

# A *contrario* edge detection with edgelets

Nicolas Widynski <sup>#1</sup>, Max Mignotte <sup>#2</sup>

<sup>#</sup> *Department of Computer Science and Operations Research (DIRO), University of Montreal  
C.P. 6128, succ. Centre-Ville, , Montreal (Quebec), H3C 3J7, Canada*

<sup>1</sup>widynski@iro.umontreal.ca

<sup>2</sup>mignotte@iro.umontreal.ca

**Abstract**—Edge detection remains an active problem in the image processing community, because of the high complexity of natural images. In the last decade, Desolneux et al. proposed a novel detection approach, parameter free, based on the Helmholtz principle. Applied to the edge detection field, this means that observing a true edge in random and independent conditions is very unlikely, and then considered as meaningful. However, overdetection may occur, partly due to the use of a single pixel-wise feature. In this paper, we propose to introduce higher level information in the *a contrario* framework, by computing several features along a set of connected pixels (an edgelet). Among the features, we introduce a shape prior, learned on a database. We propose to estimate online the *a contrario* distributions of the two other features, namely the gradient and the texture, by a Monte-Carlo simulation approach. Experiments show that our method improves the original one, by decreasing the number of non relevant edges while preserving the true ones.

## I. INTRODUCTION

Edge detection is one of the oldest image processing problematic. Due to high variability between natural images and the large quantity of information they contain (resolution, color), this problem has not been solved so far.

Among classical methods, we can cite the Sobel filter, the Canny filter [1], and the Deriche filter [2]. The aim of last two was to propose an algorithm satisfying the well known Canny properties: good detection, good localisation, and minimal response. However, besides being sensitive to noise, these methods need supervised parameters tuning and are often too local to provide relevant detections.

More recently, A. Desolneux, L. Moisan, J.M. Morel proposed in [3] a formal modeling of the Helmholtz principle, which enunciates that relevant geometric structures have a very low probability to occur in a random context. The detected structures are declared as meaningful. The general methodology consists in defining the randomness to detect meaningful events, hence yielding the so-called *a contrario* method. Structures are defined by the laws of Gestalt Theory [4], underlining the importance of perceptual grouping for human visual perception. Desolneux et al. proposed to apply their statistical framework to detect alignments in images [3], clusters [5], and edges [6]. Since then, the *a contrario* framework have been used in various purposes, such as motion detection [7], shape recognition [8], object matching [9], and local features matching [10].

The first step in *a contrario* methods is to define a set of possibly meaningful events. In the context of edge detection,

two approaches have been proposed. In [6], [11], [12], candidate edges are extracted from the level lines of the image. Given a grey level image  $u : \Omega \rightarrow \mathbb{R}$ , level lines are defined as closed Jordan curves contained in the boundary of a level set with level  $\lambda$ ,

$$\mathcal{X}_\lambda = \{x \in \Omega, u(x) \leq \lambda\} \quad \text{and} \quad \mathcal{X}^\lambda = \{x \in \Omega, u(x) \geq \lambda\},$$

with  $\mathcal{X}_\lambda$  and  $\mathcal{X}^\lambda$  respectively the lower and upper sets, and  $\Omega$  the image domain. This representation contained several advantages. First, no information is lost from the image and this one can be reconstructed from the family of lower (or upper) sets [13], [11]. Moreover, some level sets boundaries locally coincide with image contours [6]. Hence, filtering the image using level sets does not alter contours. These approaches use the fact that a long and contrasted level line is very unlikely in the randomness. In [14], [15], authors proposed to use candidates from an algorithm producing oversegmented regions. These regions are then merged upon the *a contrario* hypothesis that two adjacent regions have, in noise, a very low probability to present identical characteristics. However, in an edge detection problematic, using oversegmented regions provides to small edges to guarantee a high level relevancy.

The *a contrario* detection is related to the probability of an event under the random hypothesis. Defining this distribution is of prime interest, since it directly impacts on the accuracy of the method. In the previous mentioned approaches, the edge *a contrario* distribution is defined by pixel-wise based on saliency characteristics. Moreover, mainly because of the difficulty of defining a multi-features closed form *a contrario* distribution, the detection often use just one local feature, namely the gradient of the image. In this paper, and like in the Desolneux et al. framework, edge candidates are extracted from level lines. We propose several contributions. First, we introduce prior information in the *a contrario* model, in order to detect part of curves whose shapes are relevant according to a shape database. Then, by estimating the *a contrario* distributions using Monte Carlo simulations, we model higher level information, by reasoning on sets of connected pixels, that we call edgelets. This allows defining prior, gradient and textural features. These features are then combined in the *a contrario* framework.

This paper is organized as follows. The *a contrario* framework for edge detection as proposed in [6] is first presented in Section II. The proposed approach is described in Section III.

We show experimental results in Section IV, before concluding in Section V.

## II. $\epsilon$ -MEANINGFUL EDGES

We recall here the  $\epsilon$ -meaningful edge detector proposed in [6]. A candidate edge is declared as meaningful if its length is too long and its contrast is too strong to have been likely generated by random noise. The quantity  $\epsilon$  holds that the expected number of meaningful events happening by chance is, on the average, less than  $\epsilon$ .

### A. Defining the number of false alarm

Let  $u : \Omega \rightarrow \mathbb{R}$  be a grey level image. Since the image gradient norm is different from one image to another, the meaningfulness of a candidate should depend on its distribution. Then, we first define the distribution of the norm gradient of  $u$ :

$$\forall \mu > 0, H(\mu) = P(X > \mu) = \frac{\#\{x \in \Omega, |\nabla u(x)| > \mu\}}{\#\{x \in \Omega, |\nabla u(x)| > 0\}} \quad (1)$$

with  $|\nabla u(x)|$  the norm of the gradient computed on point  $x$  with a  $2 \times 2$  neighborhood. A level line containing a critical point (i.e. whose gradient norm is null) is ignored in the detection process. Hence  $G$  is defined only for  $\mu > 0$ , and is so normalized by the number of points satisfying this condition.

Let  $E$  be a candidate edge and  $x_E^i$  be the  $i$ -th point of  $E$ . The candidate  $E$  is composed by  $l_E$  independent points, so according to the Nyquist distance, the distance between two consecutive points of  $E$  is greater than 2. We define  $PFA(E)$ , the probability of false alarm of the event  $E$  by:

$$PFA(E) = H\left(\min_{i=1..l_E} |\nabla u(x_E^i)|\right)^{l_E} \quad (2)$$

The  $PFA$  equation fits the intuition that observing  $l_E$  independent points is very unlikely in noise, as soon as  $l_E$  and the minimum of gradient along the curve  $E$  growth.

The originality of the *a contrario* approach consists in, instead of considering the probability of accepting of a false alarm, limiting its expected number, whose computation is easier thanks to the linear property of the expected value. Let  $C = \{E_1, \dots, E_{N_E}\}$  the set of  $N_E$  edge candidates. An edge candidate  $E_n$  is an  $\epsilon$ -meaningful edge if

$$NFA(E_n) = N_E \times PFA(E_n) < \epsilon \quad (3)$$

with  $NFA(E_n)$  the number of false alarm of  $E_n$ . The demonstration of this proposition can be found in [11], using the lemma that, for  $X$  a real random variable, the inequality  $P(P(X > \mu) < t) \leq t$  holds,  $\forall t \in [0, 1]$ .

### B. Discussion

As discussed in [6], [11], the detection method is considered parameter free, since it only depends on  $\epsilon$  whose influence is logarithmic (this is due to the exponent  $l_E$  in Eq. 2). In practice, we set  $\epsilon = 1$ .

In the Desolneux et al. framework, and in this paper, edge candidates are extracted from the level lines of the image.

This is done using the Fast Level Lines Transform (FLLT) algorithm proposed in [16]. Let  $L_i$  be a level line of length  $l_{L_i}$ , so the number of edge candidates it contains is equal to  $\sum_{i=1}^{l_{L_i}} i$ . Considering  $N_L$  level lines, the total number  $N_E$  of edge candidates over all level lines is:

$$N_E = \sum_{i=1}^{N_L} \frac{l_{L_i}(l_{L_i} + 1)}{2}$$

Desolneux et al. proposed a maximality principle in order to find optimal edges among the detected ones. An edge  $E_n$  is said maximal meaningful if its  $NFA$  is lower to the edges it includes, and strictly lower to the edges that include itself:

- $\forall E_m$   $\epsilon$ -meaningful,  $E_m \subseteq E_n \implies NFA(E_n) \leq NFA(E_m)$
- $\forall E_m$   $\epsilon$ -meaningful,  $E_n \subsetneq E_m \implies NFA(E_n) < NFA(E_m)$

This definition implies that two different maximal meaningful edges on the same level line cannot intersect.

Using the same gradient feature, Desolneux et al. also proposed a boundary detector. A detected boundary is an  $\epsilon$ -meaningful level line, and then it has the nice property to be closed. Experiments in [6] concluded that the two detectors give comparable results. A first idea would so be to propose a region based on feature, for example by supporting flat and large regions. Nevertheless, in natural images, it is very unlikely to find a level line representing a whole object, due to texture, compression and quantification. Along the level lines, objects appear as fragmented, since just some parts of some level lines fit well the contours (Figure 1). In this scenario, defining a region term might not be appropriate. This is the reason why we choose to focus on edge detection, which allows splitting level lines and hence only detecting the relevant parts. The main idea of our contribution is to compute the features on edgelets instead of pixels. This allows us to define higher level discriminative features, as we will see in the next section.



Fig. 1.

Example of a level line in the Valbonne church image.

## III. PROPOSED APPROACH

We propose several contributions to the *a contrario* edge detector. First, atomic elements, i.e. pixels in the Desolneux et al. framework, are replaced by connected pixels, called

edgelets. This allows us to handle higher level information features. The first one is an edgelet prior, automatically learned using a shape database, which is also a novelty. Moreover, we model a gradient and a textural features. This last one feature aims at minimizing response along contours surrounded by redundant information. The *a contrario* distributions need to be estimated using a Monte-Carlo procedure. But first, we propose to combine several features into the *a contrario* framework.

#### A. Combining several features

A legitimate idea would be to consider the joint distribution over  $F$  features,  $PFA(\mu_1, \dots, \mu_F) = P(X_1 > \mu_1, \dots, X_F > \mu_F)$ . In this situation, the  $\epsilon$ -meaningfulness of the detector is not guaranteed, since proving that  $P(P(X_1 > \mu_1, \dots, X_F > \mu_F) < \epsilon) \leq \epsilon$  is not straightforward.

An easy way to use several features in a *a contrario* approach consists in using a fusion operator  $\Xi$ , such that  $\Xi(PFA_1(E), \dots, PFA_F(E)) \geq \max_{f=1, \dots, F} PFA_f(E)$ . Then, an edgelet  $E$  is said  $\epsilon$ -meaningful if it follows the inequality:

$$NFA(E) = N_E \times \Xi(PFA_1(E), \dots, PFA_F(E)) < \epsilon \quad (4)$$

The proof of this proposition is straightforward. Let  $NFA_j(E)$  the number of false alarm of an event  $E$  according to the feature  $j$ , and  $N_E$  the number of events. Since  $\forall j, P(\max_{i=1, \dots, F} NFA_i(E) < \epsilon) \leq P(NFA_j(E) < \epsilon)$ , then  $P(\max_{i=1, \dots, F} NFA_i(E) < \epsilon) \leq \epsilon/N_E$ , and hence  $P(NFA(E) < \epsilon) \leq \epsilon/N_E$ , with  $NFA(E) = N_E \times \Xi(PFA_1(E), \dots, PFA_F(E))$  [11]. Following that, it directly comes that the expected number of  $\epsilon$ -meaningful edges is less or equal than  $\epsilon$ . Notice by adapting the threshold of meaningfulness, less restrictive operators may also be used, such as the min, as it has been proposed in [17]. The choice of the fusion operator depends on the features. A min operator may be adapted to complementary features, whereas with  $\Xi = \max$ , less candidates become meaningful, which is more convenient for competing features. In this article, we consider that all the features detect the good candidates, but produce different false alarms. In this case, combining several features with a max operator increases the sensitivity of the detector. We set  $\Xi = \max$  in all the considered experiments.

#### B. Edgelet features

Let  $E = (w_E^1, \dots, w_E^{k_E})$  be an edge of composed by  $k_E$  edgelets, and  $w_E^i \in \Gamma$  be the  $i$ -th edgelet, defined by a set of connected pixels  $w_E^i = (w_E^{i,1}, \dots, w_E^{i,M})$ , with  $w_E^{i,j} \in \Omega$ .  $M$  is the edgelet length, and is a parameter of our method. To guarantee the independence assumption, consecutive edgelets are separated by a distance of 2. Under the *a contrario* hypothesis, an edgelet  $w$  follows  $P(w)$ , with  $P(\cdot)$  a uniform discrete distribution on  $\Gamma$ .  $\Gamma$  is designed such that it excludes self-crossing edgelets. We propose to model three features: an edgelet prior, whose *PFA* distribution is estimated offline by a Monte-Carlo simulation approach, and a gradient and a textural features, estimated online, since they depend on the image.

1) *Prior*: The first feature is a prior on edgelet shapes. Intuitively, one could model it analytically by supporting smooth curves [11]. However, this may not be robust to rough contours, or corners. We propose to learn the *a contrario* prior distribution of an edgelet using a shape database. The goal is to learn a distribution  $P(w|\{y^1, \dots, y^O\})$ , with  $\{y^1, \dots, y^O\}$  a set of  $O$  observed edgelets extracted from a database, and  $w$  an edgelet from the *a contrario* hypothesis. The likelihood of the edgelets from the database  $\{y^1, \dots, y^O\}$  conditioned by an event edgelet  $w$  from the hypothesis *a contrario* is computed using:

$$\begin{aligned} P_{i_i}^1(w) &= P(\{y^1, \dots, y^O\}|w) \\ &\propto \frac{1}{O} \sum_{o=1}^O [d_F(y^o, w)]^{-1} \end{aligned} \quad (5)$$

with  $y^o$  and  $w$  two centered edgelets of length  $M$ ,  $[d_F(y^o, w)]^{-1}$  the matching score, and  $d_F$  the discrete Fréchet distance [18]:

$$d_F(y^o, w) = \inf_{\alpha, \beta} \max_{t \in [0,1]} d(y^o(\alpha(t)), w(\beta(t))) \quad (6)$$

with  $d$  the distance function in  $\Omega$ , here the Euclidean one. The Fréchet distance depends both on the location and the order of the points along the curves, which makes it robust. However, it is quite computationally expensive and one could prefer the Hausdorff or the Euclidean distance.

The offline estimation procedure of the posterior distribution  $P(w|\{y^1, \dots, y^O\})$  is presented in Algorithm 1. The *PFA<sub>P</sub>* of the prior feature is then computed according to:

$$PFA_P(E) = \text{Prior} \left( \min_{i=1..k_E} P_{i_i}^1(w_E^i) \right)^{k_E} \quad (7)$$

with the distribution *Prior* such that

$$\text{Prior}(\mu) = \frac{1}{O} \#\{z^{(u)}, P_{i_i}^1(z^{(u)}) > \mu\} \quad (8)$$

with  $z^{(u)} \sim P(w|\{y^1, \dots, y^O\})$  according to the Algorithm 1. Observed edgelets are obtained from the Berkeley Segmentation DataSet [19]. In each human segmented images from the learning database, we collected a fix number of edgelets by a random selection.

Introducing prior information in a *a contrario* detector may seem conflicting with the spirit of the method. Indeed, a *a contrario* method models the randomness to underline Gestalt structures. However, modeling the noise instead of a reference model is not the only one advantage of using the *a contrario* framework: it also gives a mathematical structure to detect meaningful events without any parameter.

Most of *a contrario* works model *a contrario* distributions in a closed form. Among the exceptions, we can cite for example [10], in which the *PFA* distribution is learned from a database. In [14], random simulations are used to estimate a joint *PFA* distribution. For the two following features, we propose to learn their *PFA* online, using the weighted sampling procedure described in Algorithm 1. This methodology is different from the ones proposed before, and

---

**Algorithm 1:** Weighed sampling Monte-Carlo procedure

---

**Input:** Likelihood distribution  $P_{l_i}(w)$ ,  $\Gamma$ **Output:** Empiric posterior distribution  $P_U(w)$ **begin****for**  $u = 1, \dots, U$  **do**-  $z^{(u)} \sim \mathcal{U}(\Gamma)$ .- Compute  $q^{(u)} \propto P_{l_i}(z^{(u)})$  such that $\sum_{v=1}^U q^{(v)} = 1$ .**return**  $P_U(w) = \sum_{u=1}^U q^{(u)} \delta_{z^{(u)}}(w)$ 

---

is a part of our contribution. This allows us to model more complex features than in the closed form case.

2) *Gradient*: The second feature uses gradient norm of the image  $|\nabla u|$ . Unlike the Desolneux et al. framework [6], gradient is computed along an edgelet, which allows more flexibility in its formulation. The probability of false alarm is:

$$PFA_G(E) = \text{Gradient} \left( \min_{i=1..k_E} |g(w_E^i)| \right)^{k_E} \quad (9)$$

with  $g(w)$  the gradient computation along the edgelet  $w$ :

$$g(w_E^i) = \Phi \left( \left| \nabla u \left( w_E^{i,1} \right) \right|, \dots, \left| \nabla u \left( w_E^{i,M} \right) \right| \right) \quad (10)$$

The flexibility comes from the fusion function  $\Phi$ . One can set  $\Phi = \min$ , or  $\Phi = \max$  or again a ponderate mean  $\Phi(v^1, \dots, v^M) = \sum_{j=1}^M W(j) v^j$ , with  $W : \{1, \dots, M\} \rightarrow [0, 1]$  a weighting function. In our experiments, we set  $W(j) = 1/M, \forall j$ . This is less drastic than taking the minimum value over the whole edge (Section II). Since the gradient depends on the image, the distribution *Gradient* needs to be estimated online. We use the same estimation procedure as the prior one (Algorithm 1), with  $g$  the likelihood distribution:

$$\text{Gradient}(\mu) = \frac{1}{U} \#\{z^{(u)}, g(z^{(u)}) > \mu\} \quad (11)$$

3) *Textural*: The last feature aims at minimizing response on texture locations, while conserving high response values on object contours. The  $PFA_T$  for the textural feature is:

$$PFA_T(E) = \text{Textural} \left( \min_{i=1..k_E} t(w_E^i) \right)^{k_E} \quad (12)$$

Now, for a point  $w_E^{i,j}$  of an edgelet  $w_E^i$ , we consider its normal segment. The two sides of the normal segment of a point  $w_E^{i,j}$  are noted  $s_1(w_E^{i,j})$  and  $s_2(w_E^{i,j})$ . In a texture, the intuition is that value points along the first segment should not really differ from the other one. Let  $h[a] = \{h^r[a]\}_{r=1}^R$  be the histogram of a set of pixels  $a$ , and where  $r$  is the index of a bin of a histogram of length  $R$ . Distances between pairs of histograms along normal of the curves are combined to form the texture feature:

$$t(w_E^i) = \Psi \left( d_B(h[s_1(w_E^{i,1})], h[s_2(w_E^{i,1})]), \dots, d_B(h[s_1(w_E^{i,M})], h[s_2(w_E^{i,M})]) \right) \quad (13)$$

with  $\Psi$  the fusion operator. In our experiments, we set  $\Phi(v^1, \dots, v^M) = 1/M \sum_{j=1}^M v^j$ .  $d_B$  is the Bhattacharyya distance between two histograms:

$$d_B(h[a], h[b]) = \left[ 1 - \sum_{r=1}^R \sqrt{h^r[a] h^r[b]} \right]^{1/2} \quad (14)$$

Like for the gradient, the texture information depends on the image, then the distribution *Textural* is estimated online. Using the Algorithm 1, *Textural* is given by:

$$\text{Textural}(\mu) = \frac{1}{U} \#\{z^{(u)}, t(z^{(u)}) > \mu\} \quad (15)$$

For color, as for grey level images, we set in our experiments  $R = 125$ , the number of bins.

#### IV. EXPERIMENTS

The first serie of experiments concerns the results obtained by each feature, and by setting the length of edgelets to  $M = 7$  and  $M = 15$  (Figure 2). Ideally, each feature should detect the true contours, and provide different false alarms. Then, overdetection may not be a problem there are different enough to the ones provided by other features.

As we could imagine, prior feature (first column) detects mainly the smooth contours, and the effect is even more obvious when the length of the edgelet is large (Figures 2(d)). This also detects smoothes level lines in the sky, which come from low gradations and quantification. The second column presents the results obtained by the gradient feature. Compared to the classical approach (Figure 3), the feature detects more edges. This may be explained by the fact that computing the mean value along an edgelet smoothes the *a contrario* distribution and then makes the detection less radical. The last column is about the textural feature. As expected, stones are not detected, since information along the normal of an edgelet is redundant. We can observe that the detection is quite rough. This is also explained by the nature of the feature, which has a tendency to have to high response on curves parallel to the true contours.

Final results, obtained by combining the features, are presented in Figure 3, and compared to the classical approach by Desolneux et al. . Our approach detects less edges but keep the relevant ones. As we claimed before, this result is due to the combination of different competitive features, which provide the true contours, but different false alarms. Using a higher  $M$  seems to reduce the number of false alarm, but at the expense of a missing detection (trees at the right on the image, Figure 3(c)). Then, in the rest of our experiments, we set  $M = 7$  the length of edgelets.

Figure 4 presents various edge detection results obtained by our approach, on grey level and color images. We can observe that in images with smooth backgrounds (tools and telephone images), the algorithm overdetects the objects. This is called by Desolneux et al. the “blue sky effect” [6]. It happens when the background obtains very low responses to the features. Then, conditioned by the *a contrario* distributions, many parallel level lines become meaningful.

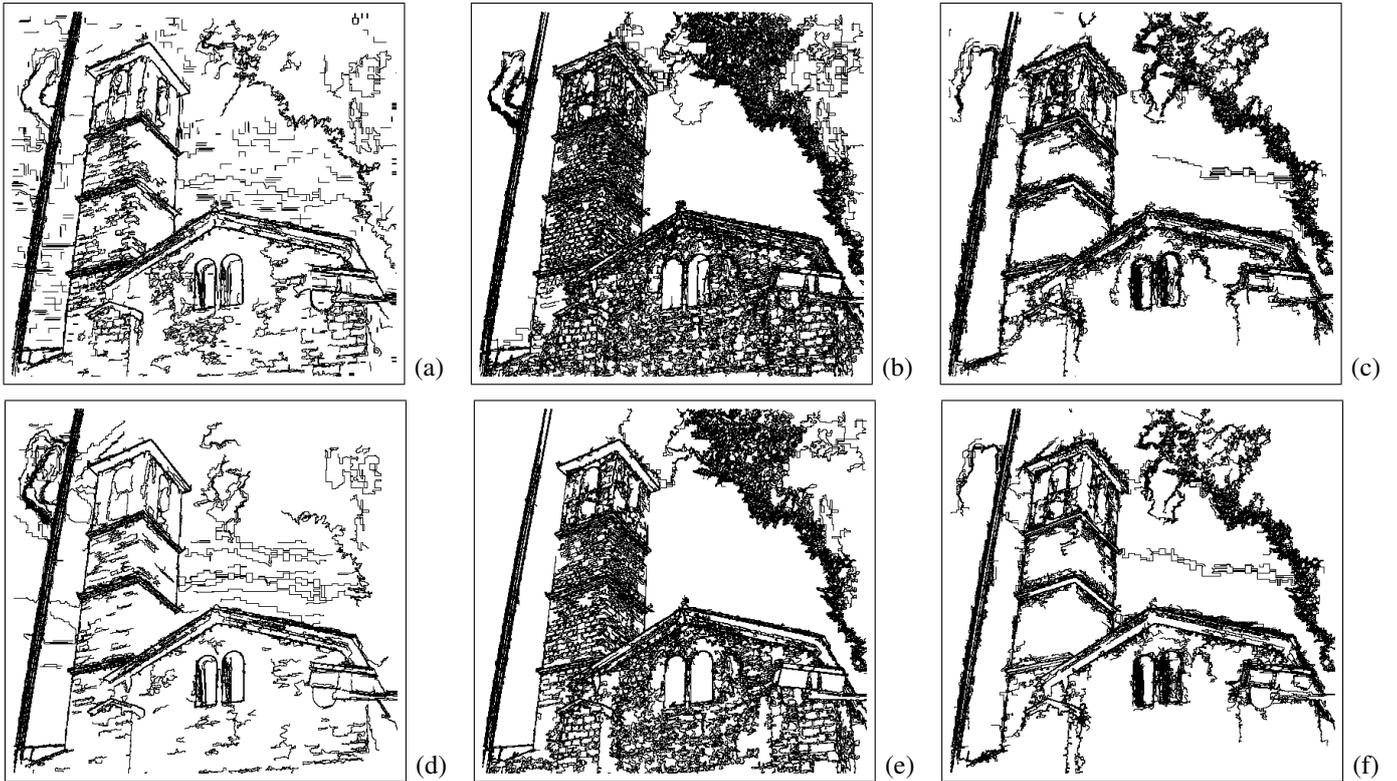


Fig. 2. Results obtained on Valbonne church image, using the features independently, and different length of edgelets. First row: length  $M = 7$ , second row: length  $M = 15$ . (a,d): prior, (b,e): gradient, and (c,f): textural.

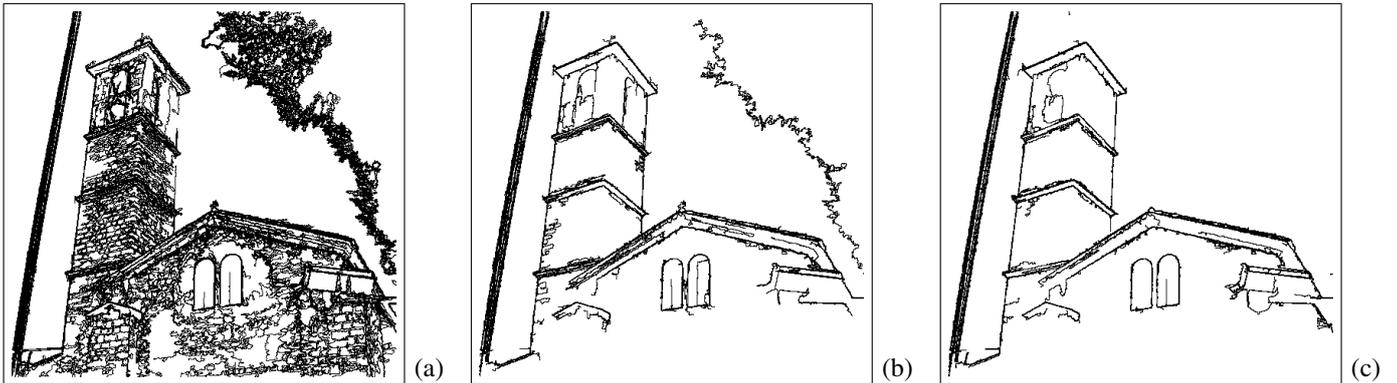


Fig. 3. Final detection results with (a) the classical approach [6], and our approach with a length of edgelets of (b)  $M = 7$ , and (c)  $M = 15$ .

## V. CONCLUSION

In this paper, we propose some improvements to the *a contrario* edge detector proposed by Desolneux et al. in [6]. Our approach aims at introducing higher level information features, by handling edgelets instead of single pixels. The first feature is an edgelet shape prior, learned using human segmentations over the Berkeley Segmentation DataSet [19]. The *a contrario* distributions of the two other features are the gradient and the texture, and are estimated online using a Monte-Carlo simulation procedure, which is another novelty of our approach. Experiments processed on classical images show that our approach reduces the number of false alarm, while keeping the most relevant object contours.

*A contrario* methods start with a set of possible candidates, and filter it to keep the most meaningful ones. As we saw, in the context of edge detection, using level lines have nice properties but possess the drawback to consider grey level images, which not be relevant for all images. A possible interesting field of investigation would be to propose more relevant edges, for example using more sophisticated preprocesses.

## REFERENCES

- [1] J. Canny, "A computational approach to edge detection," *Readings in computer vision: issues, problems, principles, and paradigms*, vol. 184, no. 87-116, p. 86, 1987.
- [2] R. Deriche, "Using canny's criteria to derive a recursively implemented optimal edge detector," *International journal of computer vision*, vol. 1, no. 2, pp. 167-187, 1987.



Fig. 4. Results of our approach obtained on various images.

- [3] A. Desolneux, L. Moisan, and J. Morel, "Meaningful alignments," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 7–23, 2000.
- [4] M. Wertheimer, "Untersuchungen zur Lehre von der Gestalt. II," *Psychological Research*, vol. 4, no. 1, pp. 301–350, 1923.
- [5] A. Desolneux, L. Moisan, and J. Morel, "A grouping principle and four applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 508–513, 2003.
- [6] —, "Edge detection by helmholtz principle," *Journal of Mathematical Imaging and Vision*, vol. 14, no. 3, pp. 271–284, 2001.
- [7] T. Veit, F. Cao, and P. Bouthemy, "An a contrario decision framework for region-based motion detection," *International journal of computer vision*, vol. 68, no. 2, pp. 163–178, 2006.
- [8] P. Musé, F. Sur, F. Cao, Y. Gousseau, and J. Morel, "An a contrario decision method for shape element recognition," *International Journal of Computer Vision*, vol. 69, no. 3, pp. 295–315, 2006.
- [9] N. Burrus, T. Bernard, and J. Jolion, "Bottom-up and top-down object matching using asynchronous agents and a contrario principles," *Computer Vision Systems*, pp. 343–352, 2008.
- [10] J. Rabin, J. Delon, and Y. Gousseau, "A statistical approach to the matching of local features," *SIAM Journal on Imaging Sciences*, vol. 2, p. 931, 2009.
- [11] F. Cao, P. Musé, and F. Sur, "Extracting meaningful curves from images," *Journal of Mathematical Imaging and Vision*, vol. 22, no. 2, pp. 159–181, 2005.
- [12] T. Hurtut and F. Cheriet, "Automatic closed edge detection using level lines selection," *Image Analysis and Recognition*, pp. 187–197, 2007.
- [13] J. Serra, *Image analysis and mathematical morphology*, 1982.
- [14] N. Burrus, T. Bernard, and J. Jolion, "Image segmentation by a contrario simulation," *Pattern Recognition*, vol. 42, no. 7, pp. 1520–1532, 2009.
- [15] J. Cardelino, V. Caselles, M. Bertalmío, and G. Randall, "A contrario hierarchical image segmentation," in *16th IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 4041–4044.
- [16] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 860–872, 2000.
- [17] N. Burrus, "Apprentissage a contrario et architecture efficace pour la detection d'vnements visuels significatifs," Thse de Doctorat en Informatique, UPMC, Dec. 2008.
- [18] T. Eiter and H. Mannila, "Computing discrete fréchet distance," *Technische Universität Wien Technical Report CD-TR*, vol. 94, p. 64, 1994.
- [19] C. Fowlkes, D. Martin, and J. Malik, "Local figure-ground cues are valid for natural images," *Journal of Vision*, vol. 7, no. 8, 2007.