance is given by

$$d(U_1, U_2) = \max\{d(U_1 \to U_2), d(U_2 \to U_1)\}. \tag{2}$$

Definition 4 (Compressibility): An edgel-graph A is *compressible with accuracy* ϵ, or ϵ-*compressible*, if for any two regular edgel-paths π_1, π_2 having the same initial vertex and the same last vertex we have

$$d(\sigma(\pi_1), \sigma(\pi_2)) < \epsilon.$$

Definition 5 (Valid Arc): Let $L_\Delta > 0, \Theta_\Delta \in [0, \pi/2], S_\Delta > 0$ be constants such that

$$S_\Delta < w_{\text{SC}}(v), \quad \forall v \in V(A). \tag{3}$$

An edgel-arc a is said to be *valid*, denoted

$$a \in \texttt{validarc}(L_\Delta, \Theta_\Delta, S_\Delta)$$

if

$$\hat{u}(a.v_1) \cdot \hat{u}(a.v_2) > \cos \Theta_\Delta \tag{4}$$
$$\hat{u}(a) \cdot \hat{u}(a.v_i) > 0, \qquad i = 1, 2 \tag{5}$$
$$|a.v_1.s - a.v_2.s| < S_\Delta \tag{6}$$
$$\|a.v_1.p - a.v_2.p\| < L_\Delta \tag{7}$$
$$\tau(a.v_1) \cap \tau(a.v_2) = \emptyset. \tag{8}$$

A set of arcs A is said to be valid, denoted

$$A \in \texttt{validgraph}(L_\Delta, \Theta_\Delta, S_\Delta)$$

if

$$a \in \texttt{validarc}(\Theta_\Delta, S_\Delta, L_\Delta), \quad \forall a \in A.$$

Definition 6 (Covering Graph): Let γ be a flat contour. Let $X_0 > 0, \Theta_0 \in [0, \pi/2], S_0 > 0$ be constants. The graph A is said to *cover* γ with accuracies X_0, Θ_0, S_0, if there exists a regular path π in A such that

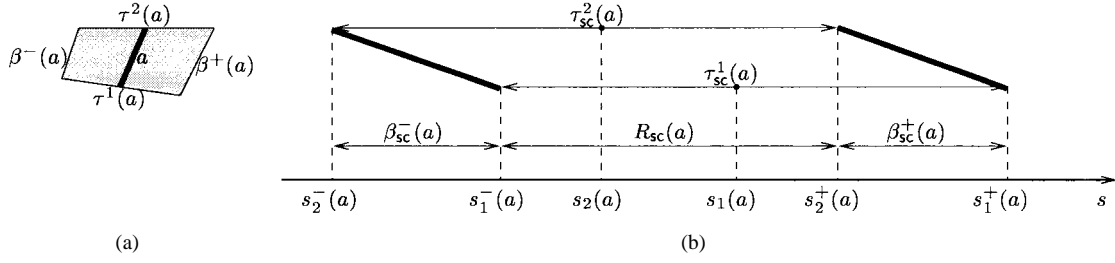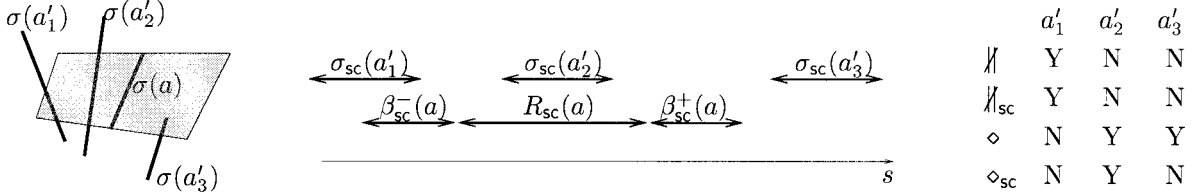$$d(\sigma(\gamma) \to \sigma(\pi)) < X_0 \tag{9}$$

and, for $0 \le i \le l$

$$d(\pi.v_i.p \to \sigma(\gamma)) < X_0; \tag{10}$$
$$d(\pi.v_i.\theta, \gamma.\theta) < \Theta_0 \tag{11}$$
$$|\pi.v_i.s - \gamma.s| < S_0. \tag{12}$$

If A covers γ, the *covering subgraph* of γ, denoted $A|\gamma$, is given by the set of arcs whose two vertices v_1 and v_2 satisfy (10)–(12).

Definition 7 (Divergent Arcs): Let a and a' be edgel-arcs. We say that a' is *nondivergent in space* from a, denoted $a'\|a$, if $\sigma(a) \cap \beta(a') = \emptyset$. If not, then a' is said to be *divergent in space* from a, denoted $a' \nmid a$.

Fig. 5. (a) Space entities of an edgel-arc a. (b) scale entities of a.

Fig. 6. The relationship of a with a_1', a_2', and a_3' is described in the table. For instance, a and a_1' are space divergent, space nonoverlapping, scale divergent and scale nonoverlapping: $a' \not\parallel a \wedge a_1' \not\diamond_{\mathsf{SC}} a \wedge a' \not\parallel_{\mathsf{SC}} a \wedge a_1' \not\diamond_{\mathsf{SC}} a$.

C. Notation and Definitions in Scale-Space

For the following notation refer to Fig. 5.

$$s^-(v) \qquad s^-(v) = v.s - w_{\mathsf{SC}}(v).$$
$$s^+(v) \qquad s^+(v) = v.s + w_{\mathsf{SC}}(v).$$
$$s_i^-(a) \qquad s_i^-(a) = s^-(a.v_i), i = 1, 2.$$
$$s_i^+(a) \qquad s_i^+(a) = s^+(a.v_i), i = 1, 2.$$
$$\tau_{\mathsf{SC}}(v) \qquad \tau_{\mathsf{SC}}(v) = [s^-(v), s^+(v)].$$
$$\tau_{\mathsf{SC}}^i(a) \qquad \tau_{\mathsf{SC}}^i(a) = \tau_{\mathsf{SC}}(a.v_i), i = 1, 2.$$
$$\sigma_{\mathsf{SC}}(a) \qquad \sigma_{\mathsf{SC}}(a) = [a.v_1.s, a.v_2.s].$$
$$R_{\mathsf{SC}}(a) \qquad R_{\mathsf{SC}}(a) = \tau_{\mathsf{SC}}^1(a) \cap \tau_{\mathsf{SC}}^2(a).$$

$\beta_{\mathsf{SC}}(a), \beta_{\mathsf{SC}}^-(a), \beta_{\mathsf{SC}}^+(a)$ are defined as follows:

$$\beta_{\mathsf{SC}}(a) = \begin{cases} (\tau_{\mathsf{SC}}^1(a) \cup \tau_{\mathsf{SC}}^2(a)) \setminus R_{\mathsf{SC}}(a), & \text{if } R_{\mathsf{SC}}(a) \neq \emptyset \\ \mathbb{R}, & \text{if } R_{\mathsf{SC}}(a) = \emptyset. \end{cases} \tag{13}$$

If $R_{\mathsf{SC}}(a)$ is not empty, then $\beta_{\mathsf{SC}}(a)$ is composed of two disjoint connected components, denoted $\beta_{\mathsf{SC}}^-(a), \beta_{\mathsf{SC}}^+(a)$ and given by

$$\beta_{\mathsf{SC}}^-(a) = [\min\{s_1^-(a), s_2^-(a)\}, \max\{s_1^-(a), s_2^-(a)\}]$$
$$\beta_{\mathsf{SC}}^+(a) = [\min\{s_1^+(a), s_2^+(a)\}, \max\{s_1^+(a), s_2^+(a)\}].$$

If $R_{\mathsf{SC}}(a) = \emptyset$, then let $\beta_{\mathsf{SC}}^-(a) = \beta_{\mathsf{SC}}^+(a) = \mathbb{R}$.

Definition 8 (Divergent Arcs): Let a, a' be two arcs in A. Then

- a' is *nondivergent in space from* a, denoted $a' \| a$, if

$$\sigma(a') \cap \beta(a) = \emptyset; \tag{14}$$

- a' is *nondivergent in scale from* a, denoted $a' \|_{\mathsf{SC}} a$, if

$$\sigma_{\mathsf{SC}}(a') \cap \beta_{\mathsf{SC}}(a) \neq \emptyset. \tag{15}$$

Definition 9 (Overlapping Arcs): Let a, a' be two arcs in A. Then

- a' overlaps a in space, denoted $a' \diamond a$, if

$$[a'.v_1.p \in R(a)] \vee [a'.v_2.p \in R(a)]$$
$$\vee [\sigma(a') \cap \tau^1(a) \neq \emptyset \wedge \sigma(a') \cap \tau^2(a) \neq \emptyset]; \tag{16}$$

- a' overlaps a in scale, denoted $a' \diamond_{\mathsf{SC}} a$, if

$$\sigma_{\mathsf{SC}}(a') \subset R_{\mathsf{SC}}(a) \cup \beta_{\mathsf{SC}}(a). \tag{17}$$

IV. SUFFICIENT CONDITIONS FOR COMPRESSIBILITY

A compressible graph is one where all paths between two vertices are close to each other. In a compressible graph it is possible to compute, in linear time, a set of paths which approximate (according to the directed Hausdorff distance) every other path. This is possible because multiple paths between two vertices in a compressible graph can always be safely "compressed" down to a single path.

It turns out that compressibility is a local property of a graph. That is, there exist sufficient conditions for compressibility which depend only on the geometric relationship of pairs of neighboring arcs. In the fixed scale case, where the scale dimension is projected out, two arcs are compatible (meaning that they do not violate the compressibility condition) if they are nondivergent in space (Theorem 2). This rules out the possibility of estimating distinct contours passing through the same neighborhood. Thus junctions can not be recovered from a compressible graph.

In the scale-space generalization (Theorem 4), two arcs are compatible if they are nondivergent in space whenever they overlap in scale (i.e., they have similar scales) and if they are nondivergent in scale whenever they overlap in space. This makes it possible to estimate distinct contours passing through the same neighborhood as long as they can be "separated" by using the scale dimension.

Whereas the proof of the fixed scale case is simple, the proof of the scale-space generalization is quite involved and is
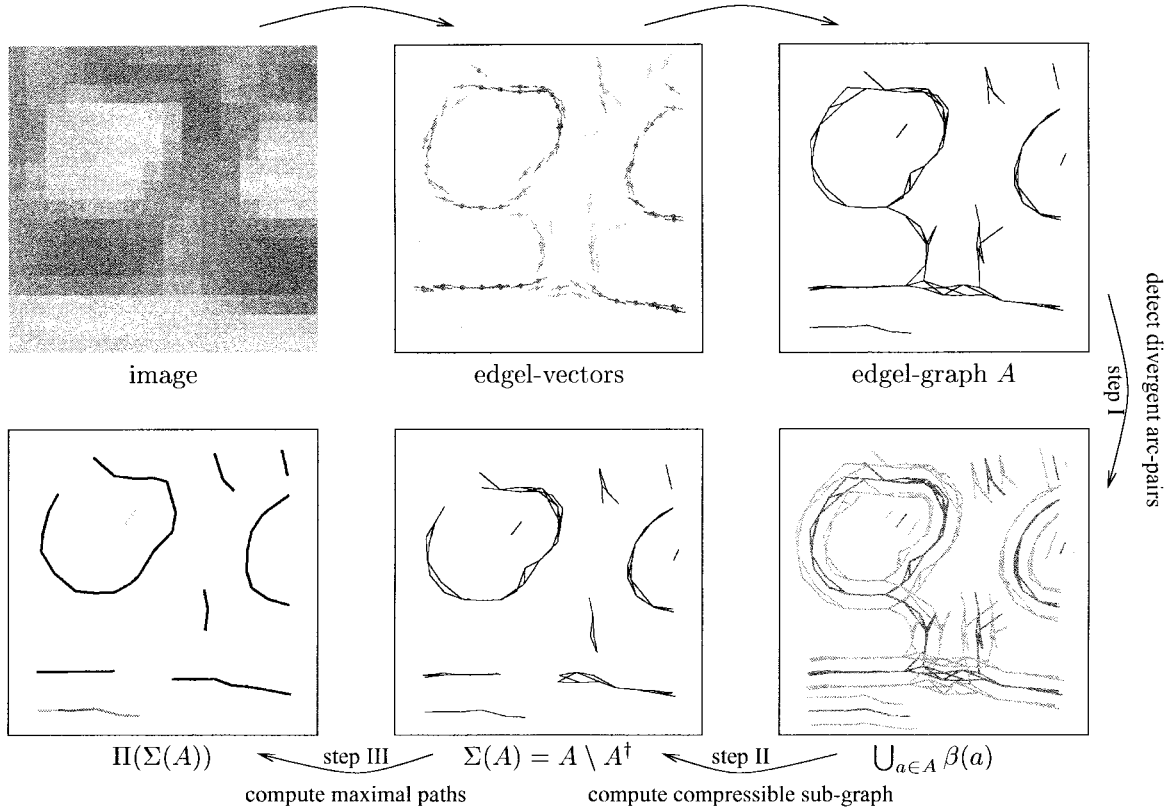
Fig. 7. The three steps of the algorithm (fixed scale case).

reported in Appendix B. The following proposition is needed to prove both theorems.

Proposition 1: Let π be a regular path in A and let $p \in R(\pi)$. Then,

$$d(p \to \sigma(\pi)) < W^{\max}(\pi).$$

Theorem 2 (Sufficient Condition for Compressibility): Let A be a valid edgel-graph. If

$$a' \| a, \quad \forall (a, a') \in A \times A \qquad (18)$$

then, for any two regular paths π_1, π_2 in A with the same initial vertex and final vertex we have

$$d(\sigma(\pi_1), \sigma(\pi_2)) < \min\{W^{\max}(\pi_1), W^{\max}(\pi_2)\}.$$

Corollary 3: If $a' \| a, \forall (a, a') \in A \times A$, then A is compressible with accuracy $W^{\max}(A)$.

Theorem 4 (Sufficient Condition for Compressibility, Multi-scale Generalization): Let A be a valid edgel-graph. If

$$a' \diamond_{\mathrm{SC}} a \Rightarrow a' \| a, \quad \forall (a, a') \in A \times A \qquad (19)$$

$$a' \diamond a \Rightarrow a' \|_{\mathrm{SC}} a, \quad \forall (a, a') \in A \times A \qquad (20)$$

then, for any two regular paths π_1, π_2 in A with the same initial vertex and final vertex

$$d(\sigma(\pi_1), \sigma(\pi_2)) < \min\{W^{\max}(\pi_1), W^{\max}(\pi_2)\}.$$

Corollary 5: If (19), (20) hold then A is compressible with accuracy $W^{\max}(A)$.

Proof of Theorem 2: From (18) we have $\sigma(\pi_1) \cap \beta(\pi_2) = \emptyset$ and $\sigma(\pi_2) \cap \beta(\pi_1) = \emptyset$. Let v_{fi} and v_{la} be the first and last vertex of π_1 and π_2 and let

$$\sigma^{\circ}(\pi_i) = \sigma(\pi_i) \setminus \{v_{\mathrm{fi}}.p, v_{\mathrm{la}}.p\}, \qquad i = 1, 2.$$

Since the paths $\pi_i, i = 1, 2$ are regular, we have

$$\sigma^{\circ}(\pi_i) \cap \tau(v_{\mathrm{fi}}) = \sigma^{\circ}(\pi_i) \cap \tau(v_{\mathrm{la}}) = \emptyset, \qquad i = 1, 2.$$

Thus since the boundary of $R(\pi_i)$ is given by $\partial R(\pi_i) = \beta(\pi_i) \cup \tau(v_{\mathrm{fi}}) \cup \tau(v_{\mathrm{la}})$, we have

$$\sigma^{\circ}(\pi_1) \cap \partial R(\pi_2) = \sigma^{\circ}(\pi_2) \cap \partial R(\pi_1) = \emptyset.$$

Then, for $i = 1, 2$, $\sigma^{\circ}(\pi_i)$ is contained in either $R(\pi_{\bar{i}})$ [where $\bar{i} = i + 1 (\bmod\ 2)$] or in the complement of $R(\pi_{\bar{i}})$ in $\mathbb{R}^2$. Since A is a valid edgel-graph, from (5) it follows that $\sigma^{\circ}(\pi_i) \cap R(\pi_{\bar{i}}) \neq \emptyset, i = 1, 2$, so that $\sigma^{\circ}(\pi_1) \subset R(\pi_2)$ and $\sigma^{\circ}(\pi_2) \subset R(\pi_1)$. The result then follows from Proposition 1. $\square$

V. ALGORITHM

The proposed algorithm takes as input an edgel-graph A, an edginess function $\phi(p, \theta, s)$, a location uncertainty function $w(v) \equiv w(p, \theta, s)$ and a scale uncertainty function $w_{\mathrm{SC}}(v) \equiv w_{\mathrm{SC}}(p, \theta, s)$. Its output is a set of polygonal lines represented by paths in A. These lines, under appropriate assumptions, approximate all the image contours with high probability (see Theorem 9). The three steps of the algorithm are illustrated in Fig. 7.

Step I: The first step of the algorithm consists in evaluating the four relations $\|,\|_{\mathrm{SC}},\diamond,\diamond_{\mathrm{SC}}$ on all pairs of edgel-arcs. For each edgel-arc, only nearby edgel-arcs need to be checked so that this step can be carried out in linear time in the number of arcs (assuming an upper bound on the density of arcs per image area).

Step II: In each pair of arcs violating the compressibility condition, the arc with minimum edginess is marked for removal. The reduced graph, denoted $\Sigma(A)$, can easily be shown to be compressible. This step can also be done in linear time. More precisely, let

$$\phi(a) = \min\{\phi(a.v_1), \phi(a.v_2)\} \qquad (21)$$

and

$$\begin{aligned}
\mathrm{div}(a) &= \{a' \in A\colon a' \nparallel a \wedge a' \diamond_{\mathrm{SC}} a\} \\
&\quad \cup \{a' \in A\colon a \| a' \wedge a \diamond_{\mathrm{SC}} a'\} \qquad (22) \\
\mathrm{div}_{\mathrm{SC}}(a) &= \{a' \in A\colon a' \nparallel_{\mathrm{SC}} a \wedge a' \diamond a\} \\
&\quad \cup \{a' \in A\colon a \nparallel_{\mathrm{SC}} a' \wedge a \diamond a'\}. \qquad (23)
\end{aligned}$$

Then, the reduced compressible graph $\Sigma(A)$ is given by

$$\begin{aligned}
\Sigma(A) &= A \setminus A^{\dagger} \\
A^{\dagger} &= \{a \in A\colon \exists a' \in \mathrm{div}(a) \cup \mathrm{div}_{\mathrm{SC}}(a), \phi(a) \le \phi(a')\}.
\end{aligned} \qquad (24)$$

In the fixed scale version of the algorithm we have (compare with Theorem 2)

$$\begin{aligned}
\Sigma(A) &= A \setminus A^{\dagger} \\
A^{\dagger} &= \{a \in A\colon \exists a' \in \mathrm{div}_0(a), \phi(a) \le \phi(a')\} \qquad (25)
\end{aligned}$$

where

$$\mathrm{div}_0(a) = \{a' \in A\colon a' \nparallel a \vee a \nparallel a'\}. \qquad (26)$$

Step III: A recursive procedure is used to extract, in linear time, one path between any two connected pairs of terminal vertices of $\Sigma(A)$ (a vertex is *terminal* if either its out-degree or its in-degree is zero). By assuming that $\Sigma(A)$ is regular (and hence it does not contain any cycle), it is easy to prove that the resulting set of paths, denoted $\Pi(\Sigma(A))$, approximates every path in $\Sigma(A)$ with accuracy $W^{\max}(A)$ according to the directed Hausdorff distance (see Theorem 7).

More precisely, let $Q(v)$ be the set of paths in $\Sigma(A)$ with initial vertex v given by the following recursive equation (assuming that $\Sigma(A)$ is regular and does not contain any cycle):

$$Q(v) = \begin{cases} \{nil\}, & \text{if } A_{\mathrm{out}}(v) = \emptyset \\ \mathit{compress}\left(\displaystyle\bigcup_{a \in A_{\mathrm{out}}(v)} \bigcup_{\pi \in Q(a.v_2)} a \circ \pi \right), \\ \hspace{2.5cm} \text{if } A_{\mathrm{out}}(v) \ne \emptyset \end{cases}$$

where $A_{\mathrm{out}}(v)$ is the set of arcs incident from v; $a \circ \pi$ denotes the path obtained by prepending the arc a to π; $\mathit{compress}\,(P)$, for any set of paths P, is a subset of P obtained by selecting

a unique representative among all paths with the same end-points. Then let

$$\Pi(\Sigma(A)) = \bigcup_{v \in V_0(\Sigma(A))} Q(v) \qquad (27)$$

where $V_0(\Sigma(A))$ denotes the set of vertices in $\Sigma(A)$ with zero in-degree. The set $\Pi(\Sigma(A))$ can be further compressed by choosing the longest path in each collection of paths having one end-point in common.

A. Results

Proposition 6: $\Sigma(A)$ is compressible with accuracy $W^{\max}(A)$.

Proof: For simplicity, we prove the result in the fixed scale case. The more general proof is similar. Let a, a' be two edgel-arcs in A such that $a' \in \mathrm{div}_0(a)$, i.e., $a' \nparallel a \vee a \nparallel a'$. Let us assume, without loss of generality, that $\phi(a) \le \phi(a')$. Since $a' \in \mathrm{div}_0(a)$, from (25) we have $a \in A^{\dagger}$ and therefore $a \notin \Sigma(A)$. Hence, for every pair of edgel-arcs $a, a' \in \Sigma(A), a' \| a$, that is, $\sigma(a) \cap \beta(a') = \emptyset$. Then from Corollary 3, it follows that $\Sigma(A)$ is compressible with accuracy $W^{\max}(A)$. $\qquad\square$

Theorem 7: If $\Sigma(A)$ is regular then for any path π in $\Sigma(A)$ there exists a path $\hat{\pi} \in \Pi(\Sigma(A))$ such that $d(\sigma(\pi) \to \sigma(\hat{\pi})) < W^{\max}(A)$.

Proof: Let π be a path in $\Sigma(A)$. Since $\Sigma(A)$ is regular and does not contain any cycle, π is a subpath of some maximal regular path π' in $\Sigma(A)$ whose end-points $v_{\mathrm{fi}}, v_{\mathrm{la}}$, are terminal vertices. By definition, $\Pi(\Sigma(A))$ contains one path $\hat{\pi}$ from v_{fi} to v_{la}. Since $\Sigma(A)$ is compressible with accuracy $W^{\max}(A)$ we have $d(\sigma(\pi'), \sigma(\hat{\pi})) < W^{\max}(A)$. Then, from $\sigma(\pi) \subset \sigma(\pi')$, it follows that $d(\sigma(\pi) \to \sigma(\hat{\pi})) < W^{\max}(A).\square$

VI. PERFORMANCE ANALYSIS

A. Assumption on the Edgel-Graph A

In order for the proposed algorithm to be able to estimate a set Γ of image contours, the edgel-graph A must satisfy two requirements:

- it has to be valid (see Definition 5);
- it must cover every image contour $\gamma \in \Gamma$ (see Definition 6).

These two requirements involve six parameters: $L_{\Delta}, \Theta_{\Delta}, S_{\Delta}, X_0, \Theta_0, \Delta_0$. The first one, L_{Δ}, is the maximum allowed distance in the image plane between two consecutive vertices in a path. Similarly, Θ_{Δ} and S_{Δ} are the maximum orientation change and the maximum scale change between two consecutive vertices. The other three parameters, X_0, Θ_0, and Δ_0 are the accuracies with which the image contours Γ are covered by A.

In principle, one can construct an edgel-graph A satisfying these requirements by sampling densely enough the space of edgel-vectors $\mathbb{R}^2 \times [0, 2\pi] \times [0, \infty]$ and by connecting all pairs of edgel-vectors which form a valid edgel-arc. This construction yields a valid edgel-graph which covers all image contours with bounded curvature whose scale parameter changes slowly

enough along the contour. Such an edge-graph will be said to be *fully dense*.

In practice, to reduce computational costs, a much smaller edge-graph A, derived from the brightness image, has to be used. It will be assumed that A is rich enough to cover each image contour $\gamma \in \Gamma$ with high probability. Clearly, a tradeoff exists between the complexity of A and the probability of covering the contours Γ. More precisely, we are going to assume that the probability that A covers Γ with accuracies X_0, Θ_0, Δ_0 is given by

$$\exp\left(-\frac{\ell(\Gamma)}{\lambda(X_0, \Theta_0, \Delta_0)}\right) \qquad (28)$$

where $\ell(\Gamma)$ is the total length of the contours in Γ. This formula states that the violation of the covering condition is a Poisson process indexed by the arc-length of each contour and that the image contours in Γ are independent from each other. This latter assumption is satisfied if we assume that image contours are sufficiently distant from each other, namely if their domains are disjoint.

B. The Detectability Condition

An important result of the work presented here is the definition of a sufficient condition which guarantees that image contours can be recovered in linear time from a given edginess function $\phi(p, \theta, s)$. Roughly speaking, this sufficient condition requires that $\phi(p, \theta, s)$ be locally maximum near image contours in a sense which takes into account the fact that the location of the maxima of $\phi(p, \theta, s)$ fluctuate around their ideal position due to noise. Furthermore, the uncertainty functions $w(p, \theta, s)$ and $w_{\mathrm{SC}}(p, \theta, s)$ must provide sufficiently accurate upper bounds to the amount of fluctuation of these maxima. An image contour which locally maximizes ϕ is the above sense is said to be *supported* by ϕ. The precise definition of this *detectability* condition requires the introduction of more parameters: $X_1, X_2, \Theta_1, \Theta_2, S_1, S_2$, which must satisfy certain constraints [see (29)–(31) and (45)–(46)]. Roughly speaking, the detectability condition is as follows.

- The edginess function $\phi(p, \theta, s)$ is larger in the X_0-neighborhood of γ than it is at a suitable range of distances from γ, denoted $[X_1, X_2]$. A similar property is needed in the orientation and scale dimensions. That is, $\phi(p, \theta, s)$ must be sufficiently low when $\theta \in [\Theta_1, \Theta_2]$ and when $s \in [S_1, S_2]$.
- The uncertainty estimates $w(p, \theta, s)$ and $w_{\mathrm{SC}}(p, \theta, s)$ are upper bounds on the displacement of the maxima of $\phi(p, \theta, s)$ from the true position and scale parameter of the nearest image contour.

More precisely, in fixed scale case, the condition is as follows.

Definition 10 (Detectability Condition): Let X_1, X_2, Θ_1 be such that

$$\Theta_0 < \Theta_1 \qquad (29)$$

$$X_0 < X_1 < \frac{\cos\Theta_1}{2}\, W^{\min}(A) \qquad (30)$$

$$X_2 > X_1 + \max\{L_\Delta, W^{\max}(A)\}. \qquad (31)$$

An image contour γ is said to be *supported* by the edginess function $\phi(p, \theta, s)$ if for any p, θ, s and p', θ', s' we have $\phi(p, \theta, s) > \phi(p', \theta', s')$ whenever

$$d(p \to \sigma(\gamma)) \leq X_0 \qquad (32)$$

$$d(\theta, \gamma.\theta) \leq \Theta_0 \qquad (33)$$

$$|s - \gamma.s| \leq S_0 \qquad (34)$$

$$\|p' - p\| \leq L_{\mathrm{div}} \qquad (35)$$

and at least one of the two following conditions hold true:

$$d(p' \to \sigma(\gamma)) \in [X_1, X_2] \qquad (36)$$

$$d(p', \sigma(\gamma)) \leq X_1 \wedge d(\theta', \gamma.\theta) \in [\Theta_1, 2\pi]. \qquad (37)$$

In (35), L_{div} is the maximum distance between any two points $p \in \sigma(a), p' \in \sigma(a')$, over all arc pairs $(a, a') \in (A \times A)$ which violate the compressibility condition. That is,

$$L_{\mathrm{div}} = \max\{\mathrm{diam}(\sigma(a) \cup \sigma(a')): (a, a'),$$
$$\text{Violate compress. cond.}\}$$

C. Results on the Detectability of Image Contours

From Theorem 7, it follows that an image contour γ covered by A is going to be detected correctly if the reduced compressible graph $\Sigma(A)$ also covers γ. The following results guarantees that the condition introduced in Definition 10 is sufficient for this to happen. The multiscale generalization is given by Theorem 10 in Section VI-E.

Theorem 8 (Preservation of Covering Arcs): Let $A \in$ `validgraph`$(L_\Delta, \Theta_\Delta, S_\Delta)$. Let γ be a flat contour covered by A and supported by ϕ according to Definition 10. Let $\Sigma(A)$ be given by (25). Then $A|\gamma \subset \Sigma(A)$.

Proof: Let γ be a flat contour which satisfies the detectability condition. From (25), we have to prove that $\phi(a) > \phi(a')$ for every $a \in A|\gamma$ and $a' \in \mathrm{div}_0(a)$. Let then $a \in A|\gamma$ and $a' \in \mathrm{div}_0(a)$. One needs to prove that for every $i \in \{1, 2\}$ there exists $j \in \{1, 2\}$ such that

$$\phi(p_i(a), \theta_i(a), s_i(a)) > \phi(p_j(a'), \theta_j(a'), s_j(a')).$$

A stronger statement will be proven, namely that there exists $j \in \{1, 2\}$ for which this inequality holds for both $i \in \{1, 2\}$. Let us make the following substitutions in (32)–(37): $p = p_i(a), \theta = \theta_i(a), s = s_i(a); p' = p_j(a'), \theta' = \theta_j(a'), s' = s_j(a)$. Notice that since $a \in A|\gamma$ we have that (32)–(34) hold with the above substitutions for $i = 1, 2$. Furthermore, for $i = 1, 2$ and $j = 1, 2$, we have $\|p_i(a) - p_j(a')\| \leq L_{\mathrm{div}}$ from the definition of L_{div}. Thus both vertices of a satisfy (32)–(35). It remains to prove that at least one of the two vertices of a' satisfies either (36) or (37). First, let us assume that

$$d(p_j(a') \to \sigma(\gamma)) \leq X_1 \Rightarrow d(\theta_j(a'), \gamma.\theta) < \Theta_1$$
$$j = 1, 2 \qquad (38)$$

so that condition (64) of Proposition 12 (in Appendix A) holds true. Then, from Proposition 12 it follows that one of the two vertices of a' satisfies (36). Let then assume that (38) is false, namely that there exists $j \in \{1, 2\}$ such that $d(p_j(a') \to \sigma(\gamma)) \leq X_1$ and $d(\theta_j(a'), \gamma.\theta) \geq \Theta_1$. Then (37) is satisfied by $a'.v_j$. $\square$

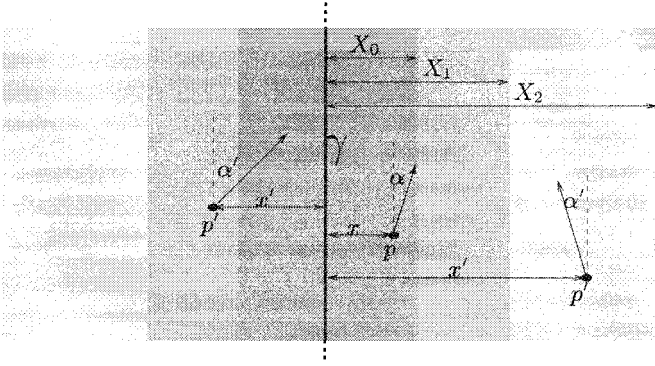Fig. 8. Violation of the detectability condition. One instance of (p', θ') is such that $x' \in [X_1, X_2]$. The other one satisfies $x' < X_1$ and $\alpha' > \Theta_1$.

D. Probability of Misdetection

To find an expression for the probability of misdetecting a set of image contours, let us consider the fixed scale case and let us assume that the uncertainty function w is constant, $w(p, \theta, s) \equiv W$. Let us assume that the edginess function $\phi(p, \theta, s) \equiv \phi(p, \theta)$ in the neighborhood of a flat image contour is given by

$$\phi(p, \theta) = \psi(d(p \rightarrow \sigma(\gamma)), d(\theta, \gamma.\theta)) + \nu(p, \theta) \qquad (39)$$

where $\psi: [0, \infty] \times [0, \pi] \rightarrow \mathbb{R}$ is a monotonically decreasing function of both variables and $\nu(p, \theta)$ is noise. Let (p, θ) be an edgel-vector which satisfies conditions (32) and (33) of Theorem 8 (condition (34) can be ignored). Let x denote the distance from p to $\sigma(\gamma)$ and let $\alpha = d(\theta, \gamma.\theta)$. See Fig. 8 for the notation. A violation of the detectability condition occurs if there exists (p', θ') such that $\phi(p', \theta') \geq \phi(p, \theta); \|p' - p\| \leq L_{\mathrm{div}}$ and either $x' \in [X_1, X_2]$ or $x' < X_1 \wedge \alpha' > \Theta_1$, where $x' = d(p' \rightarrow \sigma(\gamma))$ and $\alpha' = d(\theta', \gamma.\theta)$. By using (39), $\phi(p', \theta') \geq \phi(p, \theta)$ can be rewritten as

$$\nu(p', \theta') - \nu(p, \theta) \geq \psi(x, \alpha) - \psi(x', \alpha'). \qquad (40)$$

Since ψ is a decreasing function and $x < X_0, \alpha < \Theta_0$, we have $\psi(x, \alpha) > \psi(X_0, \Theta_0)$. Similarly, $x' \in [X_1, X_2] \Rightarrow \psi(x', \alpha') < \psi(X_1, 0)$ and $x' < X_1 \wedge \alpha' > \Theta_1 \Rightarrow \psi(x', \alpha') < \psi(0, \Theta_1)$.

Let $P_{\mathrm{err}}^x(X_1)$ be the probability that there exists p', θ' such that

$$\nu(p', \theta') - \nu(p, \theta) \geq \psi(X_0, \Theta_0) - \psi(X_1, 0).$$

Similarly, let $P_{\mathrm{err}}^\theta(\Theta_1)$ be the probability that there exists p', θ' such that

$$\nu(p', \theta') - \nu(p, \theta) \geq \psi(X_0, \Theta_0) - \psi(0, \Theta_1).$$

Then, the probability that the detectability condition is violated at a specific point along the contour is upper bounded by $P_{\mathrm{err}}^x(X_1) + P_{\mathrm{err}}^\theta(\Theta_1)$. Notice that, since $\psi(X_0, \Theta_0) - \psi(X_1, 0) > 0$ and $\psi(X_0, \Theta_0) - \psi(0, \Theta_1) > 0$, in the limit where the variance of the noise goes to zero, the error probability $P_{\mathrm{err}}^x(X_1) + P_{\mathrm{err}}^\theta(\Theta_1)$ also goes to zero. Thus in the noise-free limit case, the algorithm correctly detects all the image contours covered by A.

Recall that the Theorem 8 holds for any X_1, X_2, Θ_1 satisfying (29)–(31). Since we assumed that the uncertainty function w is constant let us substitute $W^{\min}(A) = W^{\max}(A) = W$ in (29)–(31). Furthermore, let us assume that L_Δ is chosen small enough so that $L_\Delta < W$. Then, for any fixed X_2 and W, let us minimize $P_{\mathrm{err}}^x(X_1) + P_{\mathrm{err}}^\theta(\Theta_1)$ over all X_1 and Θ_1 satisfying (29)–(31) and let $P_{\mathrm{err}}(W, X_2)$ be the optimal error upper bound. Notice that W is the accuracy error with which contours are reconstructed (compare with Theorem 7) and X_2 is the maximum distance from the contour $\sigma(\gamma)$ at which the edginess function is required to obey the contour model (39). The X_2-neighborhood of $\sigma(\gamma)$ is the *domain* of γ.

If the image contains a set of image contours Γ, then the domains of these contours must be disjoint so that the edginess function inside each domain is influenced by exactly one contour model (compare with Fig. 2). Such a set of contours is said to be *independent*.

We assume that the violation of the detectability condition is a Poisson process indexed by the contour's arc-length. Thus if Γ is a set of independent contours covered by the input edgel-graph A, an upper bound to the probability of violating the detectability condition (and of misdetecting Γ) is given by

$$1 - \exp\left(-\frac{\ell(\Gamma)}{L_{\mathrm{corr}}} \log P_{\mathrm{err}}(W, X_2)\right) \qquad (41)$$

where L_{corr} is a "correlation length" parameter.

By putting together Theorems 7 and 8, and assuming that the probabilities of violating the covering condition and the sufficient condition of Theorem 8 are given by (28) and (41), respectively, we have the following theorem, which holds under all the assumptions made in this section.

Theorem 9: Let Γ be an independent set of image contours in the image. If $\Sigma(A)$ is regular then, with probability at least

$$\exp\left(-\frac{\ell(\Gamma)}{\lambda(X_0, \Theta_0, \Delta_0)} - \frac{\ell(\Gamma)}{L_{\mathrm{corr}}} \log P_{\mathrm{err}}(W, X_2)\right)$$

for every $\gamma \in \Gamma$ there exists $\hat{\pi} \in \Pi(\Sigma(A))$ such that $d(\sigma(\gamma) \rightarrow \sigma(\hat{\pi})) < X_0 + W$.

Notice that in the noise-free limit case, and if A is fully dense, then the above probability estimate converges to one. Notice also that since the directed Hausdorff distance has been used to measure the error, the approximating path $\hat{\pi}$ can be longer than the actual image contour γ (see Fig. 9).

E. Detectability in Scale-Space

The following result is a generalization of Theorem 8 to the multiscale case. Let

$$L_{\mathrm{div}} = \max_{a \in A} \max_{a' \in \mathrm{div}(a) \cup \mathrm{div}_{\mathrm{SC}}(a)}$$
$$\cdot \max_{i,j \in \{1,2\}} \|p_i(a) - p_j(a')\| \qquad (42)$$

$$S_3 = \max_\gamma \max_{a \in A|\gamma} \max_{a' \in \mathrm{div}(a)} \max_{i \in \{1,2\}} |s_i(a') - \gamma.s| \qquad (43)$$

$$X_3 = \max_\gamma \max_{a \in A|\gamma} \max_{a' \in \mathrm{div}_{\mathrm{SC}}(a)} \max_{i \in \{1,2\}}$$
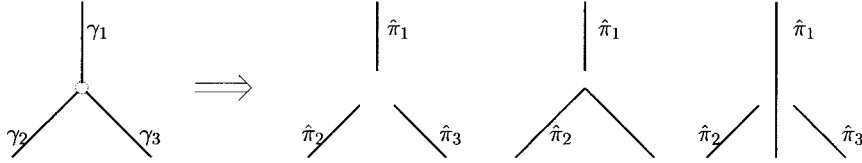$$\cdot d(p_i(a') \rightarrow \sigma(\gamma)). \qquad (44)$$

Fig. 9. Three different ways a T-junction might be approximated by the algorithm.

Definition 11 (Detectability Condition, Multiscale Generalization): Let S_1, S_2 be positive constants such that

$$S_1 \leq W_{\text{SC}}^{\min}(A) - S_\Delta - S_0 \tag{45}$$

$$S_2 \geq W_{\text{SC}}^{\max}(A) + S_\Delta + S_0 \tag{46}$$

and let X_1, X_2, Θ_1 be such that (29)–(31) hold. An image contour γ is said to be *supported* by the edginess function $\phi(p, \theta, s)$ if for any p, θ, s and p', θ', s' we have $\phi(p, \theta, s) > \phi(p', \theta', s')$ whenever

$$d(p \to \sigma(\gamma)) \leq X_0 \tag{47}$$

$$d(\theta, \gamma.\theta) \leq \Theta_0 \tag{48}$$

$$|s - \gamma.s| \leq S_0 \tag{49}$$

$$\|p' - p\| \leq L_{\text{div}} \tag{50}$$

and at least one of the following three sets of conditions hold:

$$d(p' \to \sigma(\gamma)) \in [X_1, X_2] \tag{51}$$

$$|s' - \gamma.s| \leq S_3 \tag{52}$$

$$|s' - \gamma.s| \in [S_1, S_2] \tag{53}$$

$$d(p' \to \sigma(\gamma)) \leq X_3 \tag{54}$$

$$d(p' \to \sigma(\gamma)) \leq X_1 \tag{55}$$

$$d(\theta', \gamma.\theta) \in [\Theta_1, 2\pi] \tag{56}$$

$$|s' - \gamma.s| \leq S_3. \tag{57}$$

Theorem 10 (Preservation of Covering Arcs, Multiscale Generalization): Let $A \in \texttt{validgraph}(L_\Delta, \Theta_\Delta, S_\Delta)$. Let γ be a flat contour covered by A and supported by ϕ according to Definition 11. Let $\Sigma(A)$ be given by (24). Then $A|\gamma \subset \Sigma(A)$.

Proof: The proof is similar to the proof of Theorem 8. Let γ be a flat contour for which the multiscale detectability condition is satisfied. From (24), we have to prove that $\phi(a) > \phi(a')$ for every $a \in A|\gamma$ and $a' \in \text{div}(a) \cup \text{div}_{\text{SC}}(a)$. Let then $a \in A|\gamma$ and $a' \in \text{div}(a) \cup \text{div}_{\text{SC}}(a)$. One needs to prove that for every $i \in \{1, 2\}$ there exists $j \in \{1, 2\}$ such that

$$\phi(p_i(a), \theta_i(a), s_i(a)) > \phi(p_j(a'), \theta_j(a'), s_j(a')).$$

Notice that (47)–(50) are identical to (32)–(35) of Theorem 8. If $a' \in \text{div}(a)$ (and therefore $a' \in \text{div}_0(a)$) then the results follows immediately from Theorem 8 because (52) and (57) follow from (43) and therefore the trigger condition (51) $\wedge$ (52) implies (36) and (55) $\wedge$ (56) $\wedge$ (57) implies (37). Let then $a' \in \text{div}_{\text{SC}}(a)$. It is sufficient to prove that for each vertex i of a there exists a vertex j of a' which satisfies the trigger condition (53) $\wedge$ (54). Equation (54) follows immediately from (44) for $j = 1, 2$. From $a' \in \text{div}_{\text{SC}}(a)$ it follows that

$a' \not\Vert_{\text{SC}} a \vee a \not\Vert_{\text{SC}} a'$. Hence, from the corollary of Lemma 18 in Appendix C, we have that for each i there exists j such that

$$|s_i(a) - s_j(a')| \in [w_{\text{SC}}^{\min}(a, a') - S_\Delta, w_{\text{SC}}^{\max}(a, a') + S_\Delta]$$
$$\subset [W_{\text{SC}}^{\min} - S_\Delta, W_{\text{SC}}^{\max} + S_\Delta].$$

Since $a \in A|\gamma$, we have that $|s_i(a) - s.\gamma| < S_0, i = 1, 2$, and therefore

$$|s_j(a') - \gamma.s| \in [W_{\text{SC}}^{\min} - S_\Delta - S_0, W_{\text{SC}}^{\max} + S_\Delta + S_0].$$

From (45) and (46) it follows then that $|s_j(a') - \gamma.s| \in [S_1, S_2]$. ∎

VII. IMPLEMENTATION DETAILS AND EXPERIMENTS

A. Computation of an Edgel-Graph

One possible method to compute an edgel-graph A which covers contours with high probability is to use a traditional point-based edge detector run at several scales and with very permissive thresholds. The specific method used in our implementation is a variant of the facet-model edge detector [16] and is based on a cubic polynomial approximation, $\hat{\beta}(x)$, of the brightness model across a contour, $\beta(x)$, which is assumed to be a step discontinuity smoothed with a Gaussian filter of variance s

$$\beta(x) = b_1 + (b_2 - b_1) \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/s} e^{-(1/2)u^2} \, du \tag{58}$$

where x denotes the signed distance from the contour. Initially, for each region, a linear brightness model is used to estimate the local orientation of the contour. Then, a cubic brightness model, representing the Taylor expansion of (58), is fitted to a larger rectangular region (aligned to the estimated orientation of the contour) to refine the estimate of the location of the contour and to estimate the scale s and the brightness intensities b_1, b_2. The rectangular regions used for the cubic fit are three pixel long in the direction of the contour and $2l_\perp$ wide across the contour. Three values of the parameter $l_\perp$ have been used: 2, 4, and 6. Then, every pair of valid edgel-vectors is connected with an arc. The second column of Fig. 11 shows two examples of an edgel-graph computed in this way.

B. Edginess Function

The edginess function $\phi(p, \theta, s)$ is computed by least-square fitting the 2-D generalization of the brightness contour model (58), denoted $\beta(p'|p, \theta, s, b_1, b_2)$, to the observed image $I(\cdot)$ in a rectangular region $R(p, \theta, s)$ centered at p, with orientation

θ, height equal to three pixels and width equal to $4s$. Then, $\phi(p, \theta, s)$ is defined as minus the residual of the optimal fit:

$$\phi(p, \theta, s)$$
$$= - \min_{b_1, b_2} \frac{1}{|R(p, \theta, s)|}$$
$$\cdot \left(\int_{R(p, \theta, s)} (I(p') - \beta(p'|p, \theta, s, b_1, b_2))^2 \, dp' \right)^{1/2}. \tag{59}$$

Since $\beta(p'|p, \theta, s, b_1, b_2)$ is a linear function of b_1 and b_2, the minimization in (59) can be performed by linear convolution with two appropriate filters, which depend on θ and s. Notice that, if a Gaussian noise model is assumed, and the proper normalization is chosen, then the quantity in (59) is proportional to the likelihood that a contour with orientation θ and scale s passes through the point p, maximized over the nuisance parameters b_1 and b_2. The behavior of $\phi(p, \theta, s)$ near a contour in the absense of noise is illustrated in Fig. 10.
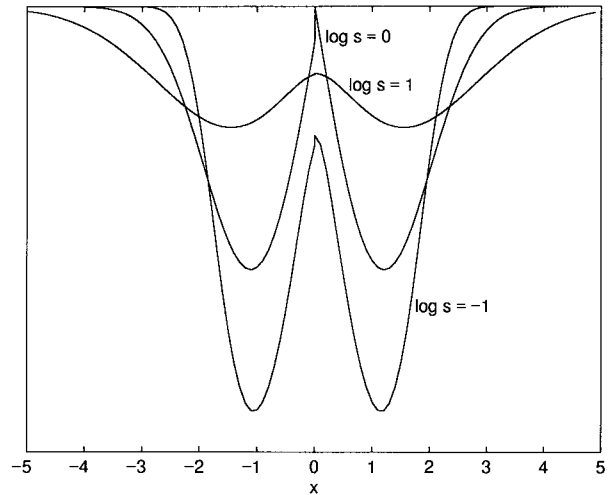
C. Experimental Results

Fig. 11 shows the polygonal lines obtained by the fixed scale version of the proposed algorithm. The results are compared with the output of Canny's algorithm, as implemented by Matlab's image processing toolbox. The function $w(v)$ has been set to 0.75 pixel everywhere. After experimenting with several images, we found, quite surprisingly, that this value is nearly optimal for most images we tested on. Since the observed value of X_0 is a small fraction of a pixel, the bound on the localization error (see Theorem 9) is about one pixel. Computation time for the current implementation is a few milliseconds per pixel.
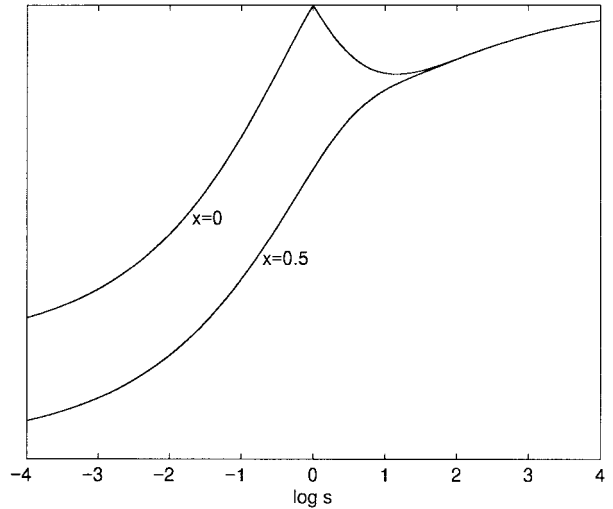
VIII. Conclusions and Future Work

Efficient computation of a curvilinear representation of the edges in an image is a challenging problem from a theoretical perspective because of the exponential size of the hypothesis space. An approach based on the notion of compressibility of a graph has been proposed to deal with curve estimation in a theoretically sound way and a specific contour model together with an estimation algorithm have been proposed to solve the problem in a multiscale setting. Probabilistic analysis of the performance has been carried out under certain probabilistic assumptions on the detectability condition. In the noise-free limit case, the image contours in the model class are recovered with probability one with an upper bound on the approximation error, measured by the directed Hausdorff distance.

On the experimental side, the results are mixed. In fact, whereas on some images the performance of the proposed algorithm compares favorably with existing methods, the extra complication of the algorithm needed to make it theoretically sound does not seem to pay off at the experimental level. However, the multiscale algorithm has not yet been fully implemented and therefore it is premature to draw a definitive conclusion. A possible situation where there might be some practical gains is when an automatic scale selection mechanism

Fig. 10. Sections of the function $\psi(x, 0, s)$, where: $\psi(x, \theta, s)$ denotes the edginess function evaluated near a flat contour γ such that $\gamma.\theta = 0, \gamma.s = 1$ and whose brightness model is given by (58) with $b_1 = -1, b_2 = 1$; x denotes the distance to γ. (a) $\psi(x, 0, s)$ as function of x. (b) $\psi(x, 0, s)$ as a function of $\log s$.

is needed. In fact, the proposed algorithm includes curvilinear constraints in the determination of the scale of the contours.

Appendix A
Result Needed for the Proof of Theorem 8

For any flat contour γ and $X > 0$, let $N_X(\gamma)$ be the neighborhood of γ of radius X

$$N_X(\gamma) = \{p \in \mathbb{R}^2 \colon d(p \to \sigma(\gamma)) \leq X\}.$$

The following proposition has been proven in [8] (see [8, Lemma 1]).

Proposition 11: Let γ be a flat contour; let $a \in A$ be such that

$$p_i(a) \in N_{X_1}(\gamma), \qquad i = 1, 2 \tag{60}$$
$$d(\theta_i(a), \gamma.\theta) < \Theta_1, \qquad i = 1, 2 \tag{61}$$
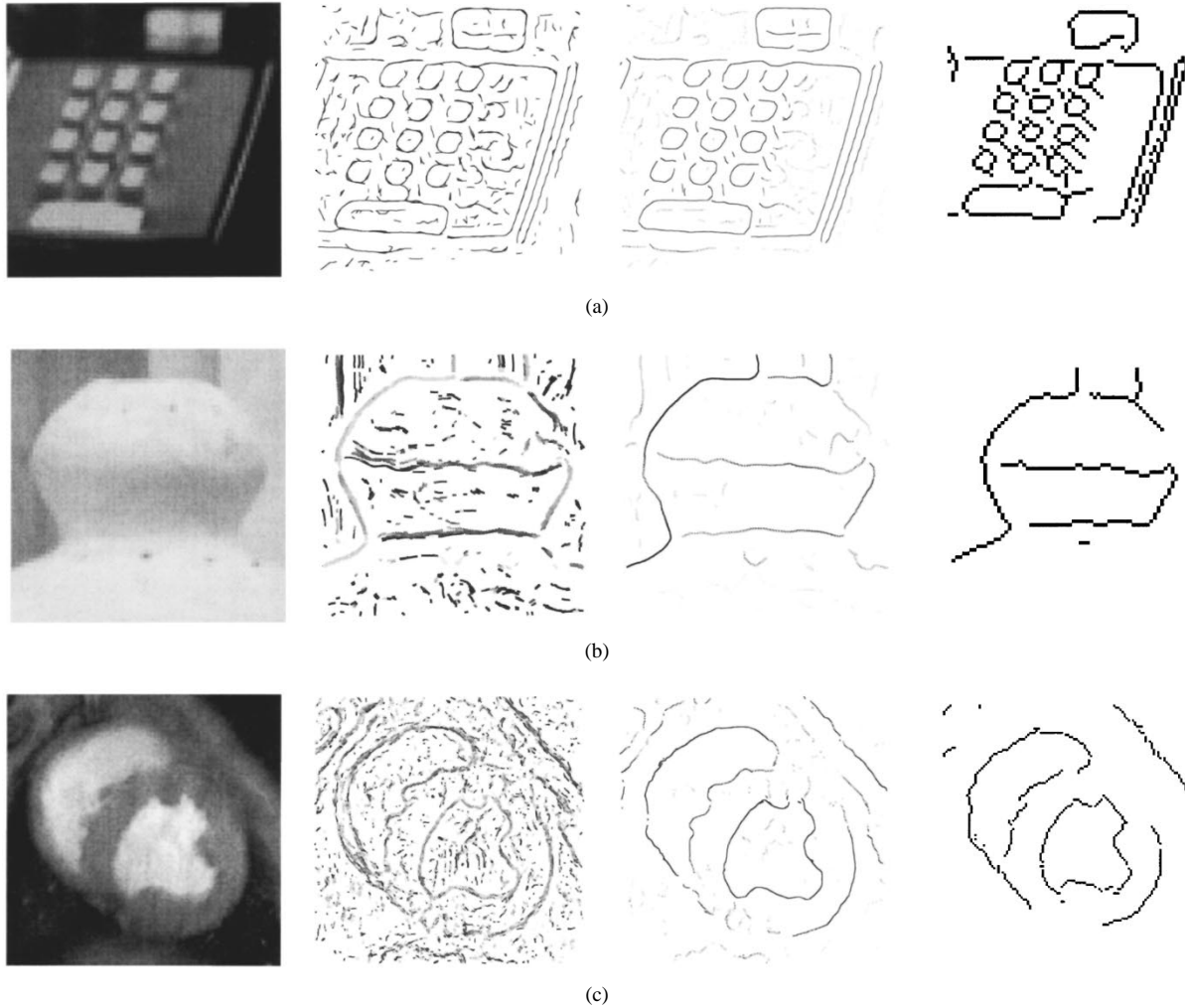
(a)

(b)

(c)

Fig. 11. Polygonal approximations of image contours. The gray value intensity in $\Pi(\Sigma(A))$ is proportional to the local brightness contrast. The blur scale parameter used in each experiment has been chosen to maximize the quality of the result (columns 3 and 4).

$$\frac{2X_1}{\cos \Theta_1} < w_i(a), \qquad i = 1, 2 \tag{62}$$

$$X_2 - X_1 > w_i(a), \qquad i = 1, 2. \tag{63}$$

Then, $\beta^+(a) \cup \beta^-(a) \subset N_{X_2}(\gamma) \setminus N_{X_1}(\gamma)$.

By using this proposition, one can prove the following result. The proof is similar to the proof of [8, Proposition 4].

Proposition 12: Let γ be a flat contour and let $a \in A|\gamma$. Let $a' \in \operatorname{div}_0(a)$ be such that

$$d(p_i(a') \to \sigma(\gamma)) \le X_1 \Rightarrow d(\theta_i(a'), \gamma.\theta) < \Theta_1. \qquad i = 1, 2. \tag{64}$$

$$\Theta_0 < \Theta_1 \tag{65}$$

$$X_0 < X_1 < \frac{\cos \Theta_1}{2} \cdot \min\{w_1(a), w_2(a), w_1(a'), w_2(a')\} \tag{66}$$

$$X_2 > X_1 + \max\{\|p_2(a') - p_1(a')\|, \\ w_1(a), w_2(a), w_1(a'), w_2(a')\} \tag{67}$$

then there exists $p' \in \{p_1(a'), p_2(a')\}$ such that

$$d(p' \to \sigma(\gamma)) \in [X_1, X_2]. \tag{68}$$

APPENDIX B
PROOF OF THEOREM 4

Definition 12: A graph for which (19) and (20) hold is said to be *separated*.

More explicitly, a graph is separated whenever the following conditions are satisfied.

(S1) If $p_1(a') \in R(a)$ or $p_2(a') \in R(a)$ then $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(a) = \emptyset$.

(S2) If $\sigma(a') \cap \tau^1(a) \neq 0$ and $\sigma(a') \cap \tau^2(a) \neq \emptyset$ then $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(a) = \emptyset$.

(S3) If $\sigma_{\mathrm{SC}}(a') \subset R_{\mathrm{SC}}(a) \cup \beta_{\mathrm{SC}}(a)$ then $\sigma(a') \cap \beta(a) = \emptyset$.

For any regular edgel-path π with arcs $(a_1, \cdots, a_n)$ let

$$\beta^-_{\mathrm{SC}}(\pi) = \bigcup_{i=1}^n \beta^-_{\mathrm{SC}}(a)$$

$$\beta^+_{\mathrm{SC}}(\pi) = \bigcup_{i=1}^n \beta^+_{\mathrm{SC}}(a)$$

and let $\beta_{\mathrm{SC}}(\pi) = \beta^-_{\mathrm{SC}}(\pi) \cup \beta^+_{\mathrm{SC}}(\pi)$. Notice that $\beta^-_{\mathrm{SC}}(\pi)$ and $\beta^+_{\mathrm{SC}}(\pi)$ are connected sets (see Fig. 12).
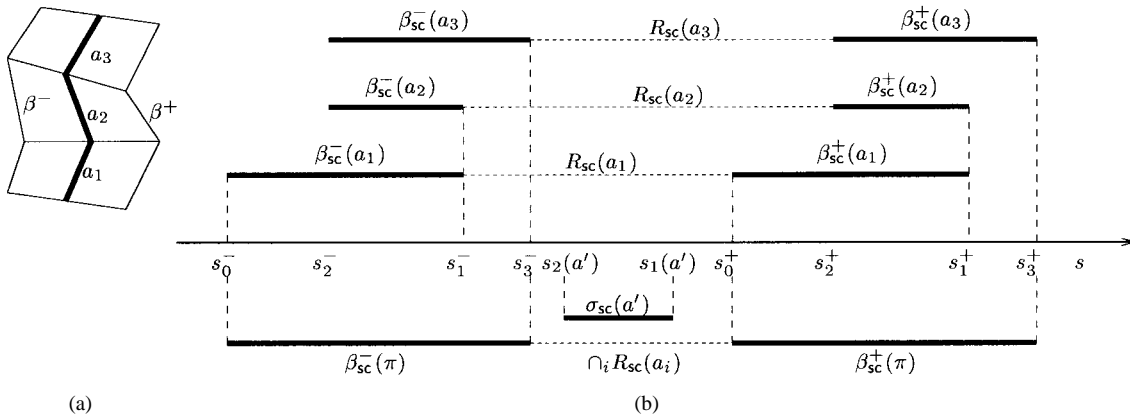
Fig. 12. An edgel-path π composed of three arcs a_1, a_2, a_3. (a) the edgel-path's entities in the image plane. (b) the edgel-path's entities on the scale axis.

Proposition 13: (Fig. 12) Let π be a regular edgel-path with arcs $(a_1, \cdots, a_n), n \geq 1$, and let $a' = (v_1', v_2')$ be an edgel-arc such that $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(\pi) = \emptyset$ and

$$\{v_1'.s, v_2'.s\} \cap \bigcup_{i=1}^{n} R_{\mathrm{SC}}(a_i) \neq \emptyset.$$

Then

$$\sigma_{\mathrm{SC}}(a') \subset \bigcap_{i=1}^{n} R_{\mathrm{SC}}(a_i). \tag{69}$$

Proof: Without loss of generality, let

$$v_1'.s \in \bigcup_{i=1}^{n} R_{\mathrm{SC}}(a_i)$$

and let k be such that

$$v_1'.s \in R_{\mathrm{SC}}(a_k), \qquad 1 \leq k \leq n.$$

Since $\beta_{\mathrm{SC}}^{-}(a_k) \leq R_{\mathrm{SC}}(a_k)$, namely, all the real numbers in $\beta_{\mathrm{SC}}^{-}(a_k)$ are less or equal to any real number in $R_{\mathrm{SC}}(a_k)$, from $v_1'.s \in R_{\mathrm{SC}}(a_k)$ it follows that $\beta_{\mathrm{SC}}^{-}(a_k) \leq v_1'.s$. Thus since $v_1'.s \in \sigma_{\mathrm{SC}}(a')$ and $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}^{-}(a_k) = \emptyset$, we have $\beta_{\mathrm{SC}}^{-}(a_k) < \sigma_{\mathrm{SC}}(a')$. Therefore, since $\beta_{\mathrm{SC}}^{-}(a_k) \subset \beta_{\mathrm{SC}}^{-}(\pi)$ and $\beta_{\mathrm{SC}}^{-}(\pi)$ is a connected set, from $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}^{-}(\pi) = \emptyset$ it follows that $\beta_{\mathrm{SC}}^{-}(\pi) < \sigma_{\mathrm{SC}}(a')$. Similarly, $\sigma_{\mathrm{SC}}(a') < \beta_{\mathrm{SC}}^{+}(\pi)$. Thus

$$\beta_{\mathrm{SC}}^{-}(a_i) < \sigma_{\mathrm{SC}}(a') < \beta_{\mathrm{SC}}^{+}(a_i), \qquad i = 1, \cdots, n,$$

that is,

$$\sigma_{\mathrm{SC}}(a') \in R_{\mathrm{SC}}(a_i), \qquad i = 1, \cdots, n. \qquad \square$$

The following notation will be used when dealing with properties and assumptions holding for sets of integers. The set of integers i such that $i \geq k$ and $i \leq k$ is denoted $\{k, \cdots, l\}$. If $l < k$ then this set is empty and therefore a property which holds true $\forall i \in \{k, \cdots, l\}$ is always true if $l < k$. The notation $i = k, \cdots, l$ is equivalent to $k \leq i \leq l$ and therefore requires that $k \leq l$.
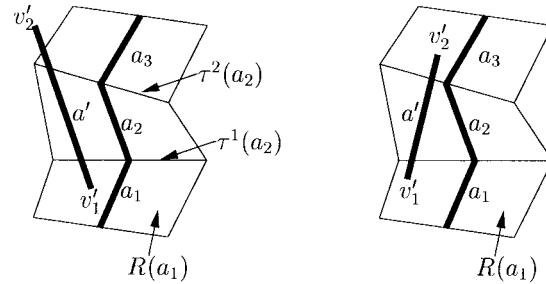

Fig. 13. Proposition 14.

Proposition 14: (Fig. 13) Let π be a regular path with arcs $(a_1, \cdots, a_n), n \geq 1$, in a separated edgel-graph A and let $a' = (v_1', v_2') \in A$ be an arc such that $v_1'.p \in R(a_1), v_1'.s \in R_{\mathrm{SC}}(a_1)$ and [Fig. 13(a)]

$$\sigma(a') \cap \tau^h(a_i) \neq \emptyset, \quad h = 1, 2; \quad \forall i \in \{2, \cdots, n-1\}. \tag{70}$$

Then

$$\sigma_{\mathrm{SC}}(a') \subset \bigcap_{i=1}^{n-1} R_{\mathrm{SC}}(a_i). \tag{71}$$

Furthermore, if $v_2'.p \in R(a_n)$ [Fig. 13(b)],

$$\sigma_{\mathrm{SC}}(a') \subset \bigcap_{i=1}^{n} R_{\mathrm{SC}}(a_i). \tag{72}$$

Proof: From the separation condition (S1) and $v_1'.p \in R(a_1)$ we have

$$\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(a_1) = \emptyset. \tag{73}$$

If $n \geq 3$, from (70) and the separation condition (S2) we have

$$\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(a_i) = \emptyset, \qquad i = 2, \cdots, n-1 \tag{74}$$

which, together with (73), yields

$$\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(\langle a_1, \cdots, a_{n-1} \rangle) = \emptyset, \qquad n \geq 2 \tag{75}$$

where $\langle a_1, \cdots, a_{n-1} \rangle$ denotes the edgel-path with arcs $a_1, \cdots, a_{n-1}$. Then, (71) follows from (75) and Proposition 13.

To prove the second part, let $v_2'.p \in R(a_n)$. Then, from the separation condition (S1) we get $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(a_n) = \emptyset$
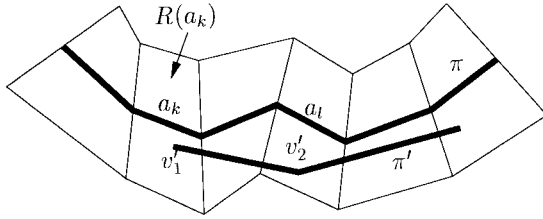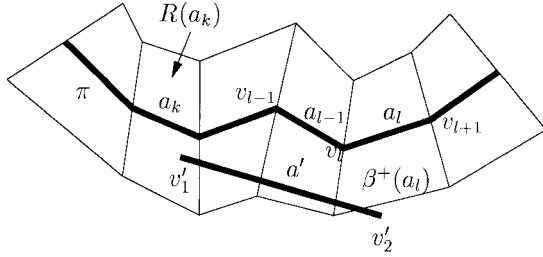
Fig. 14.   Proposition 15.

Fig. 15.   Proposition 16. The shown geometric configuration, namely $\sigma(a') \cap \beta(\pi) \neq \emptyset$, is prohibited by the proposition.

and therefore, by combining this with (75), it follows that $\sigma_{\mathrm{SC}}(a') \cap \beta_{\mathrm{SC}}(\pi) = \emptyset$ which, by Proposition 13 implies (72). $\square$

For any regular edgel-path π and any $p \in R(\pi)$ let $R_{\mathrm{SC}}^{\pi}(p) = R_{\mathrm{SC}}(a)$, where a is the unique arc in π such that $p \in R(a)$.

Proposition 15: (Fig. 14) Let π, π' be regular paths in a separated edgel-graph A such that $\sigma(\pi') \subset R(\pi)$ and $v_1'.s \in R_{\mathrm{SC}}^{\pi}(v_1'.p)$ where v_1' is the first vertex of π'. Then

$$v'.s \in R_{\mathrm{SC}}^{\pi}(v'.p)$$

for every vertex v' of π'.

Proof: First, let us assume that π' consists of one arc $a' = (v_1', v_2')$. Let $a_1, \cdots, a_n$ be the arcs of π. Let k, l be such that $v_1'.p \in R(a_k)$ and $v_2'.p \in R(a_l)$. Without loss of generality, let us assume that $l \geq k$ (otherwise, interchange v_1' with v_2' in the argument). Notice that, since $\sigma(\pi') \subset R(\pi)$

$$\sigma(a') \cap \tau^h(a_i) \neq \emptyset, \quad h = 1, 2; \quad \forall i \in \{k+1, \cdots, l-1\}.$$

Thus from the second part of Proposition 14 applied to the edgel-path $\langle a_k, \cdots, a_l \rangle$
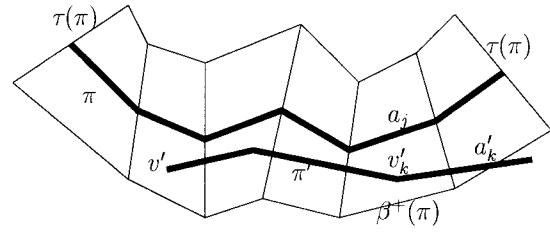
$$\sigma_{\mathrm{SC}}(a') \subset R_{\mathrm{SC}}(a_l)$$

from which

$$v_2'.s \subset R_{\mathrm{SC}}(a_l) = R_{\mathrm{SC}}^{\pi}(v_2'.p).$$

For paths with more than arc the result follows by recursion. $\square$

Proposition 16: (Fig. 15) Let π be a regular path with arcs $(a_1, \cdots, a_n), a_i = (v_i, v_{i+1})$ in a separated edgel-graph A. Let $a' = (v_1', v_2') \in A$. If $v_1'.p \in R(\pi)$ and $v_1'.s \in R_{\mathrm{SC}}^{\pi}(v_1'.p)$ then $\sigma(a') \cap \beta(\pi) = \emptyset$.

Fig. 16.   Proposition 17. The situation shown, namely $\sigma(\pi') \not\subset R(\pi)$, is prohibited by the proposition.

Proof: Let k be such that $v_1'.p \in R(a_k)$. For the purpose of contradiction, let $\sigma(a') \cap \beta(\pi) \neq \emptyset$ and let a_l be the first exit arc for $\sigma(a')$. That is, $\sigma(a') \cap \beta(a_l) \neq \emptyset$ and

$$\sigma(a') \cap \beta(a_i) = \emptyset, \quad \forall i \in \{k, \cdots, l-1\} \tag{76}$$

where we have assumed for simplicity that $k \leq l$. From the separation condition (S3), it is sufficient to prove that

$$\sigma_{\mathrm{SC}}(a') \subset R_{\mathrm{SC}}(a_l) \cup \beta_{\mathrm{SC}}(a_l). \tag{77}$$

Note that from $v_1'.s \in R_{\mathrm{SC}}^{\pi}(v_1'.p) = R_{\mathrm{SC}}(a_k)$ and the separation condition (S1) it follows that

$$\sigma_{\mathrm{SC}}(a') \subset R_{\mathrm{SC}}(a_k). \tag{78}$$

Thus (77) holds if $k = l$. If $l \geq k+1$, we will first prove that

$$\sigma_{\mathrm{SC}}(a') \subset R_{\mathrm{SC}}(a_{l-1}). \tag{79}$$

If $l = k+1$ then (79) follows immediately from (78). If $l \geq k+2$ then from (76) we have

$$\sigma(a') \cap \tau^h(a_i) \neq \emptyset, \quad h = 1, 2; \quad i = k+1, \cdots, l-1$$

which, together with the first part of Proposition 14, proves (79) for $l \geq k+2$. Then, since $R_{\mathrm{SC}}(a_{l-1}) = \tau_{\mathrm{SC}}(v_{l-1}) \cap \tau_{\mathrm{SC}}(v_l)$, from (79) we have

$$\sigma_{\mathrm{SC}}(a') \subset R_{\mathrm{SC}}(a_{l-1}) \subset \tau_{\mathrm{SC}}(v_l) \subset (\tau_{\mathrm{SC}}(v_l) \cup \tau_{\mathrm{SC}}(v_{l+1}))$$
$$= R_{\mathrm{SC}}(a_l) \cup \beta_{\mathrm{SC}}(a_l). \quad \square$$

For any edgel-path π, let $\tau(\pi) \equiv \tau(v_{\mathrm{fi}}) \cup \tau(v_{\mathrm{la}})$ where v_{fi} and v_{la} and the first and last vertex of π. Let $R^{\circ}(\pi)$ denote the interior of $R(\pi)$ and let $\sigma^{\circ}(\pi)$ denote the interior of the polygonal line $\sigma(\pi)$, namely $\sigma(\pi)$ less its two end-points.

Proposition 17: (Fig. 16) Let π, π' be paths in a separated edgel-graph A such that $v'.p \in R(\pi)$ and $v'.s \in R_{\mathrm{SC}}^{\pi}(v'.p)$ where v' is a vertex of π'. If $\sigma^{\circ}(\pi') \cap \tau(\pi) = \emptyset$ and $\sigma(\pi') \cap R^{\circ}(\pi) \neq \emptyset$, then

$$\sigma(\pi') \subset R(\pi)$$

and $v_i'.s \in R_{\mathrm{SC}}^{\pi}(v_i'.p)$ for all the vertices v_i' of π'.
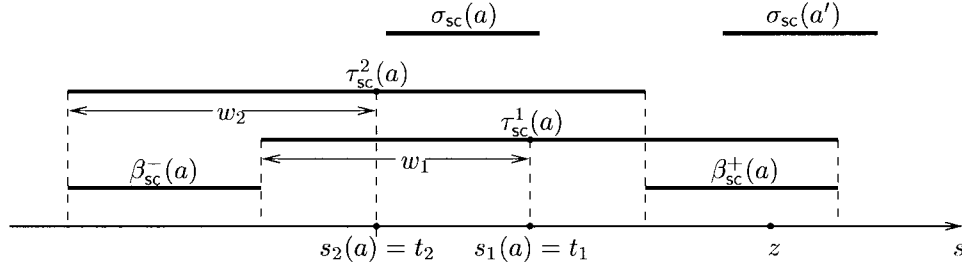
Fig. 17. Some notation for the proof of Lemma 18. z denotes a point in $\sigma_{\text{SC}}(a') \cap \beta_{\text{SC}}(a)$.

Proof: For the purpose of contradiction let us assume that $\sigma(\pi') \not\subset R(\pi)$. Since $R(\pi)$ is a closed set, then $\sigma^\circ(\pi')$ must contain a point outside $R(\pi)$. Then, since $\sigma^\circ(\pi') \cap R^\circ(\pi) \neq \emptyset$, it follows that $\sigma^\circ(\pi')$ must intersect the boundary of $R(\pi)$, which is given by $\tau(\pi) \cup \beta(\pi)$. Since $\sigma^\circ(\pi') \cap \tau(\pi) = \emptyset$ we must have $\sigma^\circ(\pi') \cap \beta(\pi) \neq \emptyset$ and therefore $\sigma(\pi') \cap \beta(\pi) \neq \emptyset$. For simplicity let us assume that v' is the first vertex of π', and that $\sigma(a'_1) \cap R^\circ(\pi) \neq \emptyset$. Let a'_k be the first arc in the path π' such that $\sigma(a'_k) \cap \beta(\pi) \neq \emptyset$. First, let us prove that

$$v'_k.s \in R^\pi_{\text{SC}}(v'_k.p). \tag{80}$$

If $k = 1$ then this follows immediately from $v'_1.s \in R^\pi_{\text{SC}}(v'_1.p)$. If $k > 1$ we have $\sigma(\pi'') \subset R(\pi)$ where π'' is the subpath with arcs $a'_1, \cdots, a'_{k-1}$ and vertices $v'_1, \cdots, v'_k$. Then (80) follows from Proposition 15. From (80) we have $v'_k.s \in R_{\text{SC}}(a_j)$ where j is such that $v'_k.p \in R(a_j)$. Then, by substituting $a'_k = (v'_k, v'_{k+1})$ for (v'_1, v'_2) in Proposition 16 one gets $\sigma(a'_k) \cap \beta(\pi) = \emptyset$ which is a contradiction.

Proof of Theorem 4: Let v_{fi} and v_{la} be the first and last vertex of π_1 and π_2. Clearly, $v_{\text{fi}}.p \in R(\pi_1)$ and $v_{\text{fi}}.s \in R^{\pi_1}_{\text{SC}}(v_{\text{fi}}.p)$. Since π_1 and π_2 are regular paths, it follows that $\sigma^\circ(\pi_2) \cap \tau(\pi_1) = \emptyset$, where $\tau(\pi_1)$ denotes $\tau(v_{\text{fi}}) \cup \tau(v_{\text{la}})$. From (5) we have that $\sigma(a'_1) \cap R^\circ(\pi_1) \neq \emptyset$ where a'_1 is the first arc of π_2. Then, from Proposition 17

$$\sigma(\pi_2) \subset R(\pi_1).$$

Similarly, by interchanging π_1 with π_2 in the above argument,

$$\sigma(\pi_1) \subset R(\pi_2).$$

The result then follows from Proposition 1. □

APPENDIX C

LEMMA FOR SECTION VI-E

For any $a = (v_1, v_2) \in A$ let

$$w^{\min}_{\text{SC}}(a) = \min\{w_{\text{SC}}(v_1), w_{\text{SC}}(v_2)\}, \tag{81}$$
$$w^{\max}_{\text{SC}}(a) = \max\{w_{\text{SC}}(v_1), w_{\text{SC}}(v_2)\}. \tag{82}$$

Lemma 18: (Fig. 17) Let $a, a' \in A$ be such that $a' \nparallel_{\text{SC}} a$ and

$$|s_1(a) - s_2(a)| < S_\Delta \tag{83}$$
$$|s_1(a') - s_2(a')| < S_\Delta. \tag{84}$$

Then,

$$\forall j \in \{1, 2\}, \exists i \in \{1, 2\}$$
$$|s_j(a') - s_i(a)| \in [w^{\min}_{\text{SC}}(a) - S_\Delta, w^{\max}_{\text{SC}}(a) + S_\Delta]$$
$$\tag{85}$$

$$\forall i \in \{1, 2\}, \exists j \in \{1, 2\},$$
$$|s_j(a') - s_i(a)| \in [w^{\min}_{\text{SC}}(a) - S_\Delta, w^{\max}_{\text{SC}}(a) + S_\Delta].$$
$$\tag{86}$$

Proof: From $a' \nparallel_{\text{SC}} a$ it follows that

$$\sigma_{\text{SC}}(a') \cap \beta_{\text{SC}}(a) \neq \emptyset$$

that is

$$\sigma_{\text{SC}}(a') \cap ((\tau^1_{\text{SC}}(a) \cup \tau^2_{\text{SC}}(a)) \setminus (\tau^1_{\text{SC}}(a) \cap \tau^2_{\text{SC}}(a))) \neq \emptyset.$$

Therefore, there exists $z \in \sigma_{\text{SC}}(a')$ such that z belongs to one of the two intervals $\tau^i_{\text{SC}}(a)$ but not to the other. Without loss of generality, let then $z \in \tau^1_{\text{SC}}(a), z \notin \tau^2_{\text{SC}}(a)$ (see Fig. 17), that is, by letting $t_1 = s_1(a), t_2 = s_2(a)$

$$|z - t_1| \leq w^1_{\text{SC}}(a), \qquad |z - t_2| \geq w^2_{\text{SC}}(a) \tag{87}$$

from which,

$$|z - t_1| \leq w^{\max}_{\text{SC}}(a), \qquad |z - t_2| \geq w^{\min}_{\text{SC}}(a). \tag{88}$$

Since $|t_1 - t_2| < S_\Delta$ by assumption, (88) yields, by letting $w_1 = w^{\min}_{\text{SC}}(a), w^2 = w^{\max}_{\text{SC}}(a)$

$$w_1 - S_\Delta < |z - t_1| \leq w_2$$
$$w_1 \leq |z - t_2| < w_2 + S_\Delta. \tag{89}$$

Let us introduce the two functions $\delta_i(s') = |s' - t_i|, i = 1, 2$. We have

$$w_1 - S_\Delta < \delta_1(z) \leq w_2 \tag{90}$$
$$w_1 \leq \delta_2(z) < w_2 + S_\Delta. \tag{91}$$

Notice that

$$|\delta_i(x) - \delta_i(x')| < |x - x'|, \quad \forall x, x' \in \mathbb{R}, \quad i = 1, 2. \tag{92}$$

Also, since $|t_1 - t_2| < S_\Delta$

$$|\delta_1(x) - \delta_2(x)| < S_\Delta, \quad \forall x \in \mathbb{R}. \tag{93}$$

To prove (85) let us fix $t'_k \in \{t'_1, t'_2\} = \{s_1(a'), s_2(a')\}$. Since $|t'_1 - t'_2| < S_\Delta$ and $z \in \sigma_{\text{SC}}(a') = [t'_1, t'_2]$, we have $|t'_k - z| < S_\Delta$ and therefore, from (92)

$$|\delta_1(t'_k) - \delta_1(z)| < S_\Delta, \tag{94}$$
$$|\delta_2(t'_k) - \delta_2(z)| < S_\Delta. \tag{95}$$

These, together with (90) and (91) yield

$$\delta_1(t'_k) < w_2 + S_\Delta \tag{96}$$
$$w_1 - S_\Delta < \delta_2(t'_k). \tag{97}$$

These two inequalities and $|\delta_1(t'_k) - \delta_2(t'_k)| < S_\Delta$ [which comes from (93)], imply that $\delta_1(t'_k)$ and $\delta_2(t'_k)$ can not be both outside the interval $[w_1 - S_\Delta, w_2 + S_\Delta]$, which proves (85).
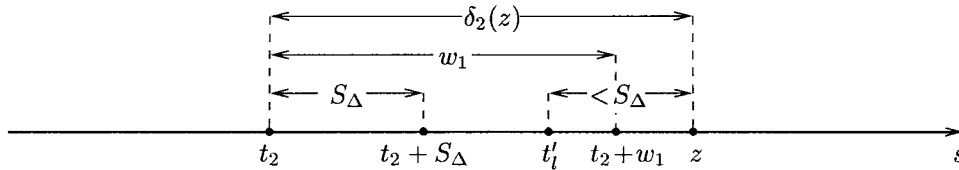
Fig. 18.   Lemma 18: existence of t'_l such that $\delta_2(t'_l) \leq \delta_2(z)$.

To prove the second part, let us fix $t_k \in \{t_1, t_2\} = \{s_1(a), s_2(a)\}$. If $k = 1$, then let t'_l be the point in $\{t'_1, t'_2\}$ where the function δ_1 is *greater* or equal to $\delta_1(z)$ (one of the two points has this property because $z \in [t'_1, t'_2]$; notice that t'_l is the point in $\{t'_1, t'_2\}$ furthest away from t_1). Thus $\delta_1(z) \leq \delta_1(t'_l) < \delta_1(z) + S_\Delta$ (the second inequality follows from (92) and $|z - t'_l| \leq |t_1 - t_2| < S_\Delta$). Then, from (90) we have that $\delta_1(t'_l)$ is in the interval $[w_1 - S_\Delta, w_2 + S_\Delta]$.

Let now $k = 2$. From (91) and (3) we have $\delta_2(z) = |z - t_2| \geq w_1 > S_\Delta$. Thus since $z \in [t'_1, t'_2]$ and $|t'_1 - t'_2| < S_\Delta$, by moving on the real axis from z toward t_2 so that δ_2 decreases, either t'_1 or t'_2 is reached before t_2. Let t'_l be this point. Clearly, $\delta_2(t'_l) \leq \delta_2(z)$. Hence, from $z \in [t'_1, t'_2]$ and $|t'_1 - t'_2| < S_\Delta$, it follows that $\delta_2(z) - S_\Delta < \delta_2(t'_l) \leq \delta_2(z)$. Thus from (91), $\delta_2(t'_l) \in [w_1 - S_\Delta, w_2 + S_\Delta]$.   □

Corollary 19: If (83) and (84) hold and $a' \nparallel_{\text{SC}} a \vee a \nparallel_{\text{SC}} a'$ then

$$\forall i \in \{1, 2\}, \exists j \in \{1, 2\}$$

$$|s_j(a') - s_i(a)| \in [w^{\min}_{\text{SC}}(a, a') - S_\Delta, w^{\max}_{\text{SC}}(a, a') + S_\Delta].$$

REFERENCES

[1] M. Basseville, M. Benveniste, A. Chou, K. C. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory*, vol. 38, pp. 529–532, Mar. 1992.

[2] F. Bergholm, "Edge focusing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, pp 726–741, Nov. 1987.

[3] M. Bertero, T. Poggio, and V. Torre, "Ill-posed problems in early vision," *Proc. IEEE*, vol. 76, no. 8, pp. 869–889, Aug. 1988.

[4] A. Blake and A. Zisserman, *Visual Reconstruction.*   Cambridge, MA: MIT Press, 1987.

[5] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 679–698, 1986.

[6] S. Casadei and S. K. Mitter, "A hierarchical approach to high resolution edge contour reconstruction," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, 1996, pp. 149–153.

[7] _____, "Hierarchical curve reconstruction. Part 1: Bifurcation analysis and recovery of smooth curves," in *Proc. European Conf. Comput. Vision*, 1996, pp. 199–208.

[8] _____, "Hierarchical image segmentation—Part I: Detection of regular curves in a vector graph," *Int. J. Comput. Vision*, vol. 27, no. 3, pp. 71–100, Mar. 1998.

[9] _____, "A perceptual organization approach to contour estimation via composition, compression and pruning of contour hypotheses," Lab. Inform. Decision Syst., Mass. Inst. Technol., Tech. Rep. LIDS-P-2415, Apr. 1998.

[10] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *Int. Conf. Comput. Vision*, 1995, pp. 694–699.

[11] I. J. Cox, J. M. Rehg, and S. Hingorani, "A Bayesian multiple-hypothesis approach to edge grouping and contour segmentation," *Int. J. Comput. Vision*, vol. 11, no. 1, pp. 5–24, Aug. 1993.

[12] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector," *Int. J. Comput. Vision*, vol. 1, pp. 167–187, 1987.

[13] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and blur estimation," in *Proc. European Conf. Comput. Vision*, vol. II, pp. 57–69, 1996.

[14] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, Nov. 1984.

[15] U. Grenander and M. I. Miller, "Representations of knowledge in complex systems," *J. Roy. Statist. Soc. B*, vol. 56, no. 4, pp. 549–603, 1994.

[16] R. M. Haralick, "Second directional derivative zero crossing detector using the cubic facet model," in *IEEE Comput. Vision Pattern Recognition or CVPR*, 1985, pp. 672–677.

[17] M. D. Heath, S. Sarkar, T. Sanocki, and K. W. Bowyer, "A robust visual method for assessing the relative performance of edge detection algorithms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 1338–1359, Dec. 1997.

[18] T. Kanungo, M. Y. Jaisimha, J. Palmer, and R. M. Haralick, "A quantitative methodology for analyzing the performance of detection algorithms," in *Proc. Int. Conf. Comput. Vision*, 1993, pp. 247–252.

[19] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vision*, vol. 1, pp. 321–331, 1988.

[20] S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models," in *Proc. Int. Conf. Comput. Vision*, 1995, pp. 810–815.

[21] H. Krim and J.-C. Pesquet, "Multiresolution analysis of a class of nonstationary processes," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1010–1020, July 1995.

[22] Y. C. Leclerc, "Constructing simple stable descriptions for image partitioning," *Int. J. Comput. Vision*, vol. 3, pp. 73–102, 1989.

[23] T. Lindeberg, "Edge detection with automatic scale selection," in *Proc. IEEE Comput. Vision Pattern Recognition*, 1996, pp. 465–470.

[24] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, pp. 617–643, Mar. 1992.

[25] D. Marr, *Vision.*   San Francisco, CA: Freeman, 1982.

[26] J. Marroquin, S. K. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Amer. Statist. Assoc.*, vol. 82, no. 397, pp. 76–89, Mar. 1987.

[27] D. Mumford and J. Shah, "Optimal approximations of piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[28] P. Parent and S. W. Zucker, "Trace inference, curvature consistency, and curve detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, Aug. 1989.

[29] P. Perona and J. Malik, "Detecting and localizing edges composed of steps, peaks and roofs," in *Int. Conf. Comput. Vision*, Osaka, Japan, 1990, pp. 52–57.

[30] T. J. Richardson and S. K. Mitter, "Approximation, computation, and distorsion in the variational formulation," in B.M. ter Haar Romeny, Ed., *Geometry-Driven Diffusion in Computer Vision.*   Amsterdam, The Netherlands: Kluwer, 1994, pp. 169–190.

[31] K. Rohr, "On the precision in estimating the location of edges and corners," *J. Math. Imaging Vision*, vol. 7, no. 1, pp. 7–22, Jan. 1997.

[32] P. W. Verbeek and L. J. Van Vliet, "On the location error of curved edges in low-pass filtered 2-d and 3-d images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 726–733, July 1994.

[33] S. J. Wang and T. O. Binford, "Detection, estimation and aggregation of three major types of discontinuities in image surfaces," in *Proc. ARPA96*, 1996, pp. 923–926.

[34] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 884–900, Sept. 1996.

[35] S. W. Zucker, C. David, A. Dobbins, and L. Iverson, "The organization of curve detection: Coarse tangent fields and fine spline coverings," in *Proc. Int. Conf. Comput. Vision*, 1988.