Journal of Mathematical Imaging and Vision 22: 159–181, 2005

Extracting Meaningful Curves from Images

FRÉDÉRIC CAO
IRISA/INRIA, Campus Universitaire de Beaulieu, 35042 Rennes, Cedex, France
Frederic.Cao@irisa.fr

PABLO MUSÉ AND FRÉDÉRIC SUR
École Normale Supérieure de Cachan, 61 avenue du Président Wilson, 94235 Cachan, Cedex, France
muse@cmla.ens-cachan.fr
sur@cmla.ens-cachan.fr

Abstract. Since the beginning, Mathematical Morphology has proposed to extract shapes from images as connected components of level sets. These methods have proved very efficient in shape recognition and shape analysis. In this paper, we present an improved method to select the most meaningful level lines (boundaries of level sets) from an image. This extraction can be based on statistical arguments, leading to a parameter free algorithm. It permits to roughly extract all pieces of level lines of an image, that coincide with pieces of edges. By this method, the number of encoded level lines is reduced by a factor 100, without any loss of shape contents. In contrast to edge detection algorithms or snakes methods, such a level lines selection method delivers accurate shape elements, without user parameter since selection parameters can be computed by the Helmholtz Principle. The paper aims at improving the original method proposed in [10]. We give a mathematical interpretation of the model, which explains why some pieces of curve are overdetected. We introduce a multiscale approach that makes the method more robust to noise. A more local algorithm is introduced, taking local contrast variations into account. Finally, we empirically prove that regularity makes detection more robust but does not qualitatively change the results.

Keywords: topographic maps, level lines, edge detection, Helmholtz principle, shapes elements

1. Introduction

Natural images are very complex, and despite the progress of modern computers, we cannot handle the huge amount of information they contain. Thus, the idea of Marr and Hildreth [25] that edges provide a good summary of images is still vivid. Since their seminal works, efforts have been carried on local methods. Marr defined edges as zero-crossings of the Laplacian [24], and Haralick [15] proposed a more correct definition which is equivalent to the zero-crossings of $D^2u(Du, Du)$ where Du and D^2u are respectively the gradient and the second derivative of the image. In his famous paper [3], Canny gives a filter that tries to optimize the edge localization, but which is equivalent to Haralick's. Although they are technically sound, local methods also have an immediate drawback: while edges are usually thought about as curves, these methods detect sets of points with an orientation (*edgels*) that have to be connected afterward. Moreover, they require different thresholds since contrast has no absolute meaning. In addition, they are sensitive to noise (since they use derivatives of the image) and should be considered through a multiscale process, by first computing edges at large scale, then tracking them back to small scales. The choice of these thresholds depends on the observed image, and is not that easy. It is also known that edge is not a completely local concept and that it does not rely entirely on contrast. Indeed, following Gestalt Theory [16, 41], shapes (and thus edges) result from the collaboration of a small set of perceptual laws (called "partial gestalts" by Desolneux, Moisan and

Morel [11]), and contrast is only one of them. Among others, we can cite alignments, symmetry, convexity, closedness, good continuation, etc.

Other theories, related to edge detection, explicitly use good continuation, which means in this case regularity of curves. The most famous is certainly the theory of active contours (or snakes) [17], where optimal boundaries result from a compromise between their intrinsic regularity and the extrinsic value of the image contrast along the active contours. The main weaknesses of this theory are the number of parameters and the sensitivity to an initial guess. More recent methods propose to initiate the detection with many contours, most of which will hopefully disappear [8]. But again, there is no measure on the certainty of the remaining detected contours.

The Mathematical Morphology school proposed an alternative to the local approaches above. Following morphologists, the image information is completely contained in a family of binary images that are obtained by thresholding the images at given values [26, 38]. This is equivalent to considering level sets; the (upper) level set of u at the value λ is

$$\chi_\lambda(u) = \{x \in \mathbb{R}^2, u(x) \geq \lambda\}. \tag{1}$$

Obviously, if we only consider a coarsely quantized set of different gray levels, information is lost, especially in textures. Nevertheless, it is worth noting how large shapes are already present with as few as 5 or 6 levels. As soon remarked by Serra [38], no information is lost at all, since we can reconstruct an image from the whole family of its level sets, by

$$u(x) = \sup\{\lambda \in \mathbb{R}, x \in \chi_\lambda(u)\}.$$

Thus, the level sets do not only give a convenient way to extract information, they provide a complete representation of images. Alternative complete representations are, for instance, Fourier or wavelets decomposition [23]. But while these last ones are very adequate for image compression (they are used in the JPEG 2000 standard), they are not very well adapted to shape analysis, since their basic elements have no immediate perceptual interpretation. (More recent decompositions as bandlets [35] or curvelets [40] try to take image geometry more into account, but they are either still too local, or need a preliminary detection step.) On the contrary, morphologists soon remarked that boundaries of level sets fit parts of objects boundaries very well. They call level lines the topological boundaries of connected components of level sets, and topographic map of an image, the collection of all its level lines. The topographic map also gives a complete representation of an image and enjoys several important advantages [6]:

- It is invariant with respect to contrast change. It is not invariant to illumination change, since in this case, the image is really different, although it represents the same scene. However, many level lines still are locally the same.
- It is not as local as sets of edges, since level lines are Jordan curves that are either closed or meet the image borders. (This property requires that the image has bounded variations [13]).
- It is a hierarchical representation: since level sets are ordered by the inclusion relation (and so are there connected components), the topographic map may be embedded in a tree structure.
- But most important regarding the main subject of this paper, object contours locally coincide with level lines very well. Basically, level lines are everywhere normal to the gradient as edges. On the other hand, level lines are accurate at occlusions. Whereas, edges detectors usually fail near T-junctions (and additional treatments are necessary), there are several level lines at a junction (See Fig. 1). The order of the multiple junction coincides with the number of level lines [5]. We shall go back to this in Section 3.2.

Figure 1. Level lines and T-junction. Depending on the grey level configuration between shapes and background, level lines may follow or not (as on the figure) the objects boundary. In any case, junctions appear where two level lines separate. Here, there are two kinds of level lines: the occluded circle and the shape composed of the union of the circle and the square. The square itself may be retrieved by difference.

The level sets representation has recently been used, with success, for image simplification and segmentation. In particular, it was shown that it allowed to define multiscale representation of images [27, 36, 37], while avoiding the main drawbacks of linear scale space theory [19, 42], namely an oversmoothing of contours.

We are convinced that level lines may directly give usable curves for any shape recognition algorithm. The main drawback of the topographic map representation is its lack of compactness. First, since it is complete, it contains all the texture information. The level lines in textures are usually very complicated, and are not always useful for shape recognition. (The opposite may be true, for instance for very accurate image registration). Moreover, because of noise and interpolation, many level lines may follow roughly one and the same contour. Thus, it is useful, for practical computational reasons, to select only the most meaningful level lines. It is also worth noting that level sets are nested, and thus can be embedded in a tree structure. Actually, this representation is not fully satisfying for two reasons. First, we can also define lower level sets by $\chi^{\lambda}(u) = \{x,\ u(x) \leq \lambda\}$. Therefore, we may construct two trees, which are usually different. Moreover, both lower and upper level sets may not be simply connected, and do not suitably represent the structure of image shapes under the occlusion phenomenon. Monasse and Guichard [29] proved that it is possible to merge the trees of upper and lower level sets into a single structure (the tree of level lines). Each node of the tree is a level line or equivalently a connected component of level set (either lower or upper) whose holes have been filled. They call a node a *Shape* since it roughly corresponds to a disoccluded shape.

Recently, Desolneux et al. proposed a parameterless algorithm using Monasse's tree to detect contrasted level lines (called *meaningful boundaries*) in grey level images [10]. Their method, which needs no parameter tuning, relies on a perceptual principle called Helmholtz Principle. Experimentally, meaningful boundaries are often very close to minimizers of any reasonable snake energy [12]. This adequation of meaningful boundaries and snakes is a bit paradoxical since, unlike snakes, no local regularity is imposed on meaningful boundaries.

However, the algorithm of Desolneux et al. raises several questions and objections. The definition of meaningful boundaries has first to be precisely interpreted in mathematical terms. Second, because of noise (and certainly partly because of quantization noise), some edges are missing (lots of them in some low contrasted images). Third, it uses a global contrast information (the histogram). This yields too many detections in regions with important contrast and too few detection ins low contrasted regions (it is the so-called blue sky effect). Finally, regularity of edges is not used for the detection.

In this paper, we discuss and answer these objections, with a significant improvement. Our conclusions are the following: the definition of meaningful boundaries themselves does not ensure that they do not contain any undesirable parts. We propose a method to remove those parts. Second, the method can be extended to several scales, and this makes the method more robust to noise. We also propose a method considering contrast in a more local way. If we use more local contrast information, we can remove edges in texture. Whether this is a nice thing or not depends on the application: for very accurate registration, texture-edges can be useful, while they must be useless for shape recognition. (For texture recognition, harmonic analysis methods are certainly more efficient.) Last, we introduce a local and stable measure of regularity of a curve and use it for smooth edges detection. As already noticed in [4], regularity is often sufficient to detect some very meaningful edges. Nevertheless, general belief is that both regularity and contrast are useful for edge detection. We experimentally check that contrast and regularity are often very redundant. This redundancy is used to make the detection even more robust, but does not change the results of contrast based detection alone. We are also able to automatically tune the relative weight of regularity and contrast, which is a recurrent question in active contours theory.

The plan is as follows. In Section 2, we recall the bases of Helmholtz Principle, the definition of meaningful boundaries of Desolneux, Moisan and Morel. We will justify and discuss this definition, which was not explicitly made in [10]. A multiscale extension is presented in Section 3. In Section 4, we describe a procedure that automatically handles local contrast variations. In Section 5, we explain how both contrast and regularity criteria can naturally be mixed in a probabilistic setting by introducing a measure of regularity on random level lines in Section 5.1. We conclude in Section 6.

2. Meaningful Boundaries

2.1. Helmholtz Principle

Helmholtz Principle is a perceptual principle asserting that conspicuous structures may be viewed as exceptions to randomness. The unexpected configurations we must be interested in, are given by the perceptual laws of Gestalt Theory [16, 41], as alignments, closedness of sets, parallelism etc. Since this principle is quite general, its formulation may slightly vary from an application to another but, we propose the following formulation. Assume that $O_1, \ldots, O_N$ are local objects in an image (for instance, $O_1, \ldots, O_N$ may be edgels, that is points assigned with a direction). We want to find out whether some of these objects must be grouped in a more global structure, with respect to some shared quality. Let us assume that we have K group candidates $G_1, \ldots G_K$. Each of the G_i gathers several of the local objects O_n, given in advance. We now consider a quality Q measured from the O_n, and given by the Gestalt laws (for instance alignment). Each measure defines a random variable X_n. We wish to determine if the G_i are meaningful groups for the quality Q. We then make the following mental experiment: assume that, anything else held equal, the quality Q is independently and identically distributed over the O_n, that is to say the X_n are i.i.d. variables. We call this hypothesis the a contrario model. If we have no a priori information on the X_n, their distribution in the a contrario model can be, for instance, their distribution in a white noise image. Assume that for some group G_i, the X_n are equal up to some precision. By definition, we will say that G_i is ε-meaningful, if in the a contrario model, the probability that all X_n in G_i are equal up to the observed precision is less than $\frac{\varepsilon}{K}$. As will be seen on a more precise example in the next sections, this definition implies that, in the a contrario model, the expected number of ε-meaningful groups is less than ε. In other words, the number of groups appearing by chance is controlled, on the average, by ε.

We refer the reader to [11] and references therein, for precise applications of this principle. The following sections are devoted to the application of this principle to the extraction of shapes elements.

Before going further, let us give a few comments on the above principle. A very important point is that it is discrete by nature. Indeed we consider a finite number of local objects, and we also consider a finite a priori number of group candidates. Moreover, the quality Q is measured with a finite accuracy. Put together, this implies that, under the a contrario model, any group has a positive probability of occurrence. This probability is decreasing with the size of the group (the number of local events it contains). Moreover, for all $\varepsilon > 0$ the number of ε-meaningful group is obviously bounded by K (the total number of group candidates). Assume now that, to the previous K group candidates, we add K' new candidates. Then, for a group G_i to be meaningful, its probability of occurrence in the random model has now to be smaller. If K' tends to $+\infty$, this probability of occurrence must go to 0. This means that the meaningfulness depends on the size of the data. This size also has to be finite, else no group can be ε-meaningful for $\varepsilon > 0$. This is completely compatible with digital image processing where image sampling implies a finite amount of information. Moreover, a digital white noise image should yield no detection. It is thus sound to construct the a contrario model such that it is true at least in the case of white noise. In order to be coherent with Shannon's sampling theory, we have to assume that the distance between the objects O_i is larger than the Nyquist distance, namely 2 pixels, for they have to be independent in noise. In the following, we shall say that two points are independent if their distance is larger than 2 pixels.

Even though digital images are discrete by nature, it is often convenient to consider grey level images as functions from $\mathbb{R}^2$ to $\mathbb{R}$, which we always do in the following. In practice, we use a bilinear interpolation, which allows us to define level lines at any level. We also use finite differences scheme to define a contrast value which is consistent with the gradient.

A last important comment is the choice of ε, which is the only decision parameter. Of course the principle can be efficient only if it is robust with respect to ε. In fact, it is often possible to prove (see [11] and the following of this paper) that the minimal size of an ε-meaningful group depends on the logarithm of ε. In practice, setting $\varepsilon = 1$ means that we have less than one detection in the a contrario model. Thus we choose $\varepsilon = 1$, and check a posteriori that changing this value does not change the results, as predicted by the theory. One main reason why this is empirically very stable is that detected structures are (very) large deviations from the a contrario model and can be detected with some values of ε even less than 10^{-10}.

2.2. Contrasted Boundaries

In order to illustrate Helmholtz principle, we recall here the definition of meaningful boundaries given in [10]. It will be also useful since we will discuss this definition in the next sections. Let $u: \mathbb{R}^2 \to \mathbb{R}$ be a differentiable grey level image. Assume that we have a measure of contrast. To simplify we take it here equal to the norm of the gradient. Assume that we know the distribution of the gradient of u given by

$$H_c(\mu) = P(|Du| > \mu).$$

In fact, we do not really care that this contrast is the gradient norm of a differentiable function, and in practice, we shall take a finite differences approximation of the gradient. In [10], Desolneux, Moisan and Morel proposed the following definition.

Definition 1 ([10]). Let E be a finite set of N_{ll} level lines of u. A level line C is an ε-meaningful boundary if

$$NFA(C) \equiv N_{ll} H_c\Big(\min_{x \in C} |\ Du(x)|\Big)^{l/2} < \varepsilon, \qquad (2)$$

where l is the length of C. This number is called number of false alarms (NFA) of C.

We first remark that ε, N_{ll} and $\min_{x \in C} |Du(x)|$ being fixed, the minimal l such that C is ε-meaningful depends on the logarithm of the other parameters. This was already discussed in [11], and in practice we can take $\varepsilon = 1$ in all experiments. This definition will be further commented in 2.4.1, and we just shortly describe its implementation. We first need an a priori on the gradient law. The approximation by the empirical histogram is used. That is, we assume that the gradient norm is distributed following the law of the positive random variable X defined by

$$\forall \mu > 0, \quad P(X > \mu) = \frac{\#\{x \in \Gamma, |Du(x)| > \mu\}}{\#\{x \in \Gamma, |Du(x)| > 0\}}, \qquad (3)$$

where the symbol # designs the cardinality of a set, Γ the finite sampling grid, $|Du|$ is computed by finite difference approximation, and we assume that it is constant in each pixel. Moreover, we need a finite and reasonable set of level lines. Since images are assumed continuous, they have an infinite number of level lines. These lines are very redundant since interpolated images are very smooth. Thus, it is soundly assumed that quantized level lines contains all the information of the image. It is perceptually known that beyond a few hundreds of grey levels, we are not able to distinguish intensity differences. So we naturally quantize 8-bits encoded images every integer levels. Dividing the quantization step by 10 will approximately multiply the number of level lines by 10. Thus the number of false alarms of a given line will also be increased by a factor 10, which has nearly no incidence on the detection. For interpolated images, quantization yields a finite number of level lines. This number N_{ll} is dependent on the image; textured images have more level lines than more simple images. To give an order of magnitude, a natural 256×256 image contains between 10^4 and 10^5 level lines.

2.3. Maximal Boundaries

As remarked by Desolneux, Moisan and Morel [9], meaningful boundaries usually appear in parallel groups, because of interpolation. Moreover, since images are made band-limited before sampling, they are blurry and there is a transition layer around objects boundaries of width at least two or three pixels. These boundaries are redundant, and in applications, it may be useful to eliminate some of them. Remark that in artificial images, edges can be arbitrarily sharp and this problem does not appear if images are not interpolated. In order to eliminate boundaries redundancy, Desolneux et al. use the structure of the level lines tree developed by Monasse. For self-containness, we recall the basic facts of this construction and refer the reader to [28] for complements.

First, we define level lines as connected components of the topological boundaries of connected components of level-sets. With suitable assumption (semi-continuity), almost all level lines are Jordan curves. From Alexandrov's Theorem [1], a Jordan curve splits the plane in two connected components, exactly one of which is bounded and is called the *interior* of the curve. It turns out that since level sets are nested, level lines are also nested and this makes sense to say that a level line L_1 is included in the interior of another line L_2. Now, in Section 1, we defined the upper level sets. We can define the lower level sets in the same way by

$$\chi^{\lambda}(u) = \{x \in \mathbb{R}^2,\ u(x) \leq \lambda\}.$$

Of course, we can also define lower level lines but it turns out that the family of upper and lower level lines

Figure 2. Maximal meaningful boundaries. (1). Original image, 83,759 level lines (2). All meaningful boundaries: 11,505 detections. (3). Maximal meaningful boundaries. Only 883 boundaries remain, while the visual loss is very weak.

are the same. Nevertheless, a grey level is naturally attached to a level line (it is the grey level of the level set it comes from) and it may differ whether we consider the level line as a lower or upper level line. Moreover, since lower level sets are nested, they also can be embedded in a tree structure (as upper level sets). In general, those two trees (the tree of upper and lower level sets) are different. Due to occlusions, connected components of level sets appear with holes (that is, they are not simply connected), and this information is not easily retrieved from either upper or lower level sets. Monasse and Guichard [29] proved that it is possible to fuse both trees into a single one. This allows to attach a grey level to a level line with no ambiguity, together with the information that the line comes from an upper (resp. lower) level set. A node in the tree is a connected component of an upper- or lower level set, with its holes filled, and is called a *Shape* by Monasse. As a consequence a node in the tree is a simply connected plane set.

Compared to (lower or upper) level set tree, Monasse's tree looses the monotonicity of grey level when following branches of the tree. On the contrary, in case of contrast reversal (because of an object occluding another one), grey level is not monotone. The image can still be reconstructed from the tree in a very fast way: indeed, the grey level of a pixel is equal to the grey level of the smallest level line containing the pixel. Remark that this reconstruction also permits to reconstruct an image from any subtree. The tree of level lines is computed by a fast region growing algorithm [22, 29]. For a bilinear interpolated image, the tree contains infinitely many nodes, but it can be retrieved from a subtree (called the fundamental bilinear tree) containing all the lines passing through pixel corners and saddle points. Considering quantized levels is equivalent to extract a finite subtree, which we do in practice.

Definition 2 ([28]). A monotone section of a level lines tree is a part of a branch such that each node has a unique son and where grey level is monotone (no contrast reversal).

A maximal monotone section is a monotone section which is not strictly included in another one.

Definition 3 ([10]). We say that a meaningful boundary is maximal meaningful if it has a minimal NFA in a maximal monotone section of the tree of meaningful level lines.

Figure 2 illustrates that the loss of information of maximal meaningful boundaries is negligible compared to the gain of information compactness.

Since maximal meaningful boundaries inherit the tree structure of the tree of level lines, they can be used to reconstruct an image, thus defining an image operator, see Fig. 3. As explained above, the grey level of a pixel can be set to the grey level of the smallest (with respect to inclusion in the tree) meaningful level line that contains this pixel. However, the grey level of a meaningful level line does not well represent the gray level in the interior of the line since, by definition, grey level has large variations along this line. For visualization purposes, it may be useful to replace the grey level of the line by averaging the original grey level. More precisely, the grey level in the interior of a level line L is averaged in the complementary of the meaningful level lines contained in the interior of L.

Remark 1. Both reconstructions above define operators that are connected operators as defined by Salembier and Serra [37], that is to say, it merges connected components where the image is constant. Equivalently, connected components of level sets can only disappear.

It is neither a contrast invariant operator, since it explicitly uses the gradient value (it only commutes with affine global contrast change).

Figure 3. Original image on the left (99,829 level lines). Middle: reconstruction from the 394 maximal meaningful boundaries. The gray level may not be really significant since, on edges, the maximal meaningful level line has an intermediate level between both sides of the edge. It is more perceptually adequate to set the grey level to the average grey level in a meaningful boundary, after removing inner meaningful boundaries interior (right image). Nevertheless, this is simply useful for visualization purposes since for contrast independent shape recognition purposes, we do not use the grey level value, but only the geometry of level lines. The main point is that we preserved the main shapes, while removing the textures.

As remarked by Salembier and Serra (see also Breen and Jones [2]), an operator pruning the topographic maps preserves edges geometry very well and this can be checked on Fig. 3.

2.4. *Discussion on the Definition of Meaningful Contrasted Boundaries*

2.4.1. Interpretation of the Number of False Alarms. In this section, we give a precise interpretation of Definition 1, which was not explicit in [10]. Let us first recall the following classical lemma.

Lemma 1. *Let X be a real random variable and $H(x) = P(X \geq x)$. Then for all $t \in [0, 1]$,*

$$P(H(X) < t) \leq t.$$

Assume that X is a real random variable described by the inverse repartition function $H(\mu) = P(X \geq \mu)$. Assume that u is a random image such that the values $|Du|$ are independent with the same law as X. Let now E be a set of random curves (C_i) in u such that $\#E$ (the cardinality of E) is independent of each C_i. For each i, we note $\mu_i = \min_{x \in C_i} |Du(x)|$. We also assume that we can choose L_i independent points on C_i (points that are afar at least by Nyquist's distance). We can think of the C_i as random walks with independent increments but since we choose a finite number of samples on each curve, the law of the C_i does not really matter. We assume that L_i is independent from the pixels crossed by C_i.

We say that C_i is ε-meaningful if

$$NFA(C_i) = \#E \cdot H(\mu_i)^{L_i} < \varepsilon.$$

Proposition 1. *The expected number of ε-meaningful curves in a random set E of random curves is smaller than ε.*

Proof: Let us denote by X_i the binary random variable equal to 1 if C_i is meaningful and to 0 else. Let also $N = \#E$. Let us denote by $\mathbb{E}(X)$ the expectation of a random variable X in the a contrario model. We then have

$$\mathbb{E}\left(\sum_{i=1}^{N} X_i\right) = \mathbb{E}\left(\mathbb{E}\left(\sum_{i=1}^{N} X_i \mid N\right)\right).$$

We have assumed that N is independent from the curves. Thus, conditionally to $N = n$, the law of $\sum_{i=1}^{N} X_i$ is the law of $\sum_{i=1}^{n} Y_i$, where Y_i is a binary variable equal to 1 if $nH(\mu_i)^{L_i} < \varepsilon$ and 0 else. By linearity of expectation,

$$\mathbb{E}\left(\sum_{i=1}^{N} X_i \mid N = n\right) = \mathbb{E}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \mathbb{E}(Y_i).$$

Since Y_i is a Bernoulli variable, $\mathbb{E}(Y_i) = P(Y_i = 1) = P(nH(\mu_i)^{L_i} < \varepsilon) = \sum_{l=0}^{\infty} P(nH(\mu_i)^{L_i} \mid L_i = l) P(L_i = l)$. Again, we have assumed that L_i is independent of the gradient distribution in the image. Thus conditionally to $L_i = l$, the law of $nH(\mu_i)^{L_i}$ is the law

of $nH(\mu_i)^l$. Let us finally denote by $(\alpha_1, \ldots, \alpha_l)$ the l (independent) values of $|Du|$ along C_i. We have

$$\begin{aligned} P(nH(\mu_i)^l < \varepsilon) &= P\left(H\left(\min_{1\le k\le l} \alpha_k\right) < \left(\frac{\varepsilon}{n}\right)^{1/l}\right) \\ &= P\left(\max_{1\le k\le l} H(\alpha_k) < \left(\frac{\varepsilon}{n}\right)^{1/l}\right) \\ &\qquad \text{since } H \text{ is nonincreasing} \\ &= \prod_{k=1}^{l} P\left(H(\alpha_k) < \left(\frac{\varepsilon}{n}\right)^{1/l}\right) \\ &\qquad \text{by independence} \\ &\le \frac{\varepsilon}{n} \text{ from Lemma 1.} \end{aligned}$$

This term does not depend upon l, thus

$$\begin{aligned} &\sum_{l=0}^{\infty} P(nH(\mu_i)^{L_i} < \varepsilon \mid L_i = l) P(L_i = l) \\ &\quad \le \frac{\varepsilon}{n} \sum_{l=0}^{\infty} P(L_i = l) = \frac{\varepsilon}{n}. \end{aligned}$$

Hence,

$$\mathbb{E}\left(\sum_{i=1}^{N} X_i \mid N = n\right) \le \varepsilon.$$

This finally implies $\mathbb{E}(\sum_{i=1}^{N} X_i) \le \varepsilon$, which exactly means that the expected number of meaningful curves is less than ε. □

In this proposition, we have not assumed a priori that the C_i are level lines of u. Indeed, in this case, we cannot certainly assert that the length (number of independent points) of the curve is independent from the values of the gradient along the curve.

2.4.2. Cleaning-up Meaningful Boundaries. Proposition 1 asserts that if a curve is a meaningful boundary, then it cannot be *entirely* generated in white noise (up to ε false detections on the average). On the other hand, can we guarantee that no part of a meaningful boundary is contained in noise? Or, for a given meaningful boundary, can we give an upper bound of the size of the part of the boundary that is likely to be contained in noise (i.e. a non-edge region)? To answer this question, we use the a posteriori length distribution

$$P\left(L \ge l \,\middle|\, \min_{x\in C} \left|Du(x)\right| \ge \mu\right). \tag{4}$$

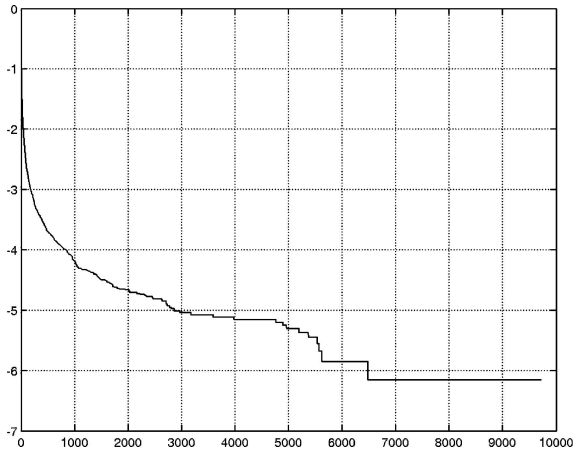

Figure 4. Log10 of the inverse repartition function of length of level lines in a white noise image. The average length is about 3.5, meaning that most level sets enclose a single pixel.

Contrarily to the probability appearing in Definition 1, this one penalizes long curves not only through the gradient value. To compute it, we need the a priori distribution $P(L \ge l)$ that a level line in noise has a length larger than l. As we do not know this distribution explicitly, we choose to estimate this law empirically. For $l \le 1000$ (to give an order of magnitude), the number of lines whose length is larger than l is still quite large (for images of size about 500×500), and we assume that the distribution is quite correctly estimated for such length (see Fig. 4). For higher values, there are too few level lines. By using Bayes' rule, we derive

$$\begin{aligned} &P\left(L \ge l \,\middle|\, \min_{x\in C} \left|Du(x)\right| \ge \mu\right) \\ &\quad = \frac{\sum_{k=l}^{\infty} P(\min_{x\in C} \mid Du(x)| \ge \mu | L = k) P(L = k)}{\sum_{k=1}^{\infty} P(\min_{x\in C} |Du(x)| \ge \mu | L = k) P(L = k)}. \end{aligned}$$

(The denominator is nothing but $P(|Du| > \mu)$). By the a contrario assumption (independence of the gradient along curves), we can still write

$$\begin{aligned} p_\mu(l) &\equiv P\left(L \ge l \,\middle|\, \min_{x\in C} \left|Du(x)\right| \ge \mu\right) \\ &= \frac{\sum_{k=l}^{\infty} H_c(\mu)^k P(L = k)}{\sum_{k=1}^{\infty} H_c(\mu)^k P(L = k)}. \end{aligned} \tag{5}$$

Let us now consider an image u with N_{ll} (quantized) level lines. We also denote by N_l the number of all possible sampled subcurves of these level lines. (N_l is the sum of the squared number of independent points of the lines if they are closed).

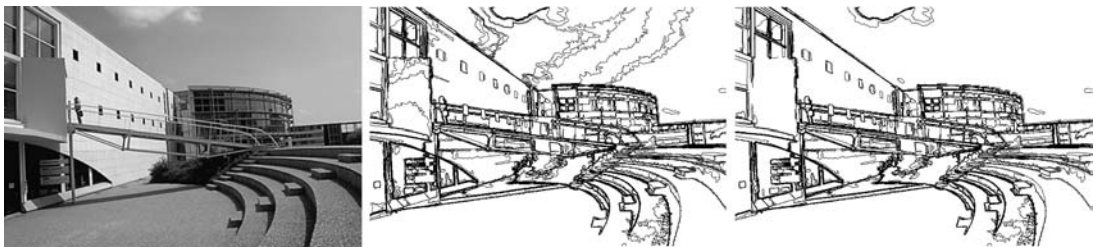

Figure 5. Meaningful boundary clean-up. On the left the original image. In the middle, the meaningful boundaries with local histograms, see Section 4. Boundaries are found in the sky. They are detected since the gradient in the sky is regular because of the smoothly changing illumination. The gradient value is about 0.2. Even though they are not smooth at small scale (they cannot be well located, due to the too small gradient), they are nearly parallel at large scales, which can be explained, a posteriori. Now, these boundaries may not be very useful for shape recognition purposes, because of their bad localization. On the right, the result after the clean-up procedure with a gradient threshold equal to 1.

Assume that C is a piece of level line with L independent points, contained in a non-edge part, described by the noise model. We want to estimate the probability that L is larger than $l > 0$, knowing that $|Du| \geq \mu$. This is exactly $p_\mu(l)$, the probability defined in (5). As in Proposition 1, we can prove that $N_l \cdot p_\mu(l)$ is an upper bound of the expected number of *pieces* of lines of length larger than l with gradient larger than μ. For a fixed μ, let be l such that $N_l \cdot p_\mu(l) \leq \varepsilon$. Then, we know that, on the average, we cannot observe more than ε pieces of level line with a length larger than l and a gradient everywhere larger than μ. We make the assumption that a point with a gradient less than μ is located in noise. Let us remove any piece of length l containing such a point. Then all remaining points belongs to a piece of curve with length larger than l with gradient larger than μ, which cannot be due to chance.

This yields a clean-up algorithm for boundary detection.

1. Detect meaningful boundaries.
2. For a fixed $\mu > 0$, let $\mathcal{L}(\mu) = \inf\{l,\ N_l \cdot p_\mu(l) < \varepsilon\}$.
3. For any meaningful boundary, remove every subcurve of length $\mathcal{L}(\mu)$ containing a point where $|Du| \leq \mu$.

This introduces a parameter, μ. When μ gets larger, $\mathcal{L}(\mu)$ decreases, so that the clean-up removes more numerous but smaller pieces of curves. The choice of μ can be determined by applicative considerations. Detected edges may be used for different purposes, for instance shape recognition or image matching. Letting $|Du|$ less than 1, means that we may detect edges with an accuracy less than one pixel. Thus choosing to eliminate pieces of curves with a gradient less than $\mu = 1$ for all images is not restrictive. In practice, the remaining pieces of level lines have a gradient much more than 1 and can be well enough located. We also check that for μ about 1, we obtain values of $\mathcal{L}(\mu)$ less than a few hundreds, which is compatible with the empirical estimation of the a priori length distribution. The result of the clean-up procedure is illustrated on Fig. 5.

3. Multiscale Meaningful Boundaries

3.1. Meaningful Boundaries by Downsampling

As previously noted above, the contrast measure is an approximation of the gradient by finite differences. More precisely, Desolneux et al. use the following scheme:

$$\frac{\partial u}{\partial x} \simeq u_x(i, j) = \frac{1}{2}(u(i+1, j) + u(i+1, j+1) - u(i, j) - u(i, j+1)), \quad (6)$$

$$\frac{\partial u}{\partial y} \simeq u_y(i, j) = \frac{1}{2}(u(i, j+1) + u(i+1, j+1) - u(i, j) - u(i+1, j)). \quad (7)$$

Using a 2×2 scheme is coherent with the application of Helmholtz principle: points afar from the Nyquist distance have independent values of contrast in white noise. On the other hand, this value is sensitive to noise. This problem was known from Marr and Hildreth [25] who considered that edge detection should be multiscale. They compute the zero-crossing of the Laplacian of the image convoluted with Gaussians with different standard deviations. Since edges at larger scale are badly located, they propose to track back the strongest edges to smaller scales, which is not obvious in practice. Smoothing introduces local

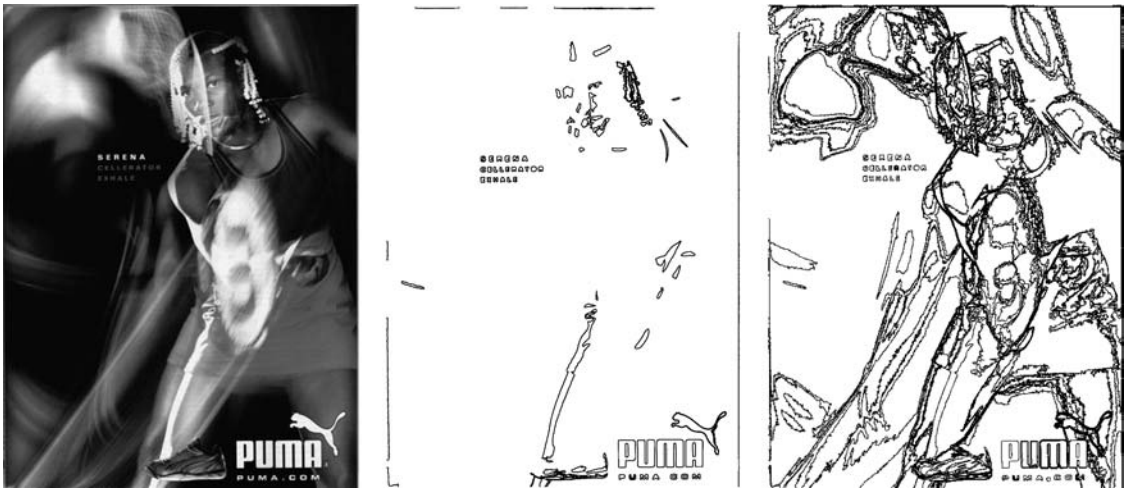

Figure 6. Influence of quantization noise on meaningful boundaries. On the left, the original image is coarsely quantized since it has a very low contrast. This leads to bad gradient estimation and a lot of missing detections (middle). Multiscale detection is less sensitive to quantization noise and leads to more correct detections (right).

dependencies between pixels, making the a contrario model false in smoothed white noise. Nevertheless, the a contrario models still applies if we downsample the image at a lower frequency, given by the amount of smoothing. More precisely, we apply the following algorithm. Consider a set $\{1, 2, \ldots, 2^{N_s-1}\}$ of N_s dyadic scales. For any level line C, we denote by C^s the curve $\frac{C}{2^s}$, obtained by scaling C by a factor 2^{-s}. We also denote by H^s the empirical contrast distribution of u^s, where u^s is obtained by downsampling u with a factor 2^s, in conformity with Shannon's theory. (That is to say, downsampling follows an adequate smoothing, for instance convolution with a prolate function.)

1. Compute the quantized level lines of the image u.
2. For each level line C with l independent points in u, compute μ^s, the minimal value of $|Du^s|$ over all pixels crossed by C^s. Let

$$NFA(C) = N_s \cdot N_{ll} \min_{s\in\{0,\ldots,N_s-1\}} (H^s(\mu^s))^{l/2^s}. \quad (8)$$

We say that C is meaningful if $NFA(C) < \varepsilon$.

Thus, a curve is meaningful if and only if there exists a scale such that it is $\frac{\varepsilon}{N_s}$ meaningful in the sense of the previous section. A direct corollary of the linearity of expectation and of Proposition 1 is that the expected number of ε-meaningful multiscale boundaries is less than ε in the a contrario model. Note that C^s is not a level line of u^s, but this is not required in Proposition 1. Moreover, if C was already $\frac{\varepsilon}{N_s}$-meaningful, then we are sure that C is still detected by the multiscale method. It is clear that we only consider a small number of dyadic scales (say 3 or 4), else images will only contain a few pixels. Since the detection depends on $\log \varepsilon$, we do not eliminate many lines by considering $\frac{\varepsilon}{N_s}$-meaningful boundaries at each scale. On the other hand, the method should be numerically less sensitive to white noise since filtering followed by downsampling reduces noise. On Figs. 6 and 7, we show the result of this multiscale method on images with quantization noise and additive Gaussian noise. On Fig. 6, the shapes are not very sharp because of motion blur and transparency. Level lines following contours are very long since they surround several objects. Moreover, the background is nearly uniform. Thus the minimal value of contrast along long level lines is all the more sensitive to the gradient computation. The effect is also dramatic in the noisy image of Fig. 7 (Gaussian noise with standard deviation 30). Because of smoothing and downsampling, the images appear more blurry, and the gradient is more stable. The fact that downsampled level lines are not level lines of the downsampled image is not an obstruction since number of false alarms may be computed on any curve in the a contrario model.

3.2. Meaningful Boundaries vs. Haralick's Detector

In this section, we comment the main differences between the meaningful boundary model and the classical

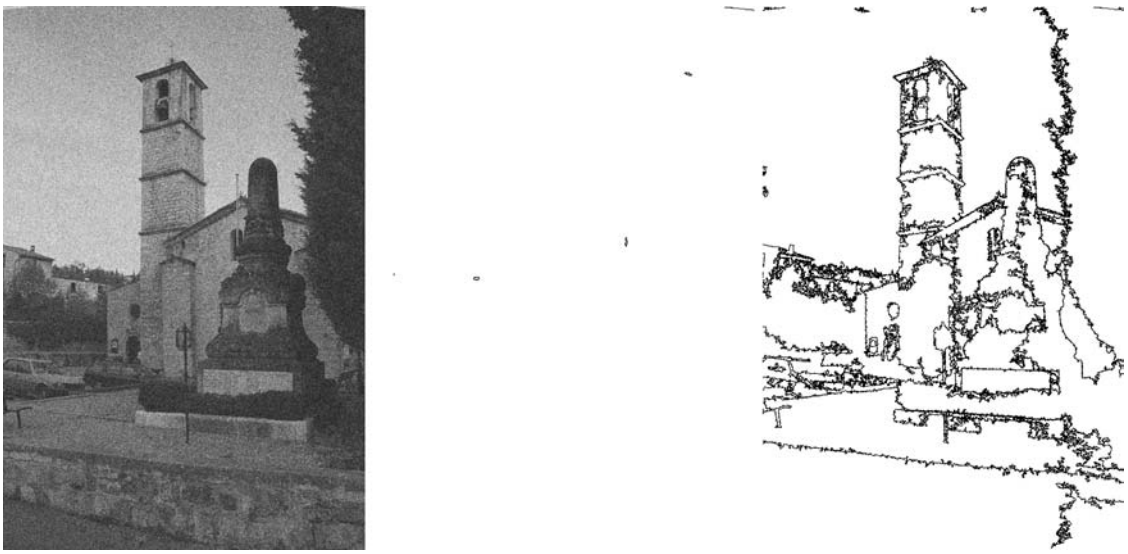

Figure 7. Multiscale meaningful boundaries and noise. Left: image of Fig. 10 with an additive white Gaussian noise of standard deviation 30. Middle: meaningful boundaries. Since noise dominates the gradient distribution, only 6 small level lines are detected. Right: multi-scale detection using 4 dyadic scales. Textures are not detected, meaning that noisy textures are in this case not different enough from noise to be detected. On the other hand, main structures remain. This allow to empirically check the stability of the topographic map in spite of the important amount of noise.

edge detector introduced by Haralick. The meaningful boundaries are based on the topographic map of grey level images, which gives a complete topological representation of grey level images. Caselles et al. [5, 6] detail all the properties of this representation. A first advantage of this representation is its stability: even with an important amount of noise, many level lines do not change much. Our multiscale approach also makes the detection quite robust (see Fig. 7). A second advantage is its invariance with respect to global contrast change. Meaningful boundaries are not contrast invariant since they use the distribution of contrast, but they are still invariant with respect to affine contrast change. But the main property is the structure of this representation: it is a set of nested curves that are either closed or meet the image boundary. As a consequence, level lines have two of the main properties usually expected in edge detection or image segmentation: they are curves (and not sets of points), and are embedded in a hierarchical structure [14, 30, 39]. Moreover, away from critical points, level lines coincide with isophotes. As a consequence, for almost any level, the gradient is almost everywhere normal to level lines, which makes level lines good candidates for edges.

Following Haralick [15], edges are the maxima of the gradient norm in the direction of the gradient, such that the gradient is larger than a given threshold. Thus, for a grey-level image u, they are the zero-crossings of $D^2u(Du, Du)$. Since, this quantity is numerically sensitive to noise, a multiscale strategy *à la* Marr is applied. Thus in practice, u is first convolved with a Gaussian with standard deviation σ (we denote by g_σ this Gaussian and $u_\sigma = g_\sigma * u$) and the points where $D^2u_\sigma(Du_\sigma, Du_\sigma)$ changes sign and $|Du_\sigma| > \mu$ are edges points. Although there have been some attempts to automatically determine the scale parameter σ [20], edge detection widely remains multiscale as predicted by Marr [24], and it is quite difficult to track edges back to small scales. The multiscale meaningful boundaries detection of the previous section allows to consider different scales, while keeping detection thresholds completely automatic. Moreover, the number of scales has a log influence. Haralick's detector provides with a set of points or a few pixels long curves. The way they should be connected is far from obvious and may lead to a very high computational complexity; this problem is structurally handled by level lines. Last but not least, Haralick's operator is inefficient for corners and junctions. Indeed, at those points, the gradient direction is very badly estimated and edges may be severely cut. Additional algorithms are necessary to reconnect pieces of edges. On the opposite, level lines bifurcate

Figure 8. Junction and level lines. On the left, the original image. Middle, Haralick's detector implemented with Canny's filter on the area designated on the left image. Note how the contour is broken at the junction, due to the bad estimate of the gradient direction, and the high number of edge pieces. Right: detailed view of meaningful boundaries on the region. There are two level lines, each corresponding to an edge part.

at junctions, thus handling the different boundaries. Figure 8 shows the meaningful boundaries and Canny's filter near two junctions. First, edges gives very small pieces of curves. Even though there are some linking procedures, we consider any additional algorithm as a drawback. The behavior of the level lines around the *T*-junctions are quite clear. When extracting shapes elements by local encoding, all the different configurations near the junctions will be considered.

4. Local Boundary Detection

In the model above, the values of the gradient are random variables whose distribution is empirically estimated. It is simply the histogram of the gradient in the image. One can argue that this distribution is too global. This also yields what we call the "blue sky effect". Consider an image containing two parts: a contrasted or textured one (e.g. ground) and a smooth one (e.g. sky). Then, we can observe an overdetection in the ground, and an underdetection in the sky. Indeed, the sky only contributes with small values in the histogram. Thus we tend to detect anything which is more contrasted than the sky, and nearly anything is detected in the ground. On the contrary, the contrasted ground makes the detection more difficult for regions with a small contrast. This is not in agreement with human vision, since we locally adapt our perception of contrast. Objects are masked in contrasted regions, while our accuracy is improved in low contrasted regions (up to some physiological thresholds).

In this section, we address this local adaptivity to contrast. It does not use new concepts and is an adaptation of the meaningful boundary model. We first describe the algorithm, then show experiments.

4.1. Algorithm

Assume that we have detected a closed boundary. Then it divides the image into two connected components: the interior and the exterior of the curve. Then, we can compute the empirical contrast distribution in the interior on the one hand and in the exterior on the other hand. We then independently detect new meaningful boundaries in each connected component. We then apply this procedure recursively. Since the size of the level line tree is finite, it is clear that we end the detection in a finite number of steps.

The situation is actually a bit more complicated. First, this method depends on the order we use to describe the image boundaries. We simply choose to start with the most meaningful boundaries. Second, boundaries are not always closed. In this case, their endpoints belong to the image border. They still cut the image into two connected components. Unfortunately, there is no clear notion of interior and exterior. An algorithmic choice is made, but is purely algorithmic and arbitrary from a perceptual point of view [28]. Thus, we cannot rely on this choice of interior, which conflicts with closed boundaries. However, we can first apply the detection to open boundaries, then to the closed ones. (Open boundaries contain all the closed ones, since level lines are nested.) More precisely, we proceed as follows.

Let us call R_0 the root boundary, that is the (non-meaningful) boundary containing all the image. If C is

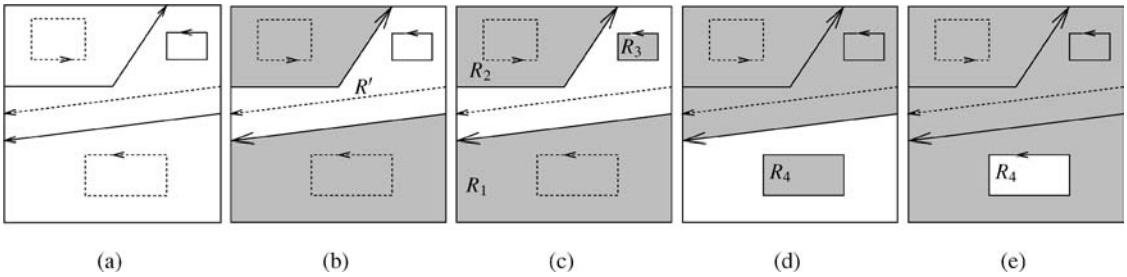

Figure 9. Example of local search of meaningful boundary. (a) the initial boundaries. They are oriented such that the tangent and the interior normal form a direct frame. We compute the NFA of each boundary. In solid line, we draw the total meaningful ones. Two are open, one is closed. Remark that the interior are disjoint, because of total maximality. While we detect some open curves, we ignore the closed ones. (b) While we detect new meaningful boundaries, we compute the contrast histogram in the complementary of the interior of the open detected boundaries and resume search in this part of the image. In R', the exterior of the detected open boundaries, we detect a total maximal boundary. Remark that this boundary may have been already detected but rejected because of open boundaries. We assume here that no new open boundaries are meaningful. Thus, we keep this closed boundary. (c) We resume the search (with recomputed histogram) in the exterior (white part) of the detected boundaries, until we cannot find new ones. When this is over, we then compute the local contrast histogram in each region R_1, R_2, R_3 and look for boundaries inside them. (d) A boundary R_4 has been detected in R_1. Compute the local histogram in $R_1 \setminus R_4$ and detect boundaries. (e) Finally, we scan for boundaries in R_4 with new local contrast histogram.

a boundary, we denote by $\operatorname{Int} C$ its interior.

1. Set $R \leftarrow R_0$. (Local root.)
2. Set $\mathcal{M}$, the set of already stored in R meaningful boundaries. Initially, $\mathcal{M}$ is empty.
3. Let $R' \leftarrow R \setminus \cup_{C \in \mathcal{M}} \operatorname{Int} C$.
4. Compute the histogram of $|Du|$ in R'.
5. Use this histogram and detect the maximal meaningful boundaries included in R'. Let $\mathcal{N}$ be the maximal meaningful boundaries defined by $C \in \mathcal{N}$ if and only if

$$\begin{cases} \operatorname{Int}(C') \subsetneq \operatorname{Int}(C) \Rightarrow NFA(C) < NFA(C') \\ \operatorname{Int}(C) \subset \operatorname{Int}(C') \Rightarrow NFA(C) \leq NFA(C'). \end{cases} \quad (9)$$

Otherwise said, the boundaries in $\mathcal{N}$ have an optimal NFA, since they are more meaningful than boundaries that contain them or are contained in them. Note that this is stronger than the maximality defined in Section 2.3 since we go across monotone sections. We call the boundaries in $\mathcal{N}$ the total maximal boundaries. The subtree with root equal to R that remains by keeping only the boundaries in $\mathcal{N}$ has only two levels: the local root R, and $\mathcal{N}$. Since the interior of open boundaries is arbitrary, we do not mix the detection of open and closed boundaries. In practice, this means that if we detect an open meaningful boundary C, we apply the definition of total maximal boundary (9) only to open boundaries containing C or contained in C.

6. If $\mathcal{N} \neq \emptyset$, then we have detected new boundaries in the complementary of the already detected ones. Then,
   (a) Set $\mathcal{M} = \mathcal{M} \cup \mathcal{N}$. By construction, all the closed boundaries in $\mathcal{M}$ have disjoint interior.
   (b) return to Step 3.
7. If $\mathcal{N} = \emptyset$, there are no new boundaries in the local root and in the complementary of the currently detected boundaries. We then continue the search at lower levels of the tree. For any boundary $C \in \mathcal{M}$,
   (a) Store C.
   (b) Set $R \leftarrow C$, and $\mathcal{M} \leftarrow \emptyset$.
   (c) Return to Step 3.

The algorithm is illustrated in Fig. 9.

Remark 2. Each boundary may be tested more than once. Thus, the number of false alarms has to be multiplied by the maximal number of visits of a boundary, which is bounded by above by the level lines tree depth. In fact, each detected boundary often lies in the middle of the local root, and this divides the tree depth by 2. Thus in practice, the maximal number of visits of a boundary is like the logarithm of the initial tree depth. In practice, it is always much smaller than 100.

4.2. *Experiments on Locally Contrasted Boundaries*

In Fig. 10, we show the difference between the detection with a global contrast histogram and the updated

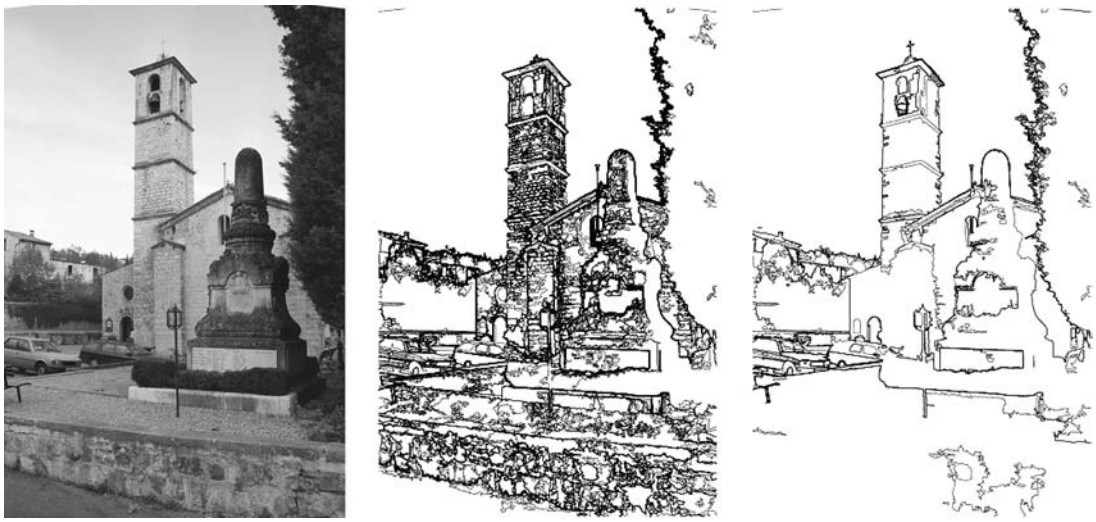

Figure 10. Influence of local contrast. From left to right: original image, maximal meaningful boundaries, local maximal meaningful boundaries. There are 280,000 boundaries in the initial image, 652 in the second one and 193 in the last one. Texture is removed since local contrast (for instance) on the church tower is much more demanding than the global histogram. As the texture is uniform, no level line is a large deviation to the empirical local contrast, yielding no detection. This is very good for shape analysis where we often want to distinguish texture from real shapes.

local histogram. To give an idea of magnitude of the number of false alarms, the boundary delimiting sky and the foreground has a NFA equal to 10^{-357}. This means, that, in order to observe such a contrasted line in noise, we need to observe on the average 10^{357} images. The smaller boundaries around the opening on the top of the tower have NFAs about 10^{-10}.

Very interestingly, using local contrast removes boundaries in texture. This is logical since the local contrast in textured regions (as on the tower) assumes larger values than in the rest of the image. Thus, this decreases the NFA of boundaries and most of them simply disappear in textured regions. This is a masking phenomenon.

Let us explain why this is useful for shape recognition. In general, a shape recognition algorithm can be divided in four steps:

1. extraction of shapes,
2. (invariant) encoding,
3. comparison: compute some distance between encoded shapes.
4. decision: accept or reject pairs of matching shapes.

Present and future applications need to compare images in huge databases, where we have no a priori that two images, or two shapes should match. Since every procedure in the above methodology is very costly, it is interesting to limit the number of encoded shapes and to try to keep the "most meaningful".

For the time being, there is no general model of shapes [43]. Nevertheless, for shape recognition algorithm, we can give empirical observations of what a "good shape" is. First, it should be stable in terms of extraction. This is generally expressed in terms of contrast and regularity, and the method we describe in this paper gives quantitative arguments. (Regularity is the object of the next section.) For encoding, a good shape should not be too simple, especially if we are interested in an invariant recognition. For instance, most convex shapes are very alike in affine invariant shape recognition. Assume that we have chosen an affine invariant distance between shapes. If we want to be sure that two convex shapes match, the distance between them has to be very small. Indeed, two convex shapes can casually be close to each other, while the probability that it occurs for more complex shapes is very small (this means that recognition is relative to the database and to the query [31]). On the other hand, a shape should not be too complex, since complexity usually makes the encoding longer and more difficult. Because of occlusions, we usually try to match pieces of shapes. Very complex shapes will be divided in numerous pieces, making computations longer.

Now, it is well known that texture is strongly damaged by compression. Thus level lines in texture may not be reliable when two images come from different sources (with different quality, compression rate etc...). Moreover, they are very complex, and yield many encoded pieces of curves. If these curves match for two different images, then those images are certainly exactly the same. The computational cost may be too high for some applications, where we may want to detect a particular shape (a logo for instance) in a database. Thus it may be useful to automatically remove contrasted regions corresponding to texture. This is what the local contrast detection makes in practice.

The argument above is reversed for stereo images registration. In this case, we have the strong a priori that the images are nearly the same, and the goal is to register them as best as possible. In this application, textures can also give some useful information (see Fig. 11).

The effect of local contrast in boundaries detection is twofold: first, textures are eliminated. On the contrary, local contrast should make curves in low contrasted areas more detectable. This is also what we empirically observe: we detect illumination gradient (see Figs. 5 and 12). They can be due to the vicinity of the light source, or to the variation of the orientation of the

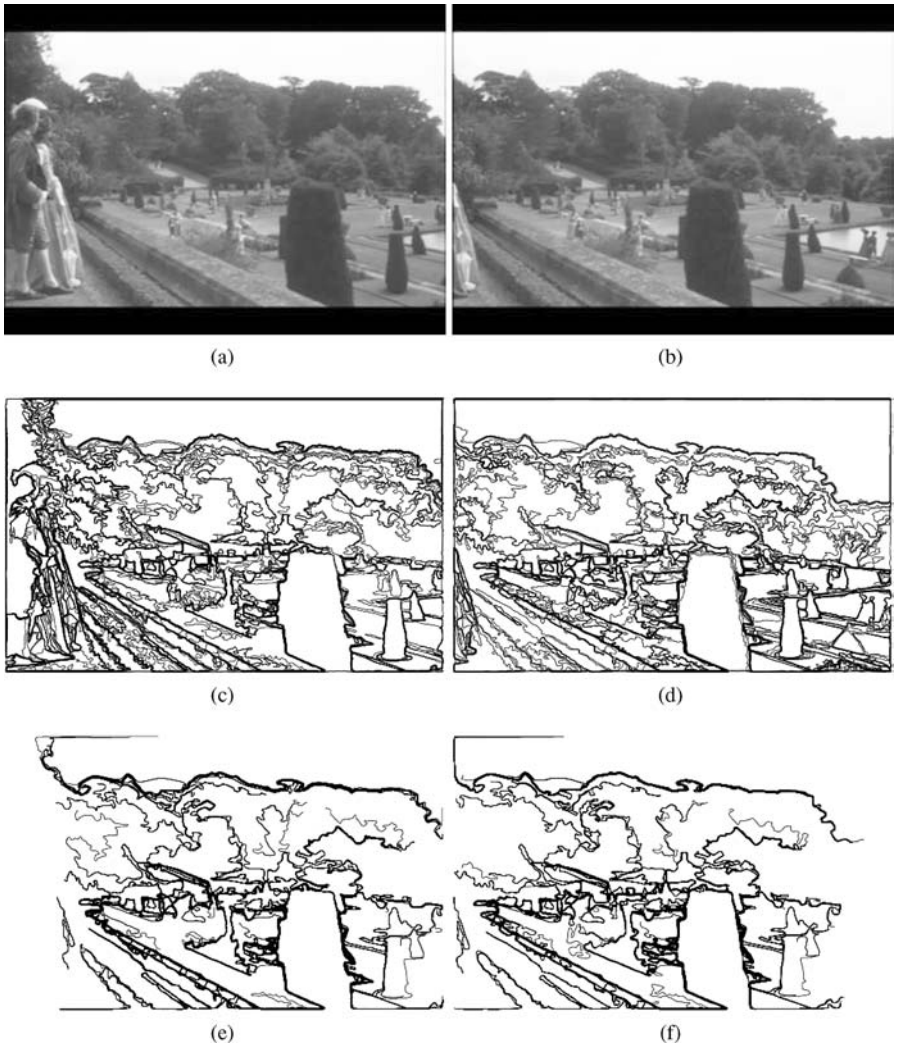

Figure 11. Image registration. (a) and (b) are two images from a movie during a rightward traveling. (c) and (d) are the meaningful boundaries in the previous images. (e) and (f) are the pieces of level lines of (c) and (d) that match with a number a false alarms less than 10^{-7}. We use the algorithm developed by Lisani et al. [21,31,32], which uses an a contrario definition of shape matching.

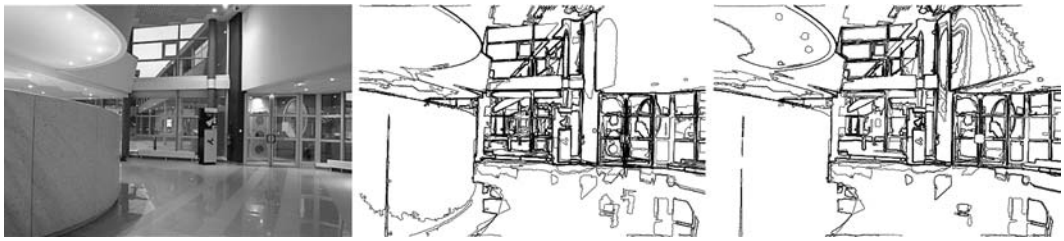

Figure 12. Illumination, local contrast and regularity. Left: original image. Middle: meaningful contrasted boundary. Right: meaningful contrasted and smooth boundary with local contrast. With contrast only, a single boundary appears on the right with the contrast due to illumination. If contrast is localized, then more boundaries are detected. If we also add a regularity constraint (see Section 5.1 below), there are still more detections. These boundaries are very different from texture since they are nearly convex and parallel. They are eliminated by the cleaning procedure described in Section 2.4.2.

surface of a three dimensional object with respect to the light source. Such lines do not correspond to the usual notion of shapes (objects). Nevertheless, it is logical to detect them as remarkable structures.

4.2.1. Geometrical Information Reduction. Caselles et al. claim that pieces of level lines are the basic objects of image analysis [5]. Then, suitably encoded (that is to say in a stable and invariant way) pieces of level lines could directly be used to feed a shape recognition algorithm. There is no theoretical obstruction to encoding all level lines. Lisani et al. [21] describe an encoding method for shape matching. Intrinsic frame for pieces of level line are defined from inflexion points or bitangent points. These normalized curves are taken as shape elements (we call them *codes*); their number is basically proportional to the number of inflexion points. Thus, long and oscillating curves are very costly, and at the time being, the method is not applicable in reasonable time if all the topographic map is encoded. The MB model is used to compute a raw primal sketch [24], that contains most information of shape contours. Experiments in [10] and this article show that the MB model gives sufficient information. It is sufficiently complete and it experimentally compresses the information so that encoding can be made more rapidly. To numerically illustrate this, we chose a small sample of 23 natural images of different types. We then compute normalized shapes elements dictionary by Lisani's encoding algorithm on a subset of images level lines, namely

1. all meaningful boundaries,
2. only maximal meaningful boundaries,
3. maximal meaningful boundaries with local contrast,
4. cleaned (see Section 2.4.2) maximal meaningful boundaries with local contrast.

We then observe the number of codes per pixel and the CPU time of the encoding per pixel (see Table 1). (Note that this time varies much from one image to another.) The gain between meaningful boundaries and maximal meaningful boundaries is obvious and due to the redundancy elimination in the tree of level lines. The gain of local meaningful boundaries is experimental since we can construct images with more local meaningful boundaries that maximal meaningful boundaries. Since the cleaning procedure removes some part of level lines, the encoding is logically faster, and shape elements dictionary shorter.

5. Meaningful Boundaries or Snakes?

Desolneux et al. [12] compared the MB model with variational snake theory. This may seem a bit weird since the MB model only uses contrast observations along a curve, while snakes are also required to be

Table 1. Number of shape elements encoded by Lisani's algorithm for shape matching [21,33]. Using local boundaries and cleaning procedure makes the encoding much faster. Moreover, the shape elements dictionary are shorter, but they experimentally contain all characteristic objects pieces of boundaries. The matching phase complexity is directly proportional to the number of codes. Thus, this simplification is algorithmically very interesting.

	#codes/pixel	CPU (s)/pixel
All MB	0.1458	0.0024
Maximal MB	0.0528	0.0006
Local MB	0.0310	0.0004
Cleaned local MB	0.0132	0.0002

smooth. In fact, the explanation for natural images is that contrasted boundaries often locally coincide with objects. Thus, they are also incidentally smooth. Whereas smoothness seems to be optional for the detection, it may give a better localization of the contour. In this section, we study the possible influence of smoothness in the detection to see, whether or not, smoothness is fundamental in the detection. We conclude the following: there are only few additional detections, while the position of the maximal meaningful boundaries may change a little bit. The NFA also significantly decreases. The small number of new detections and the fact that each partial detector can detect most image edges prove *a contrario* that contrast and regularity are not independent in natural images.

An a contrario model of regularity has been proposed in [4]. It assumes that the variation of the orientation of the tangent between two samples is a random value uniformly distributed in $(-\pi, \pi)$. Thus, the implicit a contrario model is random walks with isotropic and independent increments. This model is not really adapted for the following reason. All the curves we detect are level lines, thus boundaries of compact sets. As a consequence, they do not self-intersect. While the local influence is not clearly visible, this implies that long level lines are much more regular than random walks. This logically leads to an overdetection of long level lines because the independence assumption is strongly violated at very long range. The solution we propose is to stick to Helmholtz principle: "no detection in white noise." Thus we have to learn the regularity of level lines in white noise, and use this as the a priori distribution.

5.1. *Definition of Local Regularity*

Let $l_0 > 0$ be a fixed positive value. Let C be a rectifiable planar curve, parameterized by its length. Let $x = C(s_0) \in C$. With no loss of generality, we assume that $s_0 = 0$.

Definition 4. We call regularity of C at x (at scale l_0) the quantity

$$R_{l_0}(x) = \frac{\max(|x - C(-l_0)|, |x - C(l_0)|)}{l_0}. \quad (10)$$

Of course, this definition really makes sense if the length of C is larger than $2l_0$. This definition of regularity (see Fig. 13) is related to the Hausdorff dimension of C around x. First, $R_{l_0}(x) \leq 1$, with equality

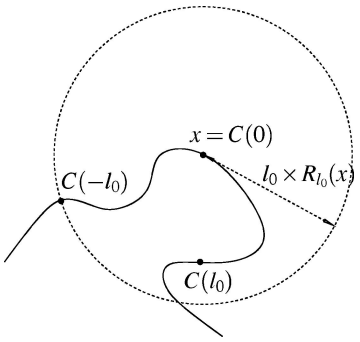

Figure 13. Regularity definition. The regularity at x is obtained by comparing the radius of the circle with l_0. The radius is equal to l_0 if and only if the curve is a straight line. If the curve has a large curvature, the radius will be small compared to l_0.

if and only if either $C((-l_0, 0))$ or $C((0, l_0))$ is a line segment. On the contrary, if $R_{l_0}(x)$ is small, then the curve is highly curved around x.

We can also interpret $R_{l_0}(x)$ as a function of the local curvature. Indeed, if C is a circle with large enough radius ρ, then

$$R_{l_0}(x) = \operatorname{sinc}\left(\frac{l_0}{2\rho}\right), \quad \text{where } \operatorname{sinc} x = \frac{\sin x}{x}. \quad (11)$$

This approximation is valid when l_0 is small compared to ρ. In this case, the regularity is a nonincreasing function of the curvature.

This definition is not purely local, but it is also less sensitive to noise compared to differential measures as the curvature. Let

$$\mathcal{H}_{l_0}(r) = P(x \in C,\ C \text{ is a white noise level line and } R_{l_0}(x) > r). \quad (12)$$

This distribution only depends on l_0 and can be empirically estimated. Of course, we learn it on level lines whose length is much larger than l_0 in order to avoid quantization effects.

Remark 3. As expected, the distribution $\mathcal{H}_{l_0}$ is very different in white noise and natural images. In natural images, the histogram of R_{l_0} has a peak at 1, corresponding to real objects boundaries (which often contain alignments). In some textured images, such as paintings, most edges are not real but subjective and this is clearly visible on the histogram of R_{l_0}. (See Fig. 14). The distribution also clearly depends on l_0.

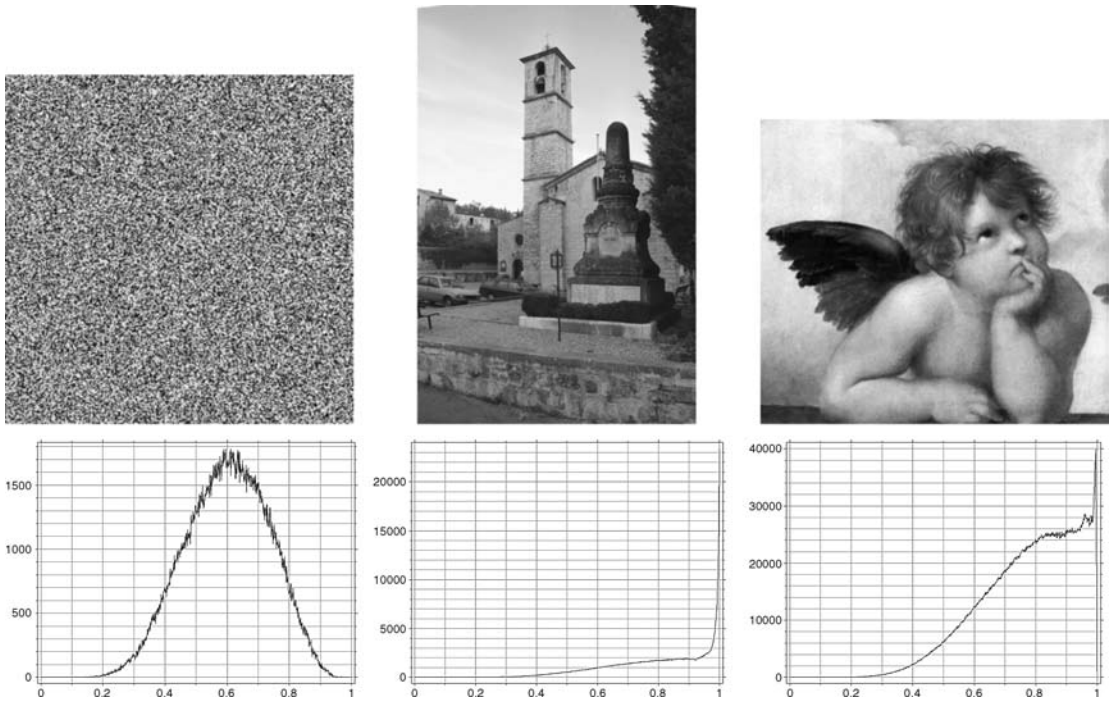

Figure 14. Regularity histograms. Upper row: a white noise image, a scanned photograph and a scanned photograph of a painting. Bottom row: the three regularity histograms for $l_0 = 10$. Since its histogram vanishes near 1, white noise does not contain any alignments or smooth curves, as foreseen. Nearly all natural images (containing true edges) have a regularity histogram like the second one. The third image contains mostly subjective edges, as it is composed of painted strokes. As a consequence, the regularity histogram is much less concentrated around 1 as for "natural" images. If we now unzoom the three images (with an adequate smoothing before downsampling), then the first histogram remains unchanged (scale invariance), while the other two have regularity histograms like the second one. Indeed, after unzooming, most textures and small scale features disappear, and small gaps get filled.

When l_0 grows, the histogram mode moves to lower values. However, we obtain the same qualitative behavior as above. In Appendix A, we use these distributions to compute the Hausdorff dimension of white noise level lines. We then quantitatively check that they are much more smooth than (self-intersecting) isotropic random walks.

Again, the choice of l_0 is a natural question. Of course l_0 should be larger than Nyquist distance. It should not be too large either. In experiments we have chosen $l_0 = 10$. But, since NFAs are additive, we may also choose several reasonable values of l_0 (say $l_0 = 5$, 10, 20) and multiply the NFAs by the number of l_0. In practice, changing l_0 influences the number of samples and best NFAs are attained for small l_0.

5.2. *Meaningful Contrasted and Smooth Boundaries*

Now that we have a background model of regularity, we use it to detect regular curves a contrario. It is natural to assume, in the background model, that contrast and regularity are independent. Thus

$$\begin{aligned} &P(C \text{ is contrasted and smooth}) \\ &\quad = P(C \text{ is smooth}) \times P(C \text{ is contrasted}). \end{aligned}$$

Definition 5. Let C be a level line. Let

$$\nu = \min\{|Du(x)|,\ x \in C\}, \tag{13}$$

$$\rho = \min\{|R_l(x)|,\ x \in C\}, \tag{14}$$

be respectively the minimal quantized contrast and regularity along C. Let

$$NFA_{\mathrm{cs}}(C) = N_{ll} H_c(\nu)^{l/2} \left(\mathcal{H}_{l_0}(\rho)\right)^{l/2l_0}. \tag{15}$$

We say that C is a ε-meaningful smooth boundary if $NFA_{cs}(C) < \varepsilon$.

The number of false alarms is the product of number of level lines and the probability that the contrast and the

regularity are simultaneously larger than the observed values along a curve with prescribed length taken in the background model. The probability is computed in the a contrario model, where contrast and regularity are independent and local observations are mutually independent.

As above, this search can be recursively performed by computing local histograms of the gradient.

In experiments, detection results are qualitatively equivalent with or without regularity. On the other hand, NFA may decrease a lot for smooth boundaries. Even though the detection is not changed in one single image, it is still interesting to decrease the NFA as much as possible. Indeed, we may want to detect boundaries not in a single image but in a database (for instance in shape recognition applications). We can consider than any database has a size much less than 10^{15}. Thus, curves with a NFA lower than 10^{-15} in a single image can also be considered as universally meaningful, since they will be detected in any database.

5.3. *Comparison with Active Contours*

Active contours is one of the most popular techniques of boundary detection. The first works of Kass, Witkin and Terzopoulos [17] have been improved and generalized by many authors. Recent models are more intrinsic, can be expressed implicitly (which ease the possible topological changes of the active contours) and can use image statistics [7, 34]. In this section, we do not focus on any particular active contour model, but try to compare a generic model with meaningful boundaries. Such a comparison has already been made by Desolneux, Moisan and Morel [12] for meaningful boundaries. Even though these boundaries are only contrast-based, they show that they are very close to active contours in general and particularly to the model of Kimmel and Bruckstein [18]. Since in this paper we have also introduced a regularity criterion, comparison is even more adequate.

Let us briefly give a generic active contour model: it is a curve that fits shape contours (hence contrast should be large along the contour) and which is also as smooth as possible. The problem usually assumes a variational formulation. An optimal curve minimizes an energy of the type

$$E(C) = \int_C g(|Du(C(s))| + \lambda h(\mathrm{curv}(C(s)))\, ds, \quad (16)$$

where Du is the gradient of a given grey-level image, g is a nonincreasing function, $\mathrm{curv}(C(s))$ is the curvature of C at point $C(s)$, h is a nondecreasing function and s is the arc-length. The optimal curve is a trade-off between the external energy depending on the image gradient, and the internal energy depending on the curve itself only. Such a model can accurately give the position of the contour. However, it has several drawbacks:

- The model assumes that there is a contour: it cannot be used as a detection algorithm. This also explains why active contours are also introduced in Bayesian models, where the real question is: knowing that one object is present, what is the best candidate?
- The initialization is crucial.
- The optimal balance parameter λ (which, for homogeneity reasons, can also be viewed as a scale parameter) is unknown and depends on the image. It has a strong influence on the result.

If we now only consider the homogeneity of the different energy terms, we have to minimize a potential of the form $Lg(|Du|) + \lambda Lh(\mathrm{curv}C)$, L being the length of the curve. Let us now consider the meaningful smooth boundary model. A meaningful curve has a small probability to occur in the a contrario model. Our regularity measure is a non increasing function of the curvature (see (11)). Thus, for a meaningful curve, the quantity

$$\left(H_c(|Du|)^{L/2}\mathcal{H}_{l_0}\left(R_{l_0}(C)\right)\right)^{L/2l_0}$$

is small. Let us now take the logarithm of this expression. We obtain an expression of the type

$$L(E_{\mathrm{ext}}(|Du|) + E_{\mathrm{int}}(\mathrm{curv}\ C)),$$

where E_{ext} is a non increasing function of $|Du|$, and E_{int} is a non decreasing function of the curvature. The model is qualitatively alike a snake model. Nevertheless, there are three major differences:

1. There is a quantitative criterion to decide if the curve has to be detected. Contrarily to snakes algorithm, meaningful boundaries detection is *not* a minimization algorithm. It is well known in active contours model that the value of the energy of the minimizer has no interpretation. All that we can say is that a candidate is better than another one. Our model gives a meaning to the energy-like term. Thus, there is no need for a minimization since we can give

thresholds under which a candidate has to be detected.
2. Meaningful boundaries are level lines. Thus, no initialization by hand is needed.
3. We do not have to fix the weight functions g and h as well as the scale parameter λ.

5.4. *Experiments on Smooth Meaningful Boundaries*

In general, adding a regularity criterion does not qualitatively change the result. This is in conformity with the observation of Desolneux et al. [12]. Remark also that adding the regularity criterion does not eliminate irregular level lines that were already detected thanks to contrast. Indeed,

$$NFA_{\mathrm{cs}}(C) \leq N_{ll}\, H_c(\nu)^{l/2},$$

(with the same notations as in Definition 5) since $\mathcal{H}_r(\rho) \leq 1$. We can only detect more lines, which is what we want: check whether or not we had misdetections because regularity was not taken into account. Of course, the NFA of smooth boundaries decrease a lot (about 10^{-15}), and this can modify maximal meaningful boundaries. As it was already observed in [4], contrast and regularity are often very redundant, and this explains why the same curves are detected.

Figure 15, (INRIA desk) is very geometrical and shows the redundancy between contrast and regularity. Since adding a regularity criterion does not change the

Figure 15. Regularity detectability. The original image is the left most. In the middle, we display the 204 detected contrasted smooth boundaries as defined in Definition 5. On the right, the 96 smooth boundaries, with no contrast information, defined in (17). All the main boundaries are already present. Of course, contrast may be the main cause of small NFA, since regularity acts at larger scales. For instance, the window panes have NFA about 10^{-150} with contrast and 10^{-15} with regularity only (which still make them detectable in any image database). The desk on the bottom right has a NFA equal to 10^{-60} with contrast and 10^{-20} with regularity, which is already very small.

Figure 16. Influence of regularity. On the left, the original texture contains a lot of elongated structure. Because the texture shows large contrast variations, meaningfulness is a very strict criterion and contrasted meaningful boundaries miss many details (middle). In this case, local regularity is important and smooth and contrasted boundaries allow to retrieve missing lines.

results, we could believe that our regularity definition is just bad and does not bring anything. This is not so. Indeed, we can also define NFA for smooth boundaries, with no care of contrast, as

$$NFA_{\text{reg}}(C) = N_{ll}(\mathcal{H}_{l_0}(\rho))^{l/2l_0}. \tag{17}$$

We retrieve most edges in the desk image with this definition. The conclusions of these experiments are the following: for natural images, there is a strong redundancy between regularity and contrast. Pieces of objects boundaries coincide with pieces of level lines, and they can be detected either by regularity or contrast, or both.

In Fig. 16, locally straight structures are also contrasted but the gradient distribution exhibits large values (since the texture variations are important). This explains why contrasted meaningful boundaries lose many lines. In this case, our local regularity criterion allows to characterize this elongated structures.

6. Conclusion

In this paper, we brought a contribution to Desolneux, Moisan and Morel's theory of meaningful boundaries. First, we gave a mathematical interpretation of the model. Basically, it means that a meaningful boundary cannot be generated only by noise. This implies that a meaningful boundary may contain some spurious parts. We proposed an algorithm to remove them. We also proposed a multiscale setting to the theory. As a result, detection is less sensitive to noise, in particular quantization noise. We also presented a method automatically to handle local contrast variations, and that do not only use a global measure of contrast. This is very useful for our purpose (shape matching) since it removes texture that usually does not yield stable shape elements. Finally, we discussed the importance of regularity in detection. Our conclusion is that it makes detection more robust, but in natural images, curves that are smooth but not contrasted are empirically quite seldom.

Experiments show that this model allows to extract a large number of shape elements from natural images. We cannot pretend to directly extract shapes of images since we believe that many contours are subjective and configurations of the type of Kanizsa's triangle often appear at lower degree. For such contours, all local methods are doomed to fail. However, for practical shape matching by shape elements comparison [21,31,32], the MB model with local contrast and cleaning-up automatically eliminates most edges due to texture or small illumination gradient. For our purpose, it is the best compromise between the compactness and the completeness of shape elements dictionaries in natural images.

Appendix A: Numerical Estimation of the Hausdorff Dimension of a Curve

In order to compute the Hausdorff dimension of identically distributed random curves from the histogram of regularity, we proceed as follows. Let C be a curve.

Definition 6. The Hausdorff measure of dimension α is defined by

$$\lim_{\delta \to 0} \inf_{(B_i)\delta-\text{covering}} \sum_i |B_i|^\alpha,$$

where the B_i form a covering of C and $|B_i|$ is the diameter of B_i. The family (B_i) is a δ-covering of C if $C \subset \cup_i B_i$ and for all i, $|B_i| < \delta$.

The problem to estimate this quantity is that it makes no sense to let $\delta \to 0$ for digital curves. Indeed, even for white noise, the precision is bounded from below by Nyquist distance. We assume that the curve is self-similar. This allows to examine it at larger and larger scales, instead of letting δ go to 0. Let us cut a curve with length $L = 2Nl$ in N chunks of length $2l$. We measure the regularity $R_l(i)$ at the middle point x_i of each piece. The balls with radius $R_l l$ nearly form a covering of C. It is not a covering because the endpoint of the curve chunk may not be the most remote point from the center (see (10)). Nevertheless, we approximate the measure of C by

$$\mathcal{H}^\alpha(C) \simeq \sum_{i=1}^{N} (2l R_l)^\alpha \simeq 2^{\alpha-1} L l^{\alpha-1} \bar{R}_l^\alpha,$$

where $\bar{R}_l$ is the mean regularity along C. Let us now consider the curve λC with $\lambda > 1$. We can make the same procedure as above with chunks whose length is equal to $2\lambda l$. Thus we evaluate the measure of λC by

$$\mathcal{H}^\alpha(\lambda C) \simeq 2^{\alpha-1} \lambda L (\lambda l)^{\alpha-1} \overline{R_{\lambda l}}^\alpha.$$

But, if we now use pieces of curves of length $2l$, we also obtain

$$\mathcal{H}^\alpha(\lambda C) \simeq 2^{\alpha-1}\lambda L l^{\alpha-1}\bar{R}_l^\alpha.$$

Thus

$$\lambda^\alpha \bar{R}_{\lambda l}^\alpha = \lambda \bar{R}_l^\alpha,$$

yielding

$$\log(\bar{R}_{\lambda l}) = \left(\frac{1}{\alpha} - 1\right)\log\lambda + \log\bar{R}_l. \qquad (18)$$

We can evaluate α by examining the histograms of R_l as a function of l.

For random walks with independent increments, we find $\alpha = 2.02$, whereas the true dimension is 2. For level lines in white noise, we find $\alpha = 1.78$. As expected, the level lines of a white noise image are more regular than random walks.

Acknowledgments

We thank Jean-Michel Morel, Agnès Desolneux and Lionel Moisan for all discussions and advice.

References

1. A. Alexandrov and Y. Reshetnyak, *General Theory of Irregular Curves*, vol. 29 of *Mathematics and Its Applications: Soviet Series*. Kluwer Academic Publishers, 1989.
2. E.J. Breen and R. Jones, "Attribute openings, thinnings and granulometries," *Computer Vision and Image Understanding*, Vol. 64, No. 3, pp. 377–389, 1996.
3. J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679–698, 1986.
4. F. Cao, "Good continuation in digital images," in *Proceeding of ICCV 03*, Nice, Vol. 1, 2003, pp. 440–447.
5. V. Caselles, B. Coll, and J.M. Morel, "A kanizsa program," in *Progress in Nonlinear Differential Equations and their Applications*, Vol. 25, 1996, pp. 35–55.
6. V. Caselles, T. Coll, and J.M. Morel, "Topographic maps and local contrast changes in natural images," *International Journal of Computer Vision*, Vol. 33, No. 1, pp. 5–27, 1999.
7. V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, Vol. 22, No. 1, pp. 61–79, 1997.
8. T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, Vol. 10, No. 2, pp. 266–277, 2001.
9. A. Desolneux, L. Moisan, and J.M. Morel, "Meaningful alignments,' *International Journal of Computer Vision*, Vol. 40, No. 1, pp. 7–23, 2000.
10. A. Desolneux, L. Moisan, and J.M. Morel, "Edge detection by Helmholtz principle," *Journal of Mathematical Imaging and Vision*, Vol. 14, No. 3, pp. 271–284, 2001.
11. A. Desolneux, L. Moisan, and J.M. Morel, "A grouping principle and four applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 4, pp. 508–513, 2003.
12. A. Desolneux, L. Moisan, and J.M. Morel, "Variational snake theory," in *Geometric Level Set Methods in Imaging, Vision, and Graphics*, S. Osher and N. Paragios (Eds.), Springer Verlag, 2003.
13. L.C. Evans and R.Gariepy, *Measure Theory and Fine Properties of Functions*, CRC Press. Ann Harbor, 1992.
14. P. Felzenszwalb and D. Huttenlocher, "Image segmentation using local variation," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 1998, pp. 98–104.
15. R. Haralick, "Digital step edges from zero crossing of second directional derivatives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 58–68, 1984.
16. G. Kanizsa, *La Grammaire du Voir*. Diderot, 1996. Original title: *Grammatica del vedere*. French translation from Italian.
17. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, Vol. 1, pp. 321–331, 1987.
18. R. Kimmel and A.M. Bruckstein, "On regularized laplacian zero crossings and other optimal edge integrators," *International Journal of Computer Vision*, Vol. 53, No. 3, pp. 225–243, 2003.
19. J.J. Koenderink, "The structure of images," *Biol. Cybern.*, Vol. 50, pp. 363–370, 1984.
20. T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, Vol. 30, No. 2, pp. 77–116, 1998.
21. J.L. Lisani, L. Moisan, P. Monasse, and J.M. Morel, "On the theory of planar shape," *SIAM Multiscale Modeling and Simulation*, Vol. 1, No. 1, pp. 1–24, 2003.
22. J.L. Lisani, P. Monasse, and L. Rudin, "Fast shape extraction and application," Preprint 16, CMLA, ENS-Cachan. Available at http://www.cmla.ens-cachan.fr, 2001.
23. S. Mallat, *A Wavelet Tour in Signal Processing*, 2nd edition. Academic Press, 1999.
24. D. Marr, *Vision*. W.H. and Co: N.York, 1982.
25. D. Marr and E. Hildreth, "Theory of edge detection," in *Proceeding of Royal Society of London*, Vol. 207, 1980, pp. 187–207.
26. G. Matheron, *Random Sets and Integral Geometry*, John Wiley: N.Y., 1975.
27. F. Meyer and P. Maragos, "Nonlinear scale-space representation with morphological levelings," *J. of Visual Comm. and Image Representation*, Vol. 11, pp. 245–265, 2000.
28. P. Monasse, "Morphological representation of digital images and application to registration," Ph.D. thesis, Université Paris IX Dauphine, 2000.
29. P. Monasse and F. Guichard, "Fast computation of a contrast invariant representation," *IEEE Transactions on Image Processing*, Vol. 9, No. 5, pp. 860–872, 2000.
30. D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Communication on Pure and Applied Mathematics*, Vol. XLII, No. 4, 1989.

31. P. Musé, F. Sur, and J.M. Morel, "Sur les seuils de reconnaissance de formes," *Traitement du Signal*, Vol. 19, Nos. 5/6, 2003.
32. P. Musé, F. Sur, F. Cao, and Y. Gousseau, "Unsupervised thresholds for shape matching," in *IEEE Int. Conf. on Image Processing, ICIP*, 2003.
33. P. Musé, F. Sur, F. Cao, Y. Gousseau, and J.M. Morel, "Accurate estimates of false alarm number in shape recognition," Technical Report 5086, INRIA, 2004. submitted.
34. N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *International Journal of Computer Vision*, Vol. 46, No. 3, pp. 223–247, 2002.
35. E. Le Pennec and S. Mallat, "Sparse geometrical image approximation with bandelets," accepted for publication in IEEE Trans. on Image Proc., 2003.
36. P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Transactions on Image Processing*, Vol. 9, No. 4, pp. 561–576, 2000.
37. P. Salembier and J. Serra, "Flat zones filtering, connected operators, and filters by reconstruction,' *IEEE Transactions on Image Processing*, Vol. 4, No. 8, pp. 1153–1160, 1995.
38. J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, 1982.
39. J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, pp. 888–905, 2000.
40. J.L. Starck, E.J. Candès, and D.L. Donoho, "The Curvelet Transform for Image Denoising," *IEEE Transactions on Image Processing*, Vol. 11, pp. 670–684, 2000.
41. M. Wertheimer, "Untersuchungen zur Lehre der Gestalt, II," *Psychologische Forschung*, Vol. 4, No. 301–350, 1923.
42. A.P. Witkin, "Scale space filtering," in *Proc. of IJCAI, Karlsruhe*, 1983, pp. 1019–1021.
43. S.C. Zhu, "Embedding gestalt laws in markov random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 11, pp. 1170–1187, 1999.

Frédéric Cao graduated from the École Polytechnique (France), and obtained his Ph.D. in 2000 at École Normale Supérieure de Cachan. From 2001, he has been with Irisa (Inria Rennes). His research interests are still image and video analysis by geometrical methods, including partial differential equations and statistical methods.

Pablo Musé was born in Montevideo, Uruguay, in 1975. He received the Electrical Engineer degree from the Universidad de la Repblica, Uruguay, in 1999, and the DEA (M.Sc.) in Mathematics, Vision and Learning from the École Normale Supérieure de Cachan, France, in 2001. He has obtained a Ph.D. in Applied Mathematics in 2004 in ENS Cachan, where he currently has a researcher position.

Frédéric Sur was born in 1976. He studied mathematics at École Normale Supérieure de Cachan from 1997 to 2001 and received the DEA Mathématiques, Vision, Apprentissage. He obtained his Ph.D. thesis in 2004 and now he has a post doc position in LORIA/CNRS.