Research Article

A Biologically Motivated Multiresolution Approach to Contour Detection

Giuseppe Papari,[1] Patrizio Campisi,[2] Nicolai Petkov,[1] and Alessandro Neri[2]

[1] *Institute of Mathematics and Computing Science, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands*
[2] *Dipartimento di Elettronica Applicata, Università degli Studi di Roma "Roma Tre", Via della Vasca Navale 84, 00146 Roma, Italy*

Standard edge detectors react to all local luminance changes, irrespective of whether they are due to the contours of the objects represented in a scene or due to natural textures like grass, foliage, water, and so forth. Moreover, edges due to texture are often stronger than edges due to object contours. This implies that further processing is needed to discriminate object contours from texture edges. In this paper, we propose a biologically motivated multiresolution contour detection method using Bayesian denoising and a surround inhibition technique. Specifically, the proposed approach deploys computation of the gradient at different resolutions, followed by Bayesian denoising of the edge image. Then, a biologically motivated surround inhibition step is applied in order to suppress edges that are due to texture. We propose an improvement of the surround suppression used in previous works. Finally, a contour-oriented binarization algorithm is used, relying on the observation that object contours lead to long connected components rather than to short rods obtained from textures. Experimental results show that our contour detection method outperforms standard edge detectors as well as other methods that deploy inhibition.

1. INTRODUCTION

Contour detection is a fundamental operation in image processing and computer vision which, despite of the large number of studies published in the last two decades, is still a fertile field of ongoing research.

Many edge detectors have been proposed in the literature. However, they react to all local luminance changes above a given threshold, irrespective of their origin—object contours or textures. Our goal is to isolate objects in a scene; therefore, some further processing is needed beyond general-purpose edge detection.

Examples of edge detectors proposed in previous works are operators that incorporate linear filtering [1–5], local orientation analysis [4, 6, 7], fitting of analytical models to the image data [8–11]. In [12], a simple energy model is introduced to simulate perception of perceptually significant elements like lines and edges. Edge detectors using local energy principles have also been proposed in [13–16]. Since these operators do not make any difference between various types of edges, such as texture edges versus object contours and region boundaries, they are

known as *noncontextual* or, simply, general edge detectors [17].

Other studies propose more elaborate edge detection techniques that take into account additional information around an edge, such as local image statistics, image topology, perceptual differences in local cues (e.g., texture, color), edge continuity and density. Examples are dual frequency band analysis, statistical analysis of the gradient field [18, 19], anisotropic diffusion [20–23], complementary analysis of boundaries and regions [24–26], use of edge density information [9] and biologically motivated surround modulation [27–30]. These operators are not aimed at detecting all luminance changes in an image but rather at selectively enhancing only those of them that are of interest in the context of a specific computer vision task, such as detecting outlines of tissues in medical images, object contours in natural image scenes, and boundaries between different texture regions. Such methods are usually referred to as *contextual* edge detectors.

Psychophysical studies on the human visual system (HVS) have given rise to biologically motivated edge detectors [3, 14, 16, 31]. In its early stages, the HVS deploys special

mechanisms to differentiate between isolated edges, such as object contours and region boundaries, on the one hand, and edges in groups, such as those in textures, on the other hand. Various psychophysical studies have shown that the perception of an oriented stimulus, for example, a line segment, can be influenced by the presence of other such stimuli (distractors) in its neighborhood. This influence can, for instance, manifest itself in the decreased saliency of a contour in presence of surrounding texture [32, 33], in the orientation contrast pop-out effect [34], or in the decreased visibility of letters, object icons, and bars embedded in texture [30, 35]. These visual perception effects are in agreement with the results of neurophysiological measurements on neural cells in the primary visual cortex. These studies show that the response of an orientation selective visual neuron to an optimal bar stimulus in its classical receptive field is reduced by the addition of other oriented stimuli to the surround [36–38]. Neurophysiologists refer to this effect as *nonclassical receptive field (non-CRF) inhibition* [37, 38] or, equivalently, *surround suppression* [39]. Statistical data [37–39] reveals that about 80% of the orientation selective cells in the primary visual cortex show this inhibitory effect. In approximately 30% of all orientation selective cells, surround stimuli of orientation that are orthogonal to the optimal central stimulus have a weaker suppression effect than stimuli of the same orientation. In 40% of the cells, the suppression effect manifests itself irrespective of the relative orientation between the surrounding stimuli and the central one. In [27, 30], it is suggested that the biological utility of surround suppression is enhancement of object contours in natural images rich in background texture. This mechanism has been shown to improve the contour detection performance of biologically motivated [26] and conventional [40] edge detection algorithms.

Other psychophysical studies [41] on the HVS have shown that image perception can be divided in two subsequent stages: the preattentive stage and the attentive stage. In the first one, which lasts the first $0.1 \div 0.3$ s after an image is projected on the retina, coarse scale information is perceived, whereas in the latter, details are identified. Some psychophysical experiments [42] indicate that the visual information in different frequency bands is processed separately. Therefore it is assumed that the retinal image is decomposed through bandpass filters, which give rise to a multichannel model [43]. Psychophysical validation of multiresolution scheme based on a local energy model is provided in [44]. These psychophysical studies suggest us to perform contour detection in a multiresolution framework [43, 45, 46].

Contour detection becomes an even more challenging task when noisy images are involved. It is well known [18] that edge extraction operators enhance noise at high-spatial frequencies. Therefore, denoising needs to be deployed. Within this framework, the definition of a priori probability model for both the noise and for images is of great importance. However, modeling the statistics of natural images is a difficult task, due to the image nonstationarity. Several attempts to model image statistics in transform domains

have been recently performed. Denoising algorithms operating in the wavelet domain have been proposed in [40, 47–52]. Specifically, in [40, 51, 52] it is assumed that the wavelet coefficients within a local neighborhood are characterized by a Gaussian scale mixture (GSM). In [53] an image denoising method based on an image representation in the edge domain and on the Bayesian estimation of the original feature is provided. Parametric probabilistic models based on Gaussian mixtures are adopted for both signal and noise edge features. Such a model is taken into account in the current study to design a Bayesian denoising step that is applied to the gradient image and that leads to an orientation-dependant zero-memory nonlinearity.

In this paper, we propose a novel, biologically motivated, multiresolution contour detector which makes use of Bayesian denoising and of an improved surround inhibition technique. Within the framework of this paper the term contour is used to represent a line delimiting an object or part of it in a scene. This is a more sophisticated concept than edge which represents a not negligible local luminance change. Therefore, in our approach, contour detection is a global concept related to the recognition of meaningful objects. Specifically, the proposed method consists in the computation of the gradient at different resolutions, followed by Bayesian denoising of the edge image. Within this framework both the a priori first-order probability density function of the edge image and of the noise are modeled as a mixture of Gaussian distributions. This approach allows us to robustly estimate the image gradient. Then, a biologically motivated surround inhibition step is applied in order to suppress the edges due to texture. When surround inhibition is applied in the way proposed in [27, 30, 54], object contours are also partially suppressed in a self-inhibition process. We propose a new inhibition scheme that overcomes this problem and allows more effective inhibition of texture edges. Finally, a binarization algorithm is used that operates on connected edge components and relies on the observation that true contours lead to long connected components rather then to short rods obtained from textures.

The paper is organized as follows. In Section 2, the proposed approach is described in detail for the single-scale case. Then it is generalized in Section 3 to the multiscale case. In Section 4 experimental results are given. Finally, conclusions are drawn in Section 5.

2. SINGLE-SCALE CONTOUR DETECTOR

The proposed single-scale contour detector is sketched in Figure 1, where $\mathbf{I} = \{I(x, y)\}$ represents the original image, $\mathbf{I}^w = \{I^w(x, y)\}$ is its observed version corrupted by an additive independent observation noise $\mathbf{W} = \{w(x, y)\}$, $\mathbf{I}^w = \mathbf{I} + \mathbf{W}$, and $\nabla_\sigma \mathbf{I}^w = \{\nabla_\sigma I^w(x, y)\} = \nabla_\sigma \mathbf{I} + \nabla_\sigma \mathbf{W}$ is the scale-dependent gradient of the noisy image, computed as described in Section 2.1. A Bayesian denoising algorithm, described in Section 2.2, is applied on the gradient of the noisy image, followed by a surround inhibition step for texture suppression (Section 2.3). In Section 2.4, a contour-based binarization algorithm is described.
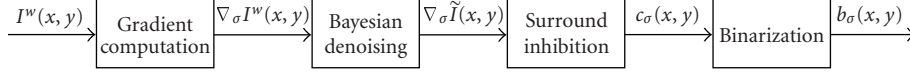
FIGURE 1: Flowchart of the proposed single-scale contour detector.

The mathematical operator that gives the binary contour map $b_\sigma(x, y)$ detected at the resolution σ from the original image $I^w(x, y)$ will be referred as RDCD_σ (resolution-dependant contour detector):

$$b_\sigma(x, y) = \mathrm{RDCD}_\sigma\{I^w(x, y)\}. \tag{1}$$

In the notation of this section, we will use the subscript σ to indicate the dependence of the introduced quantities and operators on the resolution parameter σ.

2.1. Scale-dependent gradient

Given the (noisy) input image $I^w(x, y)$, the first step toward the estimation of its contours is the computation of a scale-dependent gradient $\nabla_\sigma I^w(x, y)$, defined as follows:

$$
\begin{aligned}
\nabla_\sigma I^w(x, y) &= \nabla\{I^w * g_\sigma\}(x, y) \\
&= \begin{bmatrix} \left\{I^w * \dfrac{\partial g_\sigma}{\partial x}\right\}(x, y) \\ \left\{I^w * \dfrac{\partial g_\sigma}{\partial y}\right\}(x, y) \end{bmatrix},
\end{aligned} \tag{2}
$$

where the image $I^w(x, y)$ is convolved with the x and y derivatives of a Gaussian function [1]:

$$g_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \tag{3}$$

The operator ∇_σ defined in (2) and (3) depends on the parameter σ, that we will call the *scale* or *resolution* parameter. Gradient computation according to (2) depends on the scale parameter σ: the larger its value, the larger the spatial extent of the intensity transitions (blur) to which the operator responds.

2.2. Bayesian denoising

A Bayesian estimate $\nabla_\sigma \widetilde{\mathbf{I}}$ of $\nabla_\sigma \mathbf{I}$, given $\nabla_\sigma \mathbf{I}^w$, is obtained by the minimization of the associated absolute risk. Therefore, our goal is the minimum mean square error (MMSE) estimation of the edge image $\nabla_\sigma \mathbf{I} = \{\nabla_\sigma I(x, y)\}$, given by the a posteriori expectation $\nabla_\sigma \widetilde{\mathbf{I}} = \{\nabla_\sigma \widetilde{I}(x, y)\}$ of $\nabla_\sigma \mathbf{I} = \{\nabla_\sigma I(x, y)\}$, given $\nabla_\sigma \mathbf{I}^w = \{\nabla_\sigma I^w(x, y)\}$. Neglecting residual spatial correlation, we propose a suboptimum estimation procedure based only on the marginal a priori edge distribution. Thus, we evaluate the conditional expectation $\nabla_\sigma \widetilde{I}(x, y)$ of $\nabla_\sigma I(x, y)$ given $\nabla_\sigma I^w(x, y)$ at site (x, y) only. To this aim, let us describe the marginal distribution of $\nabla_\sigma I(x, y)$ with a rather general model constituted by a Gaussian mixture, that is, a weighted sum of Gaussian distributions:

$$p_{\nabla_\sigma I}[\nabla_\sigma I(x, y)] = \sum_{i=1}^{K} \lambda_i \mathcal{N}_2[\nabla_\sigma I(x, y), 0, \mathbf{R}_{\nabla_\sigma I_i}(x, y)], \tag{4}$$

where $\mathcal{N}_2[\chi, \mu, \mathbf{R}]$ denotes the Gaussian probability density function (p.d.f.) of a bivariate random variable $\mathbf{f}$ with expectation $\mathbf{m}$ and covariance matrix $\mathbf{R}$:

$$
\begin{aligned}
&\mathcal{N}_2[\mathbf{f}, \mathbf{m}, \mathbf{R}] \\
&= \frac{1}{2\pi[\det(\mathbf{R})]^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{f} - \mathbf{m})^T \mathbf{R}^{-1}(\mathbf{f} - \mathbf{m})\right\}.
\end{aligned} \tag{5}
$$

As to the gradient of the observation noise, we model again the p.d.f. of the random variable $\nabla_s W(x, y)$ as a zero mean Gaussian mixture with mixing parameters β_j, namely,

$$p_{\nabla_\sigma W}[\nabla_\sigma W(x, y)] = \sum_{j=1}^{M} \beta_j \mathcal{N}_2[\nabla_\sigma W(x, y), 0, \mathbf{R}_{\nabla_\sigma W_j}(x, y)]. \tag{6}$$

Derivation of the suboptimum Bayesian estimator based on edges requires the calculation of the a posteriori p.d.f. of $\nabla_\sigma I(x, y)$ given $\nabla_\sigma I^w(x, y)$. Applying Bayes rule and dropping the location (x, y) for the sake of compactness we obtain

$$
\begin{aligned}
&p_{\nabla_\sigma I / \nabla_\sigma I^w}(\nabla_\sigma I / \nabla_\sigma I^w) \\
&= \frac{p_{\nabla_\sigma I^w / \nabla_\sigma I}(\nabla_\sigma I^w / \nabla_\sigma I)\, p_{\nabla_\sigma I}(\nabla_\sigma I)}{\int p_{\nabla_\sigma I^w / \nabla_\sigma I}(\nabla_\sigma I^w / \nabla_\sigma I)\, p_{\nabla_\sigma I}(\nabla_\sigma I)\, d(\nabla_\sigma I)} \\
&= \frac{\sum_j^M \sum_i^N \beta_j \lambda_i \mathcal{N}_2[\nabla_\sigma I^w, \nabla_\sigma I, \mathbf{R}_{\nabla_\sigma W_j}] \mathcal{N}_2[\nabla_\sigma I, 0, \mathbf{R}_{\nabla_\sigma I_i}]}{\sum_j^M \sum_i^N \beta_j \lambda_i \mathcal{N}_2[\nabla_\sigma I^w, \nabla_\sigma I, \mathbf{R}_{\nabla_\sigma W_j} + \mathbf{R}_{\nabla_\sigma I_i}]}.
\end{aligned} \tag{7,8}
$$

The evaluation of the conditional expectation $\nabla_\sigma \widetilde{I}(x, y) = E[\nabla_\sigma I(x, y) / \nabla_\sigma I^w(x, y)]$ associated with (8) can be written as

$$
\begin{aligned}
\nabla_\sigma I(x, y) &= \mathrm{ZNL}[\nabla_\sigma I^w(x, y)] \\
&= \sum_{j=1}^{M} \sum_{i=1}^{N} \eta_{ij}[\nabla_\sigma I^w(x, y)] \mathbf{R}_{\nabla_\sigma I_i}(x, y) \\
&\quad \times (\mathbf{R}_{\nabla_\sigma W_i}(x, y) + \mathbf{R}_{\nabla_\sigma I_i}(x, y))^{-1} \nabla_\sigma I^w(x, y),
\end{aligned} \tag{9}
$$

where ZNL stands for zero-memory nonlinearity, and

$$
\begin{aligned}
&\eta_{ij}[\nabla_\sigma I^w(x, y)] \\
&= \frac{\beta_j \lambda_i \mathcal{N}_2[\nabla_\sigma I^w(x, y), 0, \mathbf{R}_{\nabla_\sigma W_j}(x, y) + \mathbf{R}_{\nabla_\sigma I_i}(x, y)]}{\sum_j^M \sum_i^N \mathcal{N}_2[\nabla_\sigma I^w(x, y), 0, \mathbf{R}_{\nabla_\sigma W_j}(x, y) + \mathbf{R}_{\nabla_\sigma I_i}(x, y)]}.
\end{aligned} \tag{10}
$$

Equation (9) says that in general, for signal and noise Gaussian mixtures, the MMSE estimator is a nonlinear combination of conditionally optimal linear estimators, with

<center>(a)   (b)</center>
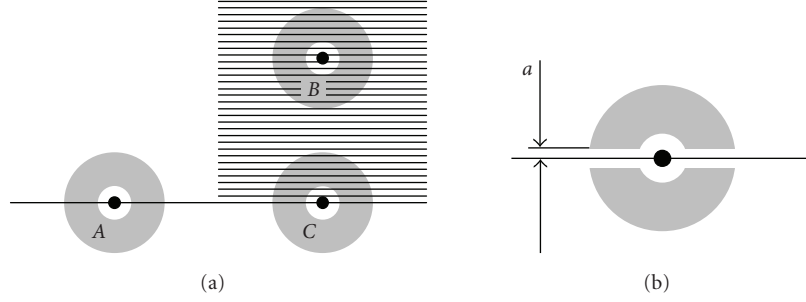
FIGURE 2: The inhibition term for a given point is computed by weighted summation of the response in the shaded surroundings of that point. (a) The annular surround proposed in [27, 30, 54] is effective for dense texture areas (point B) but leads to undesirable partial self-inhibition of isolated edges (point A) and considerable inhibition of texture region boundaries (point C). (b) In the current paper, the inhibition surround is split into two truncated half-rings oriented along the concerned edge and the inhibition term is computed as the minimum of the two weighted averages of $M_\sigma(x, y)$ on these two truncated half-rings.

gains $\mathbf{R}_{\nabla_\sigma I_i}(x, y)(\mathbf{R}_{\nabla_\sigma W_i}(x, y) + \mathbf{R}_{\nabla_\sigma I_i}(x, y))^{-1}$ each matched to a pair (x, y) of Gaussian submodels. The weights $\eta_{ij}[\nabla_\sigma I^w(x, y)]$ represent the posterior probabilities of each submodel pair.

2.3. Surround inhibition

2.3.1. Previous work

Next, following [54], we deploy a surround inhibition mechanism that takes into account the context influence of the surroundings of each point. It consists in computing an inhibition term as an integral of the gradient magnitude in the surroundings of a point and subtracting this term from the gradient magnitude in the concerned point. The inhibition term is supposed to be large in textured areas and low on object contours thus leading to the suppression of texture while retaining contours. This operator is motivated by psychophysical and neurophysiological findings (see [26] for arguments and further references).

Let $M_\sigma(x, y)$ be the gradient magnitude:

$$
\begin{aligned}
M_\sigma(x, y) &= |\nabla_\sigma I(x, y)| \\
&= \sqrt{\left[\left\{I * \frac{\partial g_\sigma}{\partial x}\right\}(x, y)\right]^2 + \left[\left\{I * \frac{\partial g_\sigma}{\partial y}\right\}(x, y)\right]^2}.
\end{aligned}
\tag{11}
$$

In [54], the inhibition term $T_\sigma(x, y)$ is defined as the weighted local average of $M_\sigma(x, y)$ on a ring around each pixel and it is computed as the convolution of $M_\sigma(x, y)$ and a weighting function $w_\sigma(x, y)$:

$$
t_\sigma(x, y) = \{M_\sigma * w_\sigma\}(x, y).
\tag{12}
$$

The weighting function $w_\sigma(x, y)$, according to [27, 30, 54], is a half-wave rectified and L_1-normalized difference of two concentric Gaussian functions:

$$
\begin{aligned}
\mathrm{DoG}_\sigma(x, y) &= |g_{k\sigma}(x, y) - g_\sigma(x, y)|^+ \\
w_\sigma(x, y) &= \frac{\mathrm{DoG}_\sigma(x, y)}{\iint_{R^2} \mathrm{DoG}_\sigma(x, y) dx \, dy},
\end{aligned}
\tag{13}
$$

where $| \cdot |^+$ denotes half-wave rectification,

$$
|\xi|^+ = \begin{cases} \xi, & \xi \geq 0, \\ 0, & \xi < 0. \end{cases}
\tag{14}
$$

The support of $w_\sigma(x, y)$ defines the annular surround of a point on which the gradient magnitude is integrated, thus obtaining the value of the inhibition term for that point (Figure 2(a)). The central region that is excluded from the inhibition term computation is the essential support of the gradient operator. It can be considered as an analogue of the classical receptive field (CRF) of an orientation selective neuron in the primary visual cortex. The annular area around it can be considered as the surround of that CRF. The radius ρ_0 of the concerned central region is given by

$$
\rho_0(k) = 2\sigma\sqrt{\frac{\ln k}{1 - 1/k^2}}
\tag{15}
$$

and is a slowly changing function of the parameter k. For instance, for $k = 4$, we have $\rho_0 \cong 2.5\sigma$. The weighting function $w_\sigma(x, y)$ is essential in a region of radius $k\rho_0(k)$, thus the radius of the annular surround is roughly k times larger than the radius of the central (CRF) region. In our experiments we take the value $k = 4$, corresponding to an inhibition surround being several times larger (in diameter) than the *classical* receptive field of visual neurons that exhibit surround modulation [36]. Our experiments show that the performance of the proposed method does not depend significantly on the value of this parameter: for values of k between 3 and 6, the performance change is negligible (see Section 4.2.2).

The inhibition term computed in this way will be large for points in whose surroundings there are multiple edges, such as point B in Figure 2(a). In contrast, it will be small for points along isolated edges, such as point A in Figure 2(a). Therefore, subtraction of this term from the gradient magnitude leads to texture suppression while leaving isolated contours relatively unaffected. The result $c_\sigma(x, y)$ of the inhibition is computed as follows:

$$
c_\sigma(x, y) = |M_\sigma(x, y) - \alpha t_\sigma(x, y)|^+.
\tag{16}
$$

The coefficient α, called *inhibition strength*, specifies the extent to which the inhibition term is taken into account. Depending on the value of α, the inhibition term can partially or completely suppress the response of the operator to texture edges. For this type of surround suppression, we choose the value of the inhibition coefficient α to be such that the following equation is fulfilled at the points of maximum of $M_\sigma(x, y)$, when the input image is a bar grating of bar spacing and bar width ρ_0:

$$M_\sigma(x, y) = \alpha t_\sigma(x, y). \tag{17}$$

The radius ρ_0 of the "receptive field" of the gradient operator is chosen to be equal to the bar spacing and bar width so that only one edge is visible in that field. This is the smallest value of α for which the operator will not respond to a texture input defined by such a bar grating. The idea is not only to suppress texture but to minimize the partial suppression of isolated edges and contours. The inhibition strength value which satisfies these conditions is $\alpha = 1.59$.

However, this straightforward inhibition process has two drawbacks.

(1) While being small, the inhibition term is not zero on isolated edges because parts of such an edge fall in the inhibition surround of other parts of the same edge, see point A in Figure 2(a). We refer to this effect as self-inhibition.

(2) Edges at texture borders, such as point C in Figure 2, are considerably inhibited as well, which is not desirable with respect to the detection of region boundaries.

2.3.2. Improved inhibition scheme

In this paper we propose a modification of the inhibition scheme that does not suffer the above-mentioned drawbacks. The idea is to exclude from the annular surround of a point a band region of width $2a$ oriented along the edge, as shown in Figure 2(b). We define the inhibition term T_σ as the minimum of the two weighted local averages of $M_\sigma(x, y)$ on the two resulting half-rings. More specifically, we define two weighting functions $w^+_{\sigma,\phi}(x, y)$ and $w^-_{\sigma,\phi}(x, y)$:

$$\text{DoG}^+_{\sigma,\phi}(x, y) = \text{DoG}_\sigma(x, y) \cdot U(x\cos\phi + y\sin\phi - a),$$
$$\text{DoG}^-_{\sigma,\phi}(x, y) = \text{DoG}_\sigma(x, y) \cdot U(a - x\cos\phi - y\sin\phi),$$
$$w^\pm_{\sigma,\phi}(x, y) = \frac{\text{DoG}^\pm_{\sigma,\phi}(x, y)}{\iint_{R^2} \text{DoG}^\pm_{\sigma,\phi}(x, y)dx\,dy}, \tag{18}$$

where $\phi \in [0, \pi)$ is a generic orientation and U is the step function defined as follows:

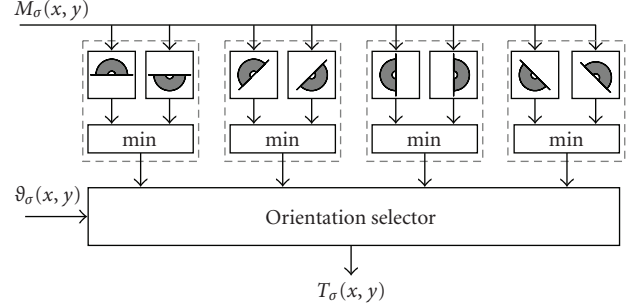$$U(\xi) = \begin{cases} 0, & \xi < 0, \\ 1, & \xi \geq 0. \end{cases} \tag{19}$$

FIGURE 3: Computation scheme of the inhibition term. For each pixel of the image, inhibition terms are computed for a number of different orientations. Then the gradient orientation information is used to select the appropriate value at each pixel.

Then, we define and compute the modified inhibition term as follows:

$$T_\sigma(x, y) = \min\left\{[M_\sigma * w^+_{\sigma,\phi}](x, y), [M_\sigma * w^-_{\sigma,\phi}](x, y)\right\}, \tag{20}$$

where $\vartheta(x, y)$ is the orientation of $\nabla_\sigma \widetilde{I}(x, y)$.

In practice, we compute the convolutions in (20) for a discrete set of orientations $\{\phi_i\}_{i=1}^{N_\phi}$, $\phi_i = \pi((i-1)/N_\phi)$, as shown in Figure 3 for $N_\phi = 4$, and then, for each pixel, we use the result obtained for the angle that is closest to the gradient orientation $\vartheta_\sigma(x, y)$ for that pixel. Our experiments show that (above a certain reasonable minimum of $N_\phi = 4$) the number of orientations used does not substantially influence the performance of the contour detection operator (see Section 4.2.2).

The exclusion of the central band region avoids the self-inhibition and is motivated by neurophysiological studies [36] according to which, inhibitory modulation originates from the regions flanking the receptive field of an orientation selective neuron on both sides of the optimal stimulus for that neuron. The parameter a controls the width of the excluded band region and we set it to be a fraction η of the radius ρ_0, $a = \eta\rho_0$. Our experiments show that for values of η around 1 the exact choice of η is not critical for the performance of our algorithm (see Section 4.2.2). Therefore, we use $\eta = 1$ in the following. As to the specific choice of the minimum function used in (20), at the current moment, this is a pure design consideration. A certain neurophysiological justification for this choice can be sought for by the fact that the inhibition surround of a neuron need not be circular symmetric. For instance, only 23% of cells in area MT/V5 show circular symmetrical surrounds while 45% of the cells have asymmetrical surrounds [55]. In this context (20) can be considered as a maximum value combination of two surround suppression operators with opposite asymmetrical surrounds as defined by the half-rings shown in Figure 2(b). The result is a computation of a directional derivative of the gradient magnitude in direction of the gradient and can be used for effectively detecting region boundaries for the gradient magnitude as illustrated by Figure 2(b).
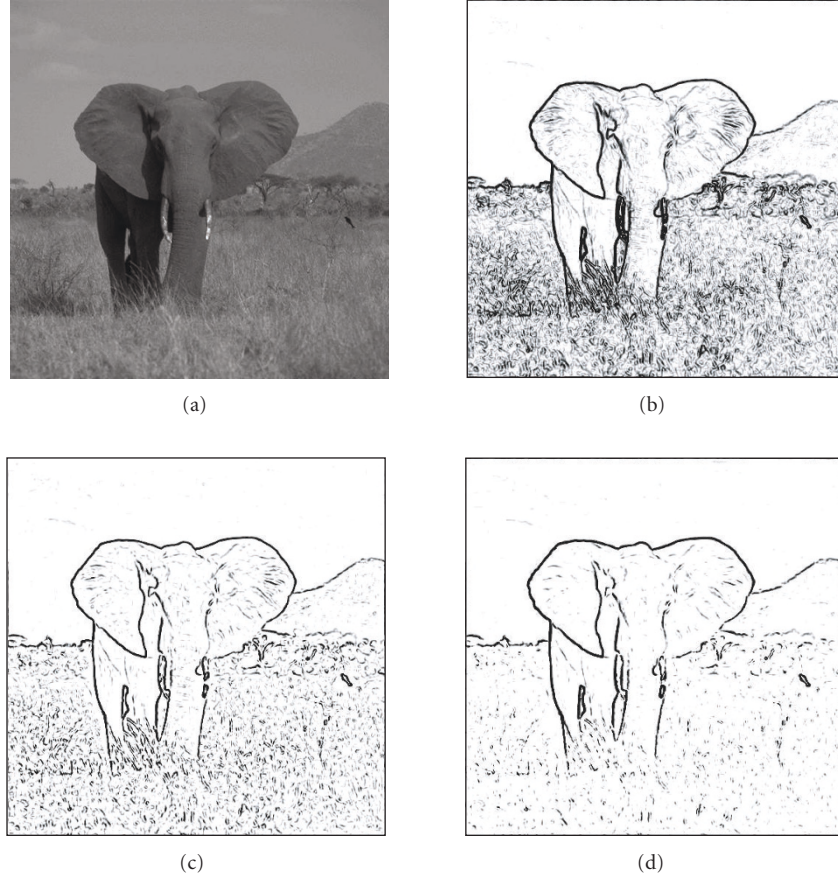
FIGURE 4: (a) A test image elephant and edge strength computed as (b) the gradient magnitude, (c) the gradient magnitude with surround inhibition according to the traditional annular surround method [54] with $\alpha = 1.59$, and (d) the gradient magnitude with surround inhibition according to the split-surround method proposed in this paper with $\alpha = 3$. For a better representation, the three edge images have been equalized and shown in negative, thus white pixels correspond to the value zero.

The edge strength $c_\sigma(x, y)$ is computed similar to (16), with the inhibition term $T_\sigma(x, y)$ according to (20),

$$c_\sigma(x, y) = \left| M_\sigma(x, y) - \alpha T_\sigma(x, y) \right|^+. \tag{21}$$

Figure 4 shows a test image elephant (Figure 4(a)) and three gray level edge images representing the gradient magnitude $M_\sigma(x, y)$ without surround inhibition (Figure 4(b)), the edge strength $c_\sigma(x, y)$ computed according to the previous inhibition scheme [54] (Figure 4(c)), and the edge strength $c_\sigma(x, y)$ computed according to the improved inhibition scheme proposed here (Figure 4(d)). Since no self-inhibition is involved in the proposed modified inhibition scheme, a higher value of the parameter α can be used without destroying weak edges. In this way texture can be suppressed more effectively.

2.4. Binarization

Similar to other methods for edge and contour detection, the last step of the algorithm comprises edge thinning by non-maxima suppression and binarization by thresholding. Traditional thresholding techniques, such as global or hysteresis thresholding [3], cannot deal adequately with texture edges that present stronger gradient magnitude values than contours, Figure 5.

In this paper we present a new thresholding algorithm, based on the observation that object contours lead to long and wide connected components of nonzero pixels, while texture edges, especially after surround inhibition, lead to relatively short and thin components. Specifically, we apply nonmaxima suppression to the signal $c_\sigma(x, y)$. Let $\mathbf{u}_\sigma(x, y)$ be the unit vector parallel to the gradient $\nabla_\sigma I(x, y)$, that is, $\nabla_\sigma I(x, y) = M_\sigma(x, y)\mathbf{u}_\sigma(x, y)$; we consider the set S_σ of all points which are local maxima of $c_\sigma(x, y)$ in the direction of $\mathbf{u}_\sigma(x, y)$:

$$S_\sigma = \left\{ (x, y) \ \middle| \ \frac{\partial c_\sigma}{\partial u_\sigma} = 0 \wedge \frac{\partial c_\sigma}{\partial u_\sigma} < 0 \right\}. \tag{22}$$

Let $C_k^{(\sigma)}$, $k = 1, \ldots, N_c$, be the connected components of the set S_σ,
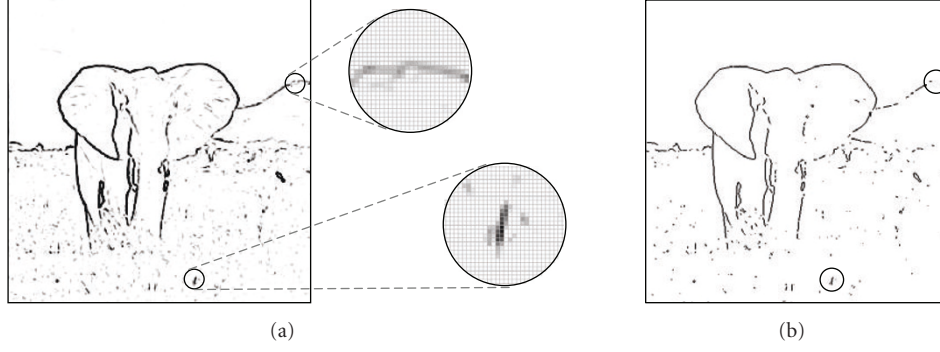
$$S_\sigma = \bigcup_k C_k^{(\sigma)}, \tag{23}$$

FIGURE 5: (a) Gray level contour image $c_\sigma(x, y)$ obtained after surround inhibition. (b) Result of traditional binarization comprising thinning by nonmaxima suppression and thresholding. Some contour pixels are weaker than some texture edges and it is not possible to select a threshold that retains the former while eliminating the latter.
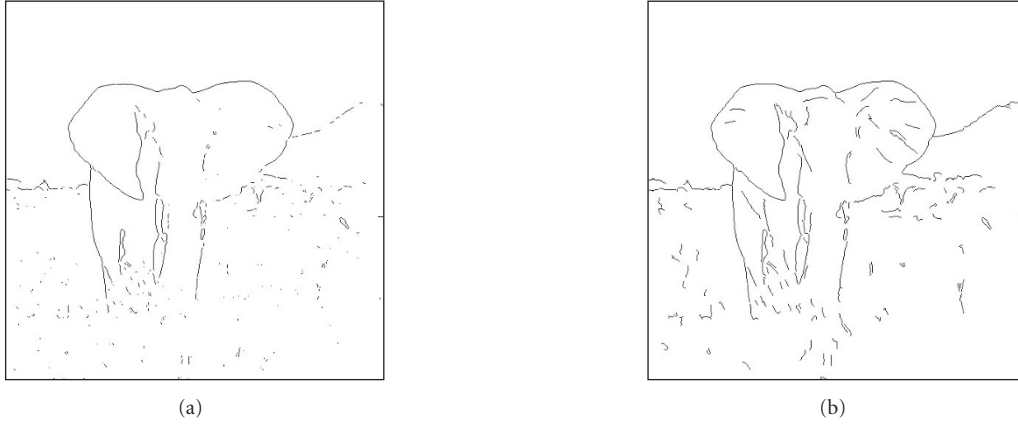

FIGURE 6: Results of binarization by (a) traditional thresholding and (b) the proposed connected component weight thresholding.

where N_c is the number of such components. We apply a morphological dilation to $C_k^{(\sigma)}$ [56, 57], with a 3×3 square q_3 as structuring element, and obtain dilated components $D_k^{(\sigma)}$:

$$D_k^{(\sigma)} = C_k^{(\sigma)} \oplus q_3. \tag{24}$$

For each connected component $C_k^{(\sigma)}$, we introduce a quantity $G_k^{(\sigma)}$, which we call *global contour weight*, defined as the sum of the values of $c_\sigma(x, y)$ over the dilated component $D_k^{(\sigma)}$:

$$G_k^{(\sigma)} = \sum_{(x,y) \in D_k^{(\sigma)}} c_\sigma(x, y). \tag{25}$$

We compute a binary contour map $b_\sigma(x, y)$ by setting to 1 the value of the pixels from all connected components $C_k^{(\sigma)}$ whose contour weights $G_k^{(\sigma)}$ are above a given threshold $G_{\min}$:

$$b_\sigma(x, y) = \begin{cases} 1, & (x, y) \in \bigcup_{G_k^{(\sigma)} > G_{\min}} C_k^{(\sigma)}, \\ 0, & \text{otherwise}. \end{cases} \tag{26}$$

The result of this type of thresholding compared to traditional thresholding is shown in Figure 6. Low-contrast contours are successfully detected and, most importantly, contours are not depleted by the binarization process.

Since the value of the contour strength G is related to the length of the contours of the object represented in an image, the value of the threshold $G_{\min}$ should be proportional to the linear size of the image. In our experiments, performed on images of size 512×512 pixels, we found empirically that connected components that contain less than 7 pixels are too small to be part of an object contour. Therefore, unless a different value is specified, in our experiments we set the value of the threshold to $G_{\min} = 7$. In Section 4 we will discuss quantitatively the dependence of the performance of the algorithm with respect to the value of $G_{\min}$.

A similar connected component thresholding method has been proposed in [58]. However, in our computational experiments we found out that without surround inhibition this thresholding technique gives bad results. The reason is that if the set S_σ of the nonzero pixels of the gradient after non-maxima suppression is computed directly from the gradient magnitude without surrounding inhibition, the
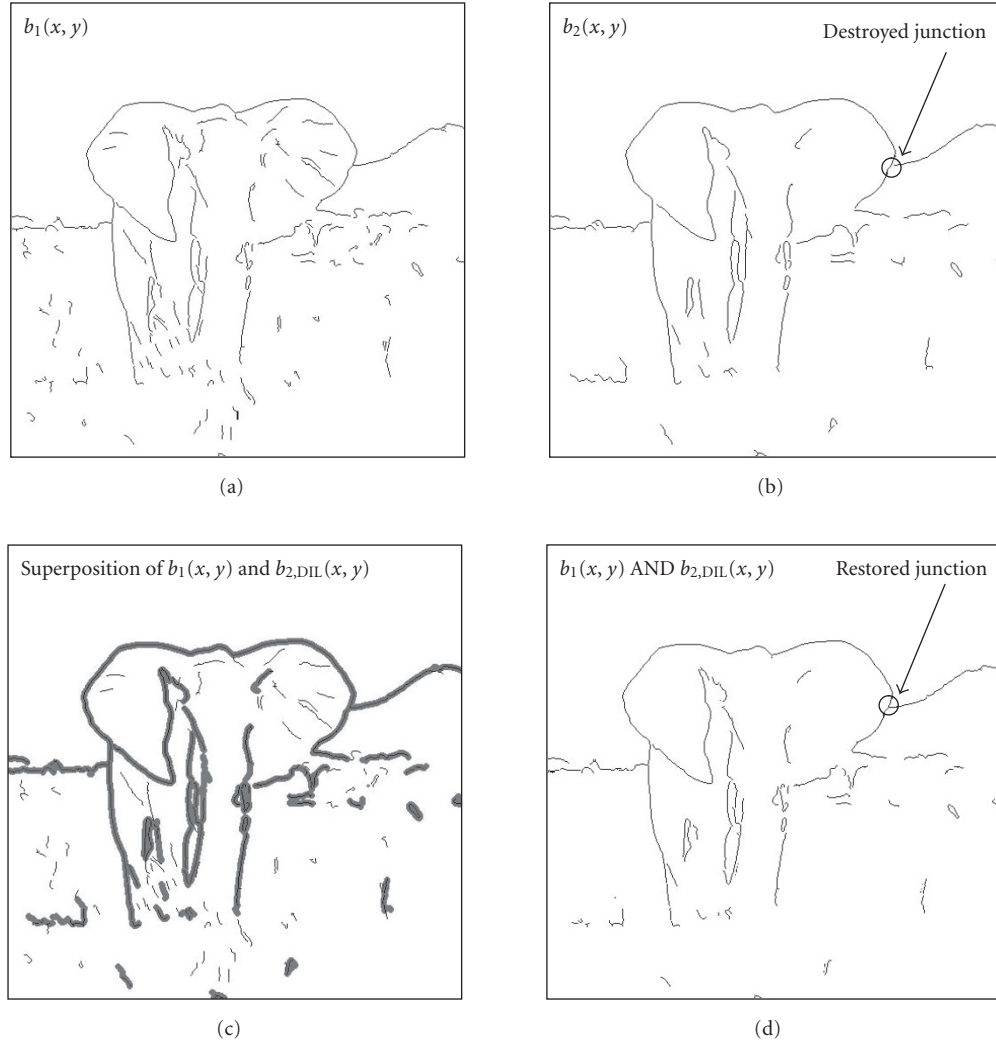
FIGURE 7: (a), (b) Binary contour maps $b_1(x, y)$ and $b_2(x, y)$ obtained with the RDCD operator introduced in Section 2, with $\sigma = 1$ and $\sigma = 2$, respectively. At the finer scale the borders are detected at their respective positions, but some texture is present; at the coarser scale, texture is reduced but the contours are shifted and some junctions are destroyed. (c) Superposition of the contour map at the fine scale (shown in black) and the morphologically dilated contour map at the coarse scale (rendered in a gray), and (d) the result of their logic AND. Texture is reduced, contours are well detailed, and the morphological dilatation restores the junctions.

result includes many large tangled connected components of nonzero pixels that originate from noisy image regions. Such components have high contour strength values and cannot be eliminated by the proposed thresholding scheme. In contrast, surround inhibition breaks such connected components into pieces that are small enough and consequently have small contour strengths and can effectively be removed by thresholding.

3. MULTISCALE CONTOUR DETECTOR

It is well known from multiresolution wavelet analysis [43] that coarser scales contain only the general morphology of the image where most of the high-frequency texture details disappear. This fact is illustrated in Figures 7(a) and 7(b), displaying two binary contour maps $b_1(x, y)$ and $b_2(x, y)$ ob-

tained with the RDCD operator defined above for $\sigma = 1$ and $\sigma = 2$, respectively. From Figure 7(a) we can see that the contours detected at the fine scale ($\sigma = 1$) are detected at their correct positions and the junctions are preserved, but at the same time much texture is present. When a coarser scale is used ($\sigma = 2$, Figure 7(b)), some texture is removed, but the contours are shifted away from their true positions [59], especially at positions of high curvature, and some junctions are destroyed by the nonmaxima suppression [60].

In order to exploit the advantages of both resolutions we superpose the two binary images and we select from $b_1(x, y)$ only those "1" pixels that are close enough to "1" pixels of $b_2(x, y)$. More specifically, we first apply a morphological dilation operator with a disk of radius 3σ as structuring element on the edge map $b_2(x, y)$ at the coarse scale and we denote the result by $b_{2,\text{DIL}}(x, y)$. Figure 7(c) shows the
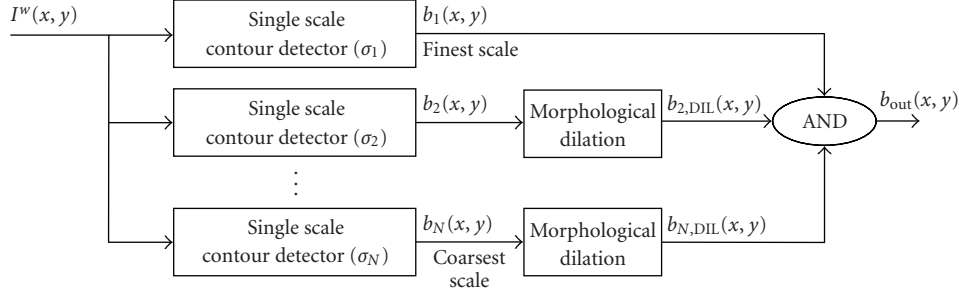
FIGURE 8: Overall scheme of our multiscale contour detector, where each block "single-scale contour detector" implements the RDRC operator with a different scale parameter σ_i.

superposition of $b_1(x, y)$ and $b_{2,\text{DIL}}(x, y)$. The object contours, well detailed and localized in $b_1(x, y)$, are contained in $b_{2,\text{DIL}}(x, y)$; on the contrary, most of the texture present in $b_1(x, y)$ is not present in $b_{2,\text{DIL}}(x, y)$. Consequently, the logic AND of $b_{2,\text{DIL}}(x, y)$ and $b_1(x, y)$ shown in Figure 7(d) has well-detailed contours similar to $b_1(x, y)$ and less texture edges similar to $b_{2,\text{DIL}}(x, y)$. The morphological dilation compensates for the contour shifting at the coarse scale and restores the junctions.

In our approach we apply an N-level multiscale analysis in order to remove the residual spurious texture still present in Figure 7(d). This algorithm relies on the observation that starting from a given scale that is determined by the object blur, object contours are present in the results at all scales, while texture appears only at the finer scales. Referring to the scheme shown in Figure 8, we first compute the binary contour maps $b_k(x, y)$ at N different scales:

$$b_k(x, y) = \text{RDCD}_{\sigma_k}\{I^w(x, y)\}, \quad k = 1, \dots, N. \quad (27)$$

Then we apply morphological dilation to all binary maps but the one that corresponds to the finest scale:

$$b_{k,\text{DIL}} = b_k \oplus D_{3\sigma}, \quad k = 2, \dots, N, \quad (28)$$

where we use as a structuring element a disk $D_{3\sigma}$ of radius 3σ. The final output is obtained by the logic AND of the binary maps at all resolutions:

$$b_{\text{out}} = b_1(x, y) \cdot \prod_{k=2}^{N} b_{k,\text{DIL}}(x, y). \quad (29)$$

In the previous discussion, the scale values σ_k have been considered as input parameters. Simple general considerations about the noise levels allow us to compute them automatically, thus making the algorithm unsupervised in this respect. The idea is that the only new information carried by the finer resolution channels with respect to the coarser ones is the *details* of the contours. However, when noise is present, human observers are not able to distinguish details of the contours and only the general shape of the objects is perceived (Figure 9(b)). Consequently, for noisy images the information carried by the edge maps at the finest resolutions can be discarded [58, 61].

With this idea in mind, we perform a preliminary estimation of the noise level and use it to determine the value of σ_1 of the finest scale, which must be larger the larger the noise is. It can be easily proved that, when the gradient is smoothed by a Gaussian mask $g_\sigma(x, y)$, the noise reduction is given by

$$\frac{N_{\text{out}}}{N_{\text{in}}} = \frac{\text{erf}\left(\pi\sigma\sqrt{2}\right)}{\sigma\sqrt{\pi}}, \quad (30)$$

where N_{in} and N_{out} are the noise levels before and after the smoothing, and

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt. \quad (31)$$

Therefore, once the noise level N_{est} of the input image has been estimated, the value of σ_1 can be obtained by solving (30), where N_{in} is set equal to N_{est}, and N_{out} is set to a fixed value, above which contours cannot be detected reliably anymore. We compute the value σ_i for the ith resolution as follows:

$$\sigma_i = \sigma_1 \cdot 2^{i-1}, \quad i = 1, \dots, N. \quad (32)$$

4. EXPERIMENTAL RESULTS

In this section some experimental results are presented and discussed. The performance of the proposed contour detector is compared with the performance of four other existing algorithms: the standard single-scale Canny edge detector [1], a modification[1] of the multiscale edge detector CARTOON [45], the single-scale surround inhibition (SSSI) contour detector proposed in [54], and the multiscale

---

[1] In the original CARTOON method as proposed in [45], only two values of σ are used and the edges are detected using the Laplacian of Gaussian filter (LoG). On the other hand, the multiscale algorithms proposed here and in [46] make use of multiple resolutions and detect edges by means of the gradient of Gaussian filter. In order to do a fair comparison between the proposed method and the CARTOON approach, we have reimplemented CARTOON by using the gradient of Gaussian filter for detecting edges and by using the same values of σ that are used in the other multiscale approaches discussed in this paper.

surround inhibition (MSSI) contour detector proposed in [46]. We performed experiments on a set of 40 images using both noiseless (SNR = ∞) and noisy image versions corrupted by additive noise with SNR equal to 10 dB, 13 dB, and 16 dB.

4.1. Qualitative comparison

Some experimental results are shown in Figures 9–22 for both noiseless images and images corrupted by additive noise of SNR = 13 dB. A larger set of examples is available on http://www.cs.rug.nl/~imaging. We would like to stress that we used the same set of parameter values for all images in the dataset as follows: inhibition strength $\alpha = 3$, binarization threshold $G_{min} = 7$, ratio of the two standard deviations in DoG $k = 4$, number of orientations for computing the inhibition term $N_\phi = 4$, number of scales $N = 3$, radius of the structuring element used for dilation $r_\sigma = 3\sigma$, and noise amplitude N_{out} equal to 8% of the average standard deviation of the input image, computed across all images.

SSSI [54] applies surround inhibition in a single-scale context. The modification of CARTOON [45] that we use here operates in a multiresolution framework without applying surround inhibition. MSSI [46] uses the surround inhibition scheme proposed in [54] in a multiscale framework. The approach proposed here is an improvement of MSSI using Bayesian denoising, a modification of the inhibition term, and a new binarization scheme.

We can see that the approach proposed in this paper outperforms the other algorithms in terms of cleanness of the detected contours, amount of suppressed texture, and robustness to noise. In particular, the results for the test images rhino and frog (Figures 11–14) show the ability of our algorithm to suppress texture while effectively detecting weak edges in low-contrast images. On the other hand, the results for the test image bear1 (Figures 15-16) show the ability of the proposed method to suppress high-contrast oriented texture like the fur of the bear. All the other studied algorithms but MSSI completely fail in removing this type of texture. Figures 17-18 show the behavior of our algorithm with respect to low-frequency texture, like the plants in the background behind the bear. Such type of texture, well removed by our contour detector (Figure 17(b)), can neither be suppressed by SSSI techniques (Figure 17(d)), nor by simply projecting the image on a coarse scale domain as CARTOON and MSSI do. The simple combination of multiscale analysis and surround inhibition also fails in this case. Finally, the examples shown in Figures 19–22 illustrate the behavior of our algorithm for images containing multiple objects of different sizes. It can be noted that some object details, like for instance the windows of the building in Figure 19(a), are detected by the single-scale contour detectors, but not by the multiscale ones. Indeed, whether they should be considered object contours or texture to be suppressed depends on the specific application. For instance, in the ground truth provided in the Berkeley image dataset [62] such details are not considered as object contours.

By comparing the results of the modification of CARTOON (Figures 9–22(e)), SSSI (Figures 9–22(d)), and MSSI

(Figures 9–22(f)), we can see that multiscale analysis and surround inhibition play complementary roles: the combination of the two approaches gives much better results than those obtained by each of them separately. The Bayesian denoising step, the modified computation of the inhibition term, and the contour-oriented binarization technique introduced here further improve the quality of the obtained results: the residual texture placed in the neighborhood of contours, still present when applying MSSI (especially well visible in Figures 11(f) and 15(f)), disappears by applying the proposed approach. Also the residual noise present when applying MSSI is removed by our approach. Figure 23 illustrates the effectiveness of the Bayesian denoising step introduced in Section 2.2 for the noisy test image *elephant* shown in Figure 10(a) (SNR = 13 dB). If the entire process explained in Sections 2 and 3 had been applied without the Bayesian denoising step, we would get the output shown in Figure 23(a). It is definitely worse than the output obtained with the algorithm proposed in this paper, where the Bayesian denoising step is performed at all resolutions (Figure 23(b)).

4.2. Quantitative performance evaluation

4.2.1. Metric definition

Methods for performance evaluation of edge detectors can be categorized as using either synthetic or natural images, with or without specified ground truth, [18, 63]. When the ground truth is given, performance evaluation can be readily carried out by comparing detected contours with the ground truth edges. Although synthetic images allow precise objective definition of ground truth and seem appropriate for any performance evaluation criterion, the conclusions drawn in most of the cases are not easily extrapolated for natural scenes [17]. Additional qualitative metrics such as smoothness, continuity, thinness, which may sometimes be computed in absence of the ground truth, do not always properly reflect performance [64]. For these reasons, most of the current evaluation methods use natural image scenes with an associated ground truth specified by a human observer [26, 40, 62, 64–66]. For a comprehensive list of performance evaluation methods for edge detection we refer to [66].

Different human observers produce different ground truth contour images for the same input image and a given pixel can be marked by some observers as a contour pixel (of value 1) and by others as a texture or background pixel (of value 0). One way to deal with this fact is to use a superposition of the binary contour maps produced by different observers [62]. Here we apply an alternative approach in which we asked 8 observers to mark the contours they see. Based on their contour drawings we defined a weighted ground truth in which a pixel (x, y) is assigned a weight $\gamma(x, y) = 1$ if 5 or more out of the 8 observers drew a contour pixel within a distance of 2 pixels and weight $\gamma(x, y) = 1/3$ if this was done by 3 or 4 observers, Figure 24.

Let DC be the set of points for which a given contour detection operator outputs "1" and let GT be the set of points for which $\gamma(x, y) > 0$. We define generalized *recall R* and
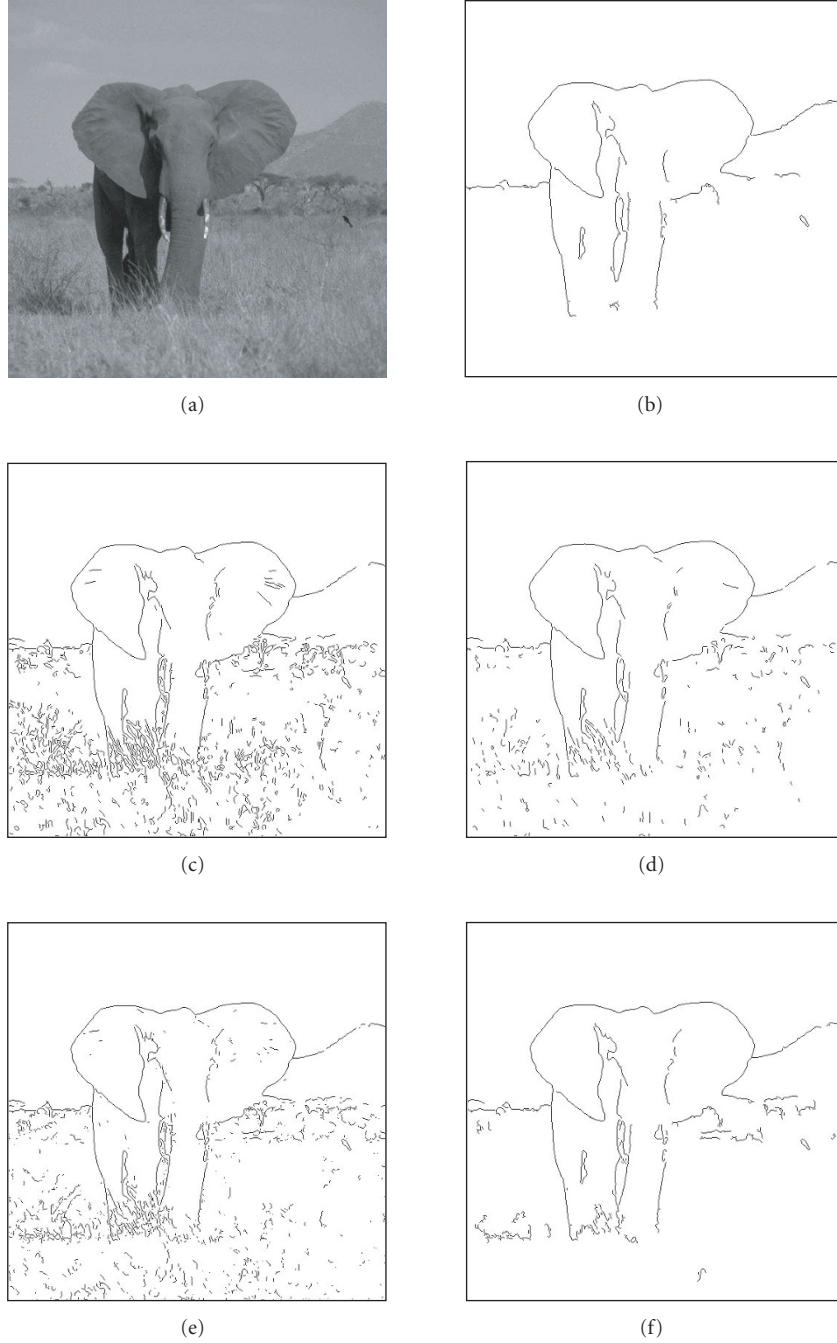
(a)

(b)

(c)

(d)

(e)

(f)

Figure 9: "Elephant" (512×512 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

precision P as follows:

$$\text{recall} = \frac{\sum_{(x,y)\in \text{DC}\cap \text{GT}} \gamma(x, y)}{\sum_{(x,y)\in \text{GT}} \gamma(x, y)},$$

$$\text{precision} = \frac{\sum_{(x,y)\in \text{DC}\cap \text{GT}} \gamma(x, y)}{\text{card}\{\text{DC}\}},$$

(33)

where $card\,(X)$ is the number of elements of the set X.

In order to compensate for small shifts of contours detected by an operator from ground truth contours, the intersection of GT and DC is computed as proposed in [30].

4.2.2. Results

For each of the five algorithms discussed above we evaluated R and P for a set of 40 images and computed the averages
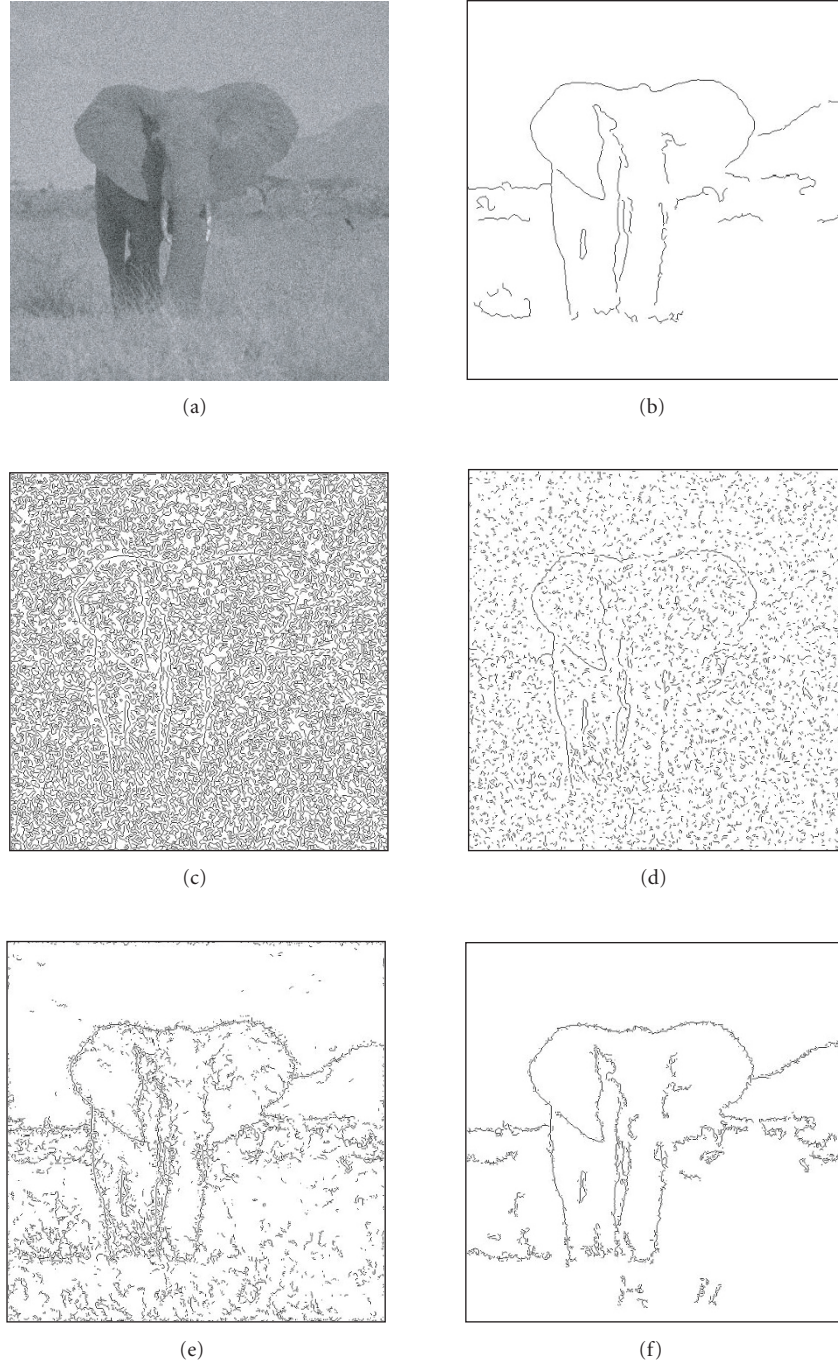
FIGURE 10: Noisy elephant (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

of the obtained values. We computed such averages for the noiseless images (SNR = ∞) and for noisy versions of these images (SNR = 10 dB, 13 dB, and 16 B).

Figure 25(a) shows the results obtained whereby for each algorithm and SNR we used the optimal parameter values as specified in the respective references. As we can see, for all SNR values our approach gives the best performance in terms

of texture suppression (e.g., $P = 0.74$ for noiseless images, SNR = ∞), while keeping a sufficient percentage of detected true contour pixels ($R = 0.4$ for noiseless images, SNR = ∞). By comparing our approach with the best of the other existing algorithms (MSSI) that we consider here, we see that the proposed approach gives a significant advantage in terms of increased precision (by a factor of at least 2) while paying a
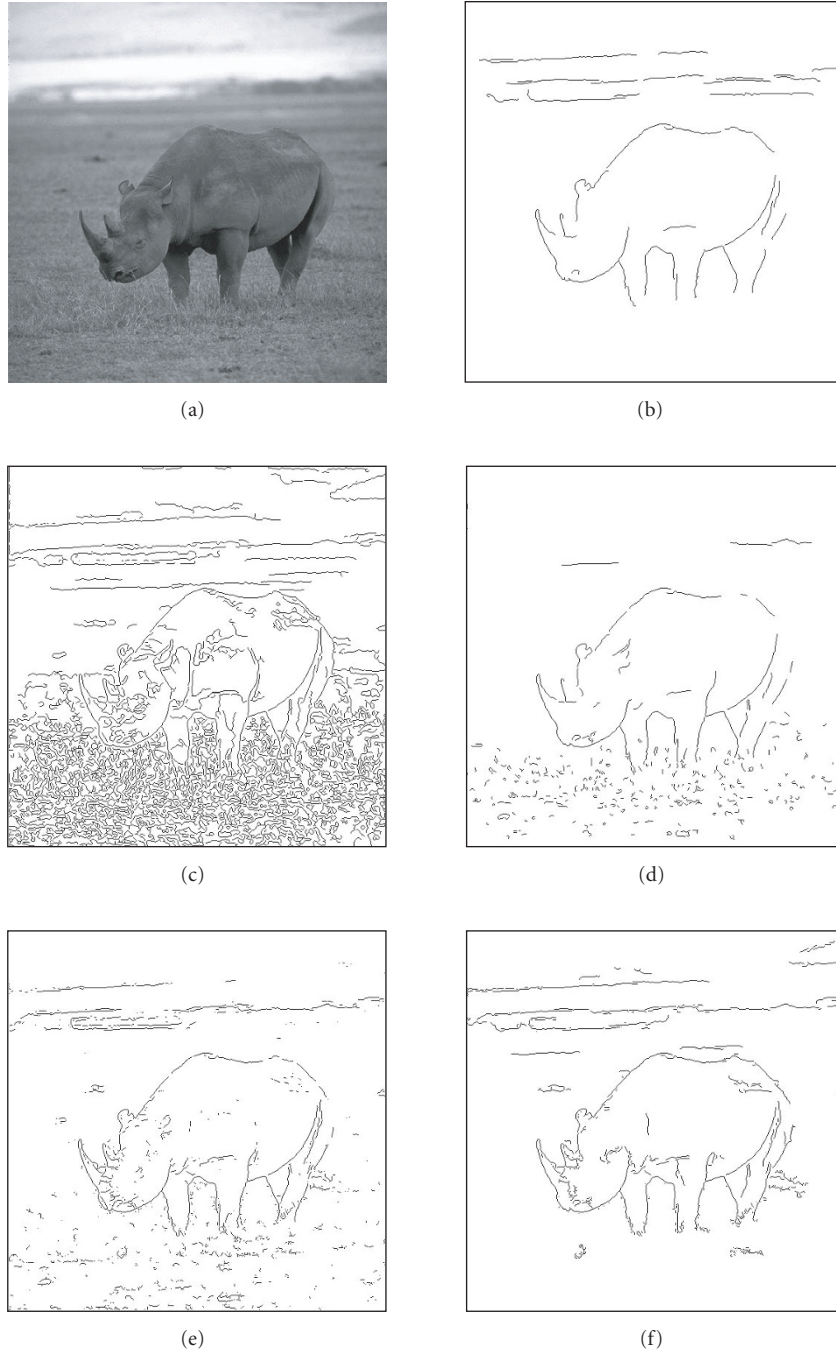
FIGURE 11: "Rhino" (512×512 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

small price in terms of decreased recall (by only 20%–25%). Canny algorithm gives good recall at the expense of a very bad precision which means that in the resulting binary map the true contours are buried in a binary texture noise (see, e.g., Figures 11(c), 13(c), 15(c), and 17(c)).

The proposed approach has also the best performance in terms of noise rejection, since the value of the preci-

sion achieved by our approach with the most noisy test images is higher than the one achieved by the other approaches for noiseless test images. The worst performance in terms of precision is exhibited by the Canny edge detector because this algorithm deploys no texture/noise suppression mechanism. The modification of CARTOON and SSSI provides approximately the same improvement, while
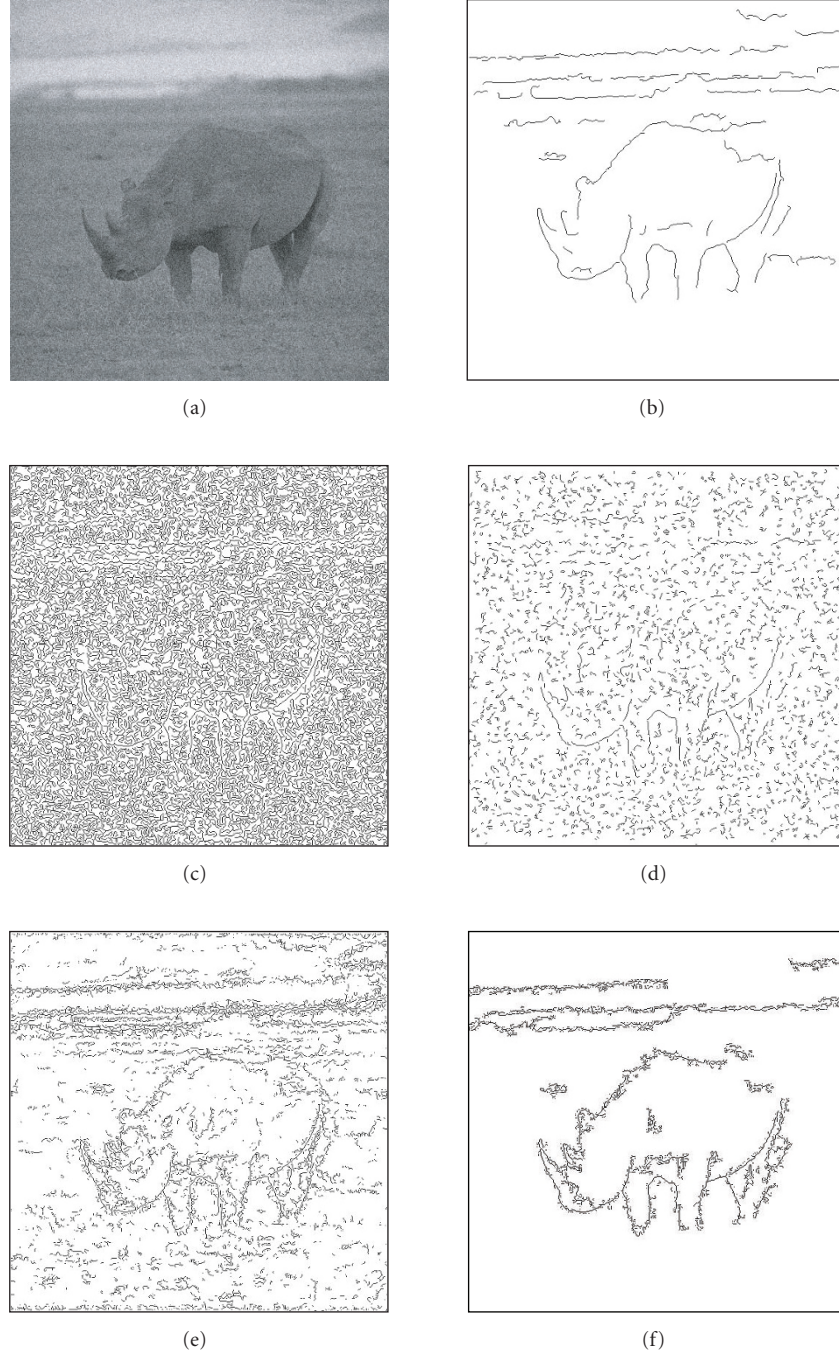
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 12: Noisy Rhino (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

SSSI performs slightly better for noiseless images but it is more sensitive to noise. The use of the surround inhibition scheme [54] in a multiscale framework as done in [46] provides a considerable performance increase in terms of precision.

It is worth pointing out that the relatively high values of precision achieved by means of the proposed algorithm are obtained at the expense of the loss of some contour details. In this respect we have to note that the values of recall of the other algorithms are over estimated due to noise or texture pixels that are not suppressed and lie near true contours. This is confirmed by the fact that the value of the recall increases as the SNR decreases for all considered algorithms but the one proposed here.
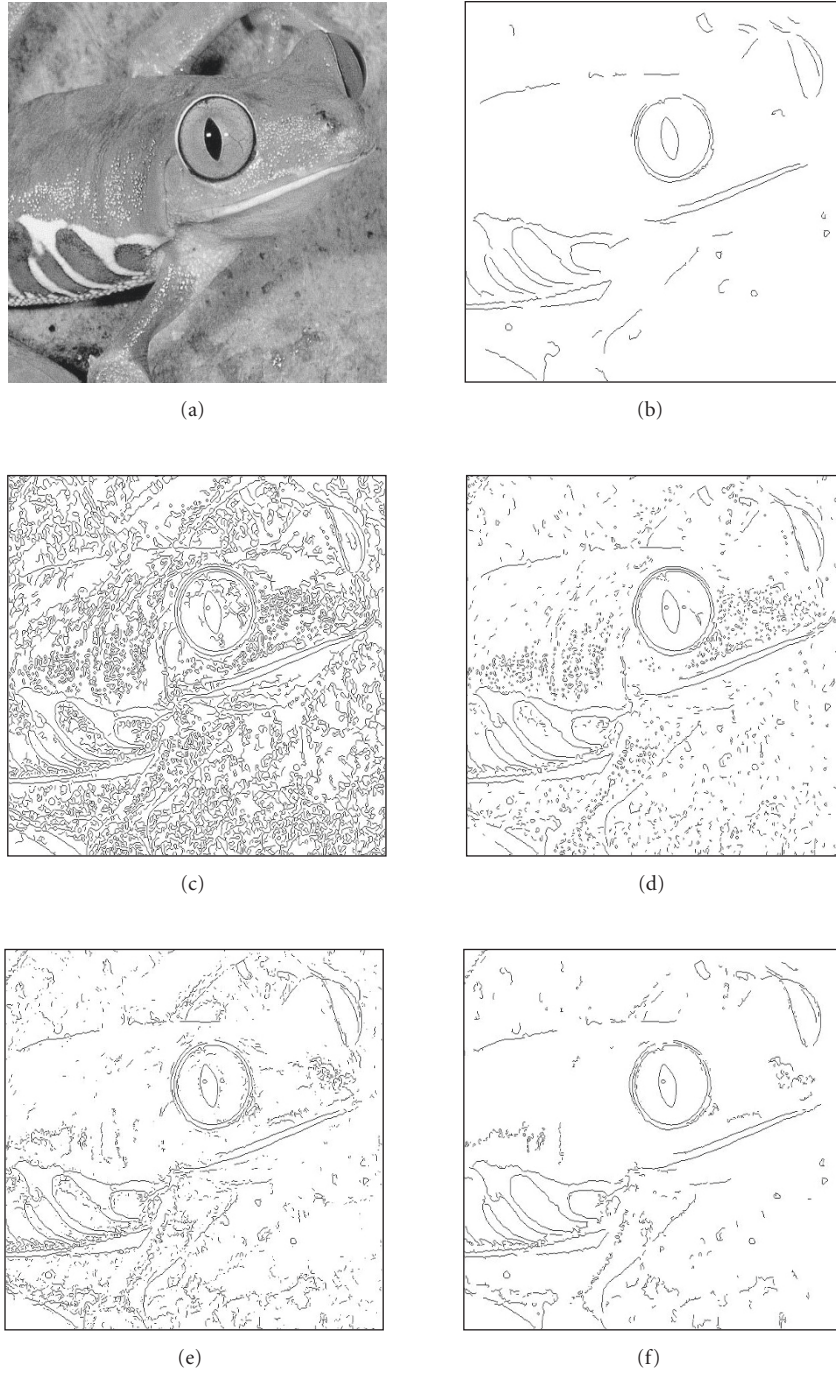
FIGURE 13: "Frog" (564×496 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

The fact that with the proposed approach the recall value decreasing with increasing noise is in agreement with the property of the human visual system that tends to detect contours less effectively in presence of noise [67].

Figure 25(b) shows the performance behavior of the proposed algorithm for various values of the threshold $G_{\min}$

($G_{\min} = 2.5, 5, 6, 7, 8, 9, 10, 12.5$). Finally, Figure 25(c) shows the performance of the Canny edge detector and the proposed algorithm for a wide range of values of the respective binarization thresholds of these algorithms. Such plots have been deployed in recent literature to study the performance of boundary detection algorithms for a broad range

(a)

(b)
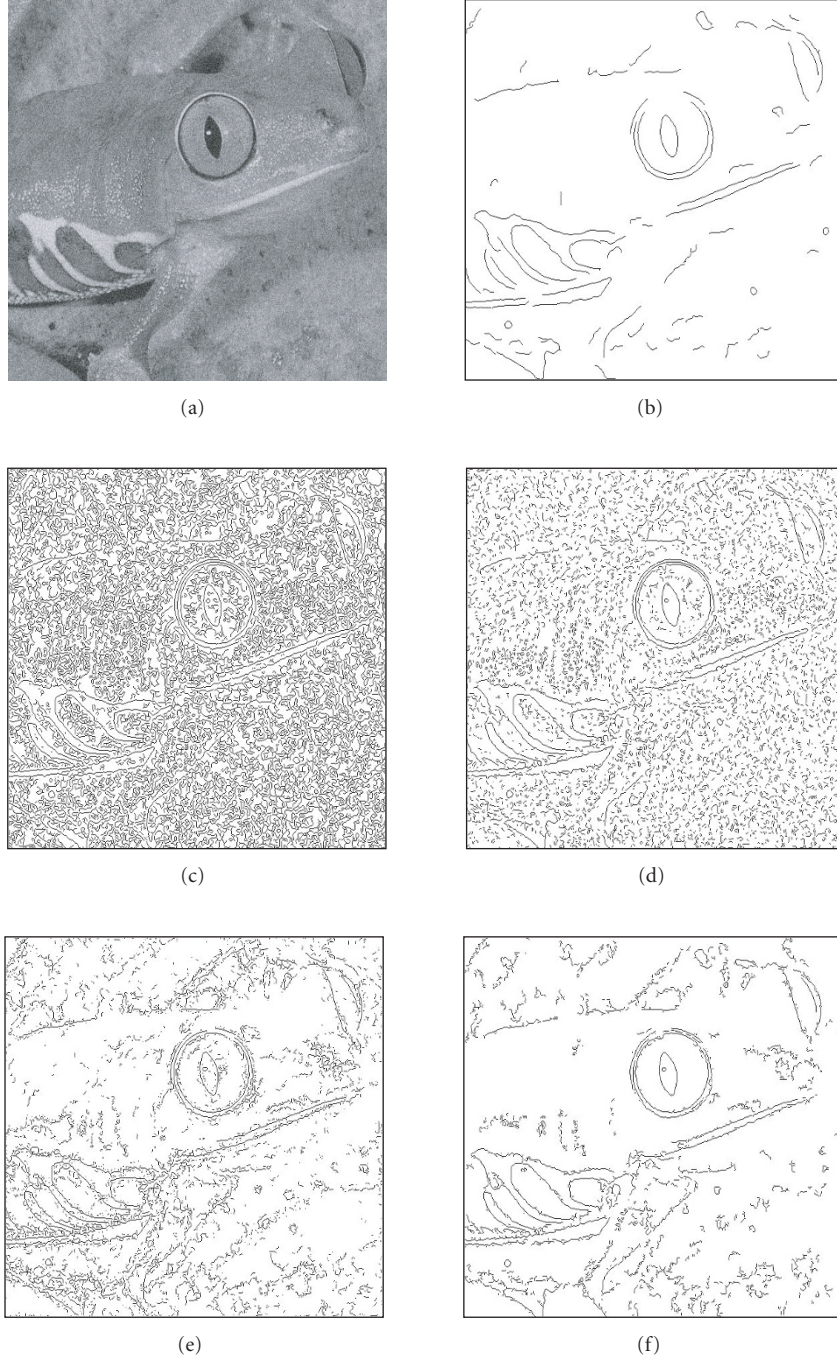

(c)

(d)

(e)

(f)

FIGURE 14: Noisy frog (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

of threshold values that leads to a broad range of recall and precision values [62]. In this respect, we think that it is worth to explore the threshold space only for values that lead to some reasonably large values of the precision and recall as illustrated in Figure 25(c).

We studied also how the performance is influenced by the parameters related to the computation of the inhibition term T_σ introduced in Section 3: the number of orientations N_ϕ, the distance a between the two half-rings r^+ and r^-, and the ratio k between the standard deviations of the two Gaussian functions of the inhibition weighting function (14). Our experiments have shown that the overall performance is not sensitive to the choice of these parameters. For N_ϕ ranging between 4 and 10 ($N_\phi = 4, 6, 8, 10$), the relative
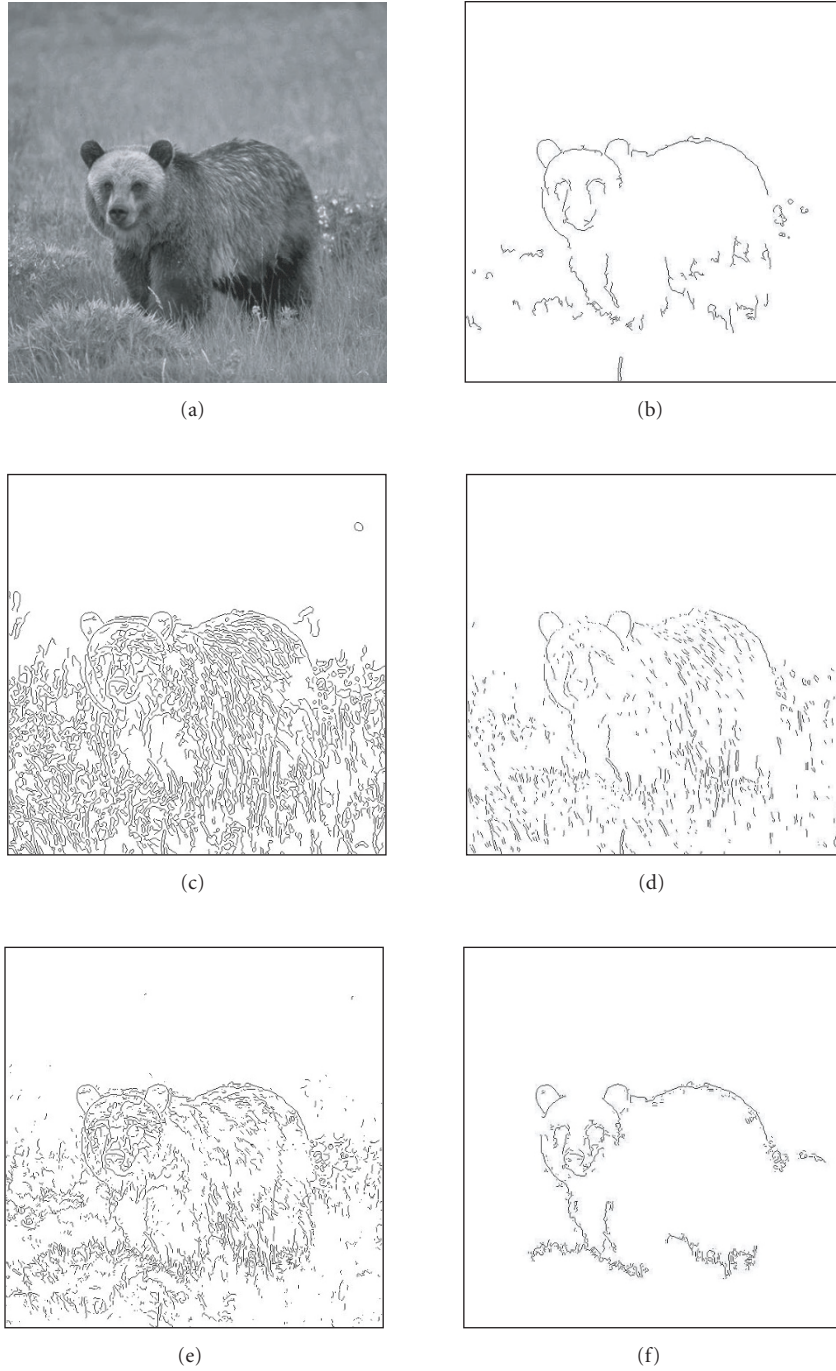
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 15: "Bear1" (512×512 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

variations of the performance indicators are only 0.62% for recall and 0.88% for precision. We considered the spacing a between the two half-rings as a fraction η of the inner radius ρ_0 and experimented with values of η ranging between 1 and 2 ($\eta = 1, 1.25, 1.5, 1.75, 2$). The relative variations of the performance indicators are of 0.34% for recall and 0.65% for precision. Finally, we considered values of k ranging between 3 and 6 ($k = 3, 4, 5, 6$) and found the relative variations of the performance indicators to be only 0.23% for recall and 1.76% for precision. To summarize, our experiments showed that the proposed algorithm is robust to variations of these parameters.
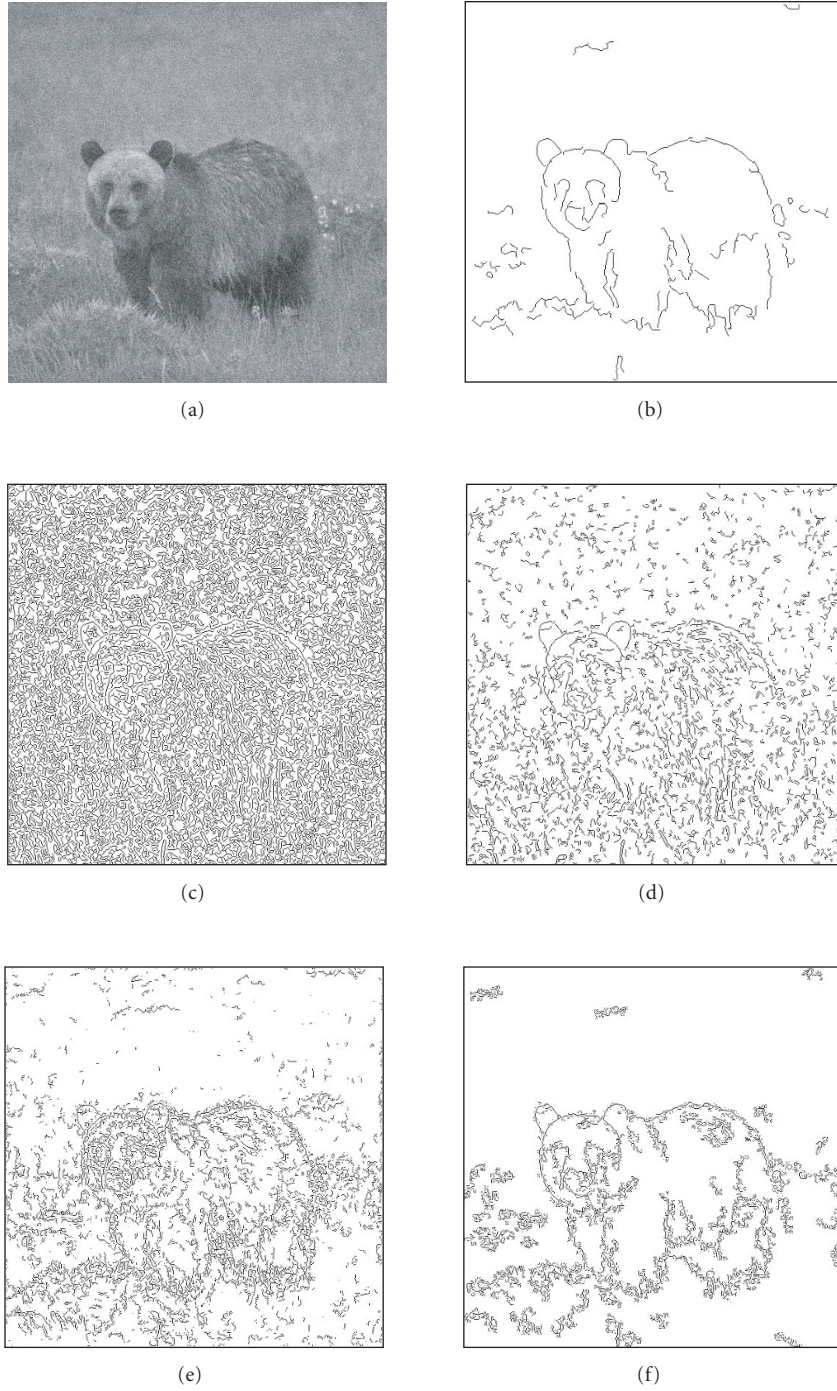
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 16: Noisy bear1 (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

5. SUMMARY AND CONCLUSIONS

In this paper we proposed a contour detection algorithm that outperforms standard edge detectors that react to all the local luminance changes, irrespective of whether they are due to object contours or due to natural textures like grass, foliage, water, and so forth. Specifically, the method we presented relies on different characteristics of the HVS. Inspired by psychological and neurophysiological studies, we incorporated in our scheme surround inhibition of texture that
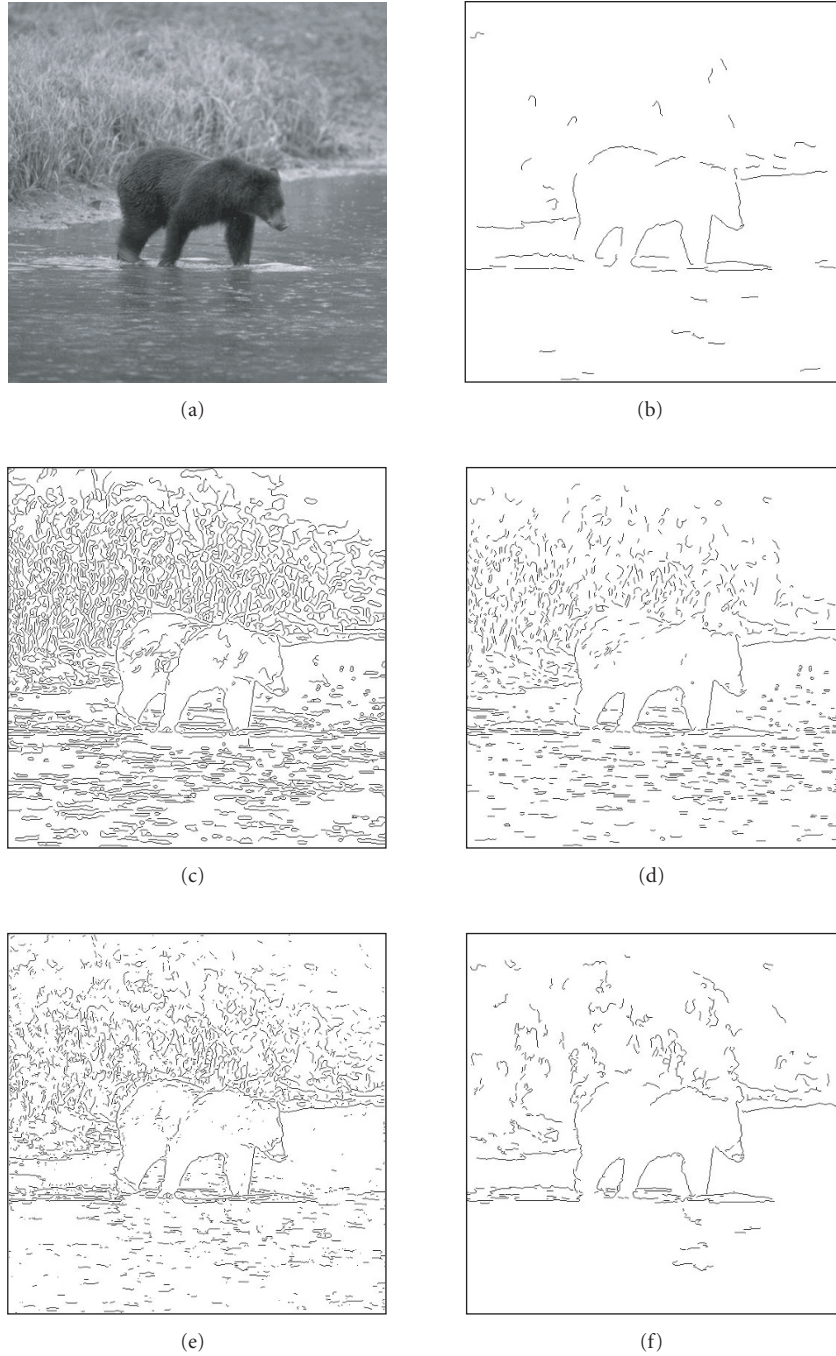
FIGURE 17: "Bear2" (512×512 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

does not affect isolated edges but that inhibits edges that are surrounded by other edge stimuli. Compared with previous inhibition schemes [27, 30, 54], the method we propose does not suffer the problem of self-inhibition of true contours and, therefore, allows for a stronger and thus more effective inhibition of texture. The design of the new inhibition scheme involves a few parameters for which we however

found that the selection of their values does not have significant effect on the performance of the algorithm so that their values can be fixed.

In order to make our algorithm robust to noise, we apply a Bayesian denoising step at each resolution before the surround inhibition step. It consists of the optimal MMSE estimator of the image gradient in additive noise for which
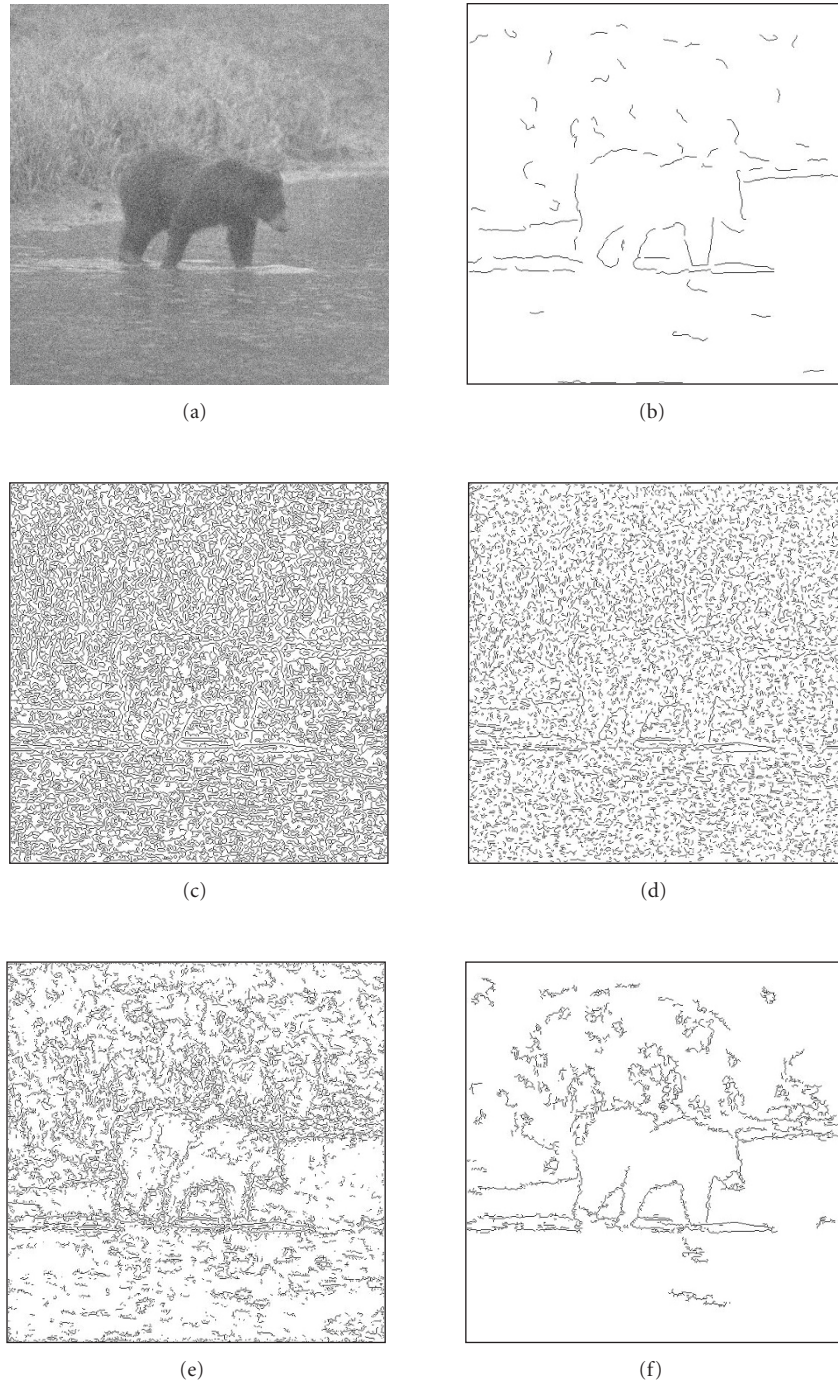
FIGURE 18: Noisy bear2 (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

a closed form is given. The a priori first-order p.d.f. of both the signal and the additive noise is assumed Gaussian scale mixture, according to previous studies of the statistics of the wavelet coefficients of natural images.

For binarization we use a thresholding technique that is based on weights computed for the connected and thinned

(by nonmaxima suppression) components of the surround inhibited and denoised gradient magnitude. We found out that for obtaining good results it is essential that this thresholding technique is applied in combination with surround inhibition which cuts into small pieces long connected edge components that are due to noise and texture.
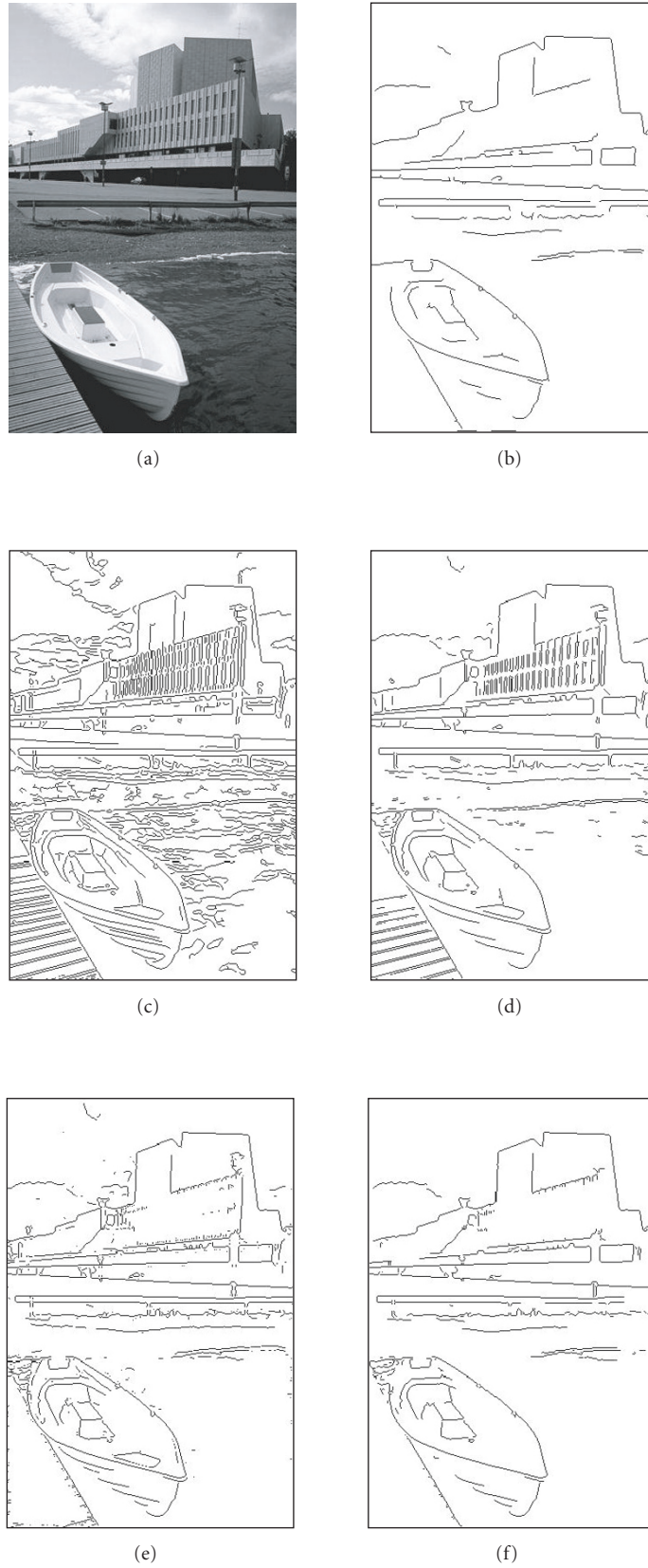
(a)

(b)

(c)

(d)

(e)

(f)

Figure 19: Boat (321 × 481 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].
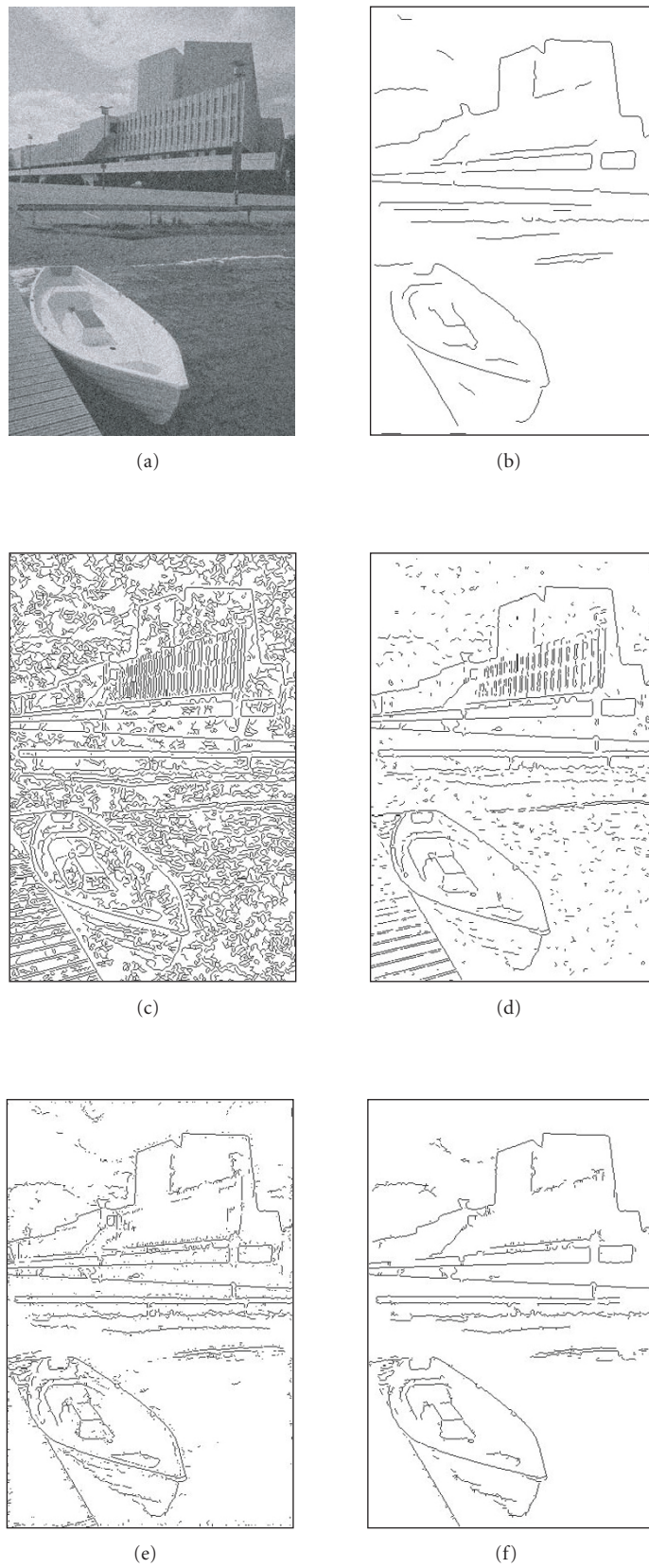
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 20: Noisy boat (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

FIGURE 21: "Man and woman" (480×320 pixels): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].
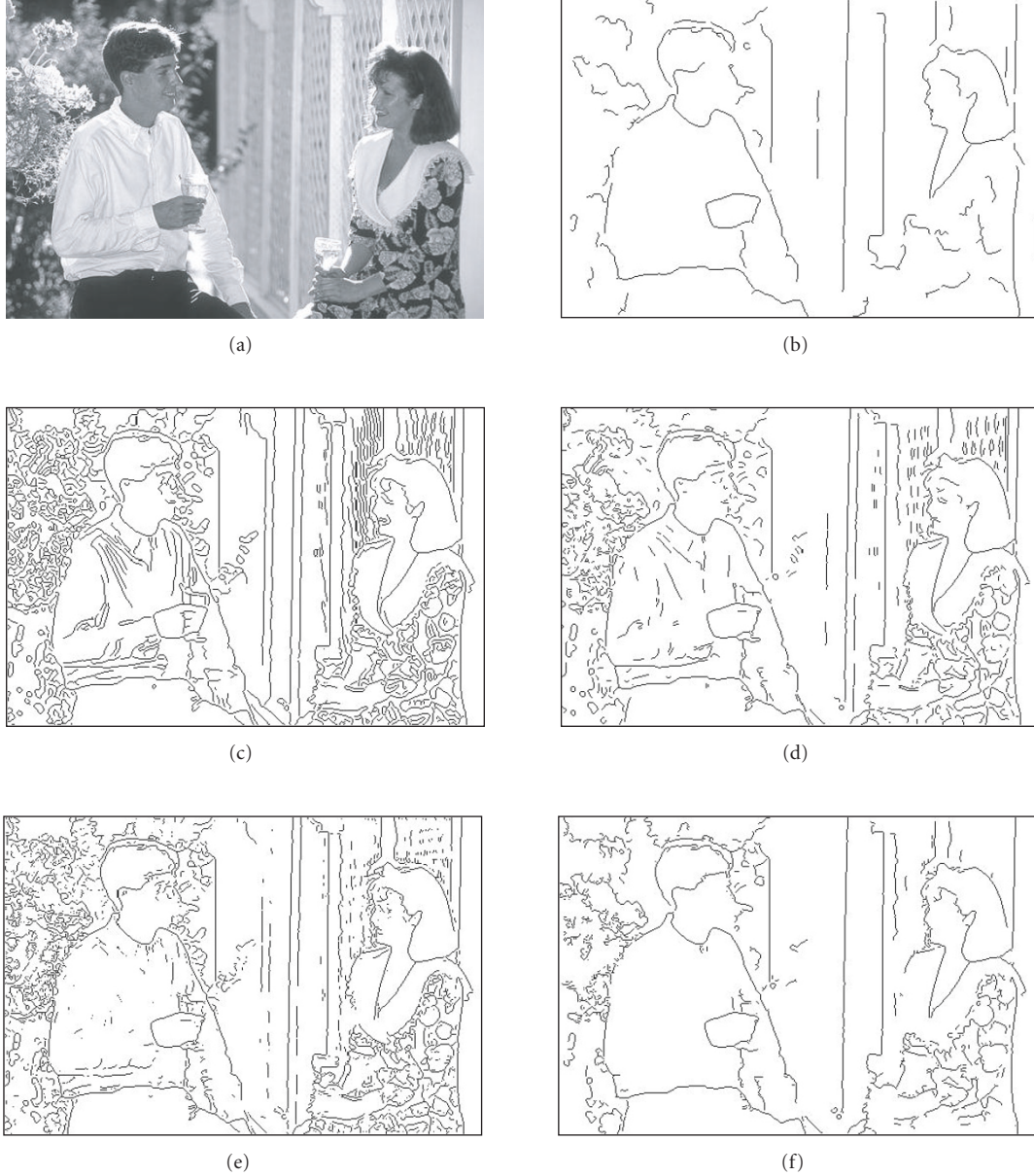
The second characteristic of the HVS taken into account is that, as pointed out by psychophysical experiments, the visual information in different frequency bands is processed separately. Following these findings, we perform contour detection in a multiresolution framework. Object contours can be discriminated from texture edges because the former are present at all scales (above a given scale that is determined by the contour blur), while the latter appear only at the finer scales. We use this fact and combine the binary contour maps obtained for different scales in such a way that texture is eliminated while contours and junctions are retained.

The entire algorithm can be easily implemented by computing convolutions, applying zero-memory nonlinearities and basic morphological operations, whereby convolutions are the most computationally demanding operations and have computational complexity $O(N_I \log N_I)$, where N_I is the linear size of the image. All the other operations can be done in linear time, therefore the overall algorithm complexity is $O(N_I \log N_I)$. To summarize, with reference to Figure 1, the gradient computation requires two convolution operations, the Bayesian denoising operation is computable in linear time, the computation of the inhibition term requires $2N\phi$
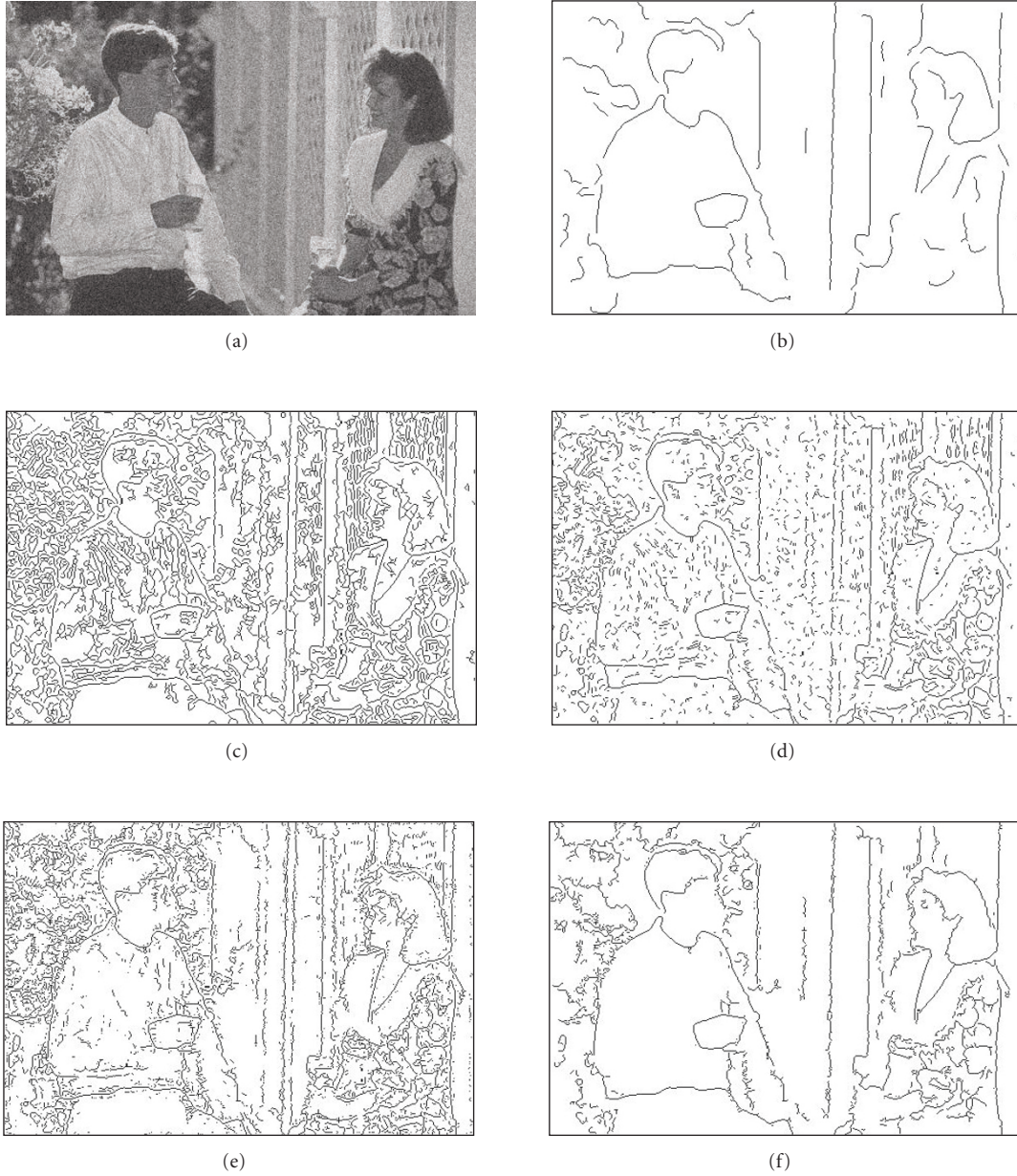
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 22: Noisy man and woman (SNR = 13 dB): (a) test image and contours detected using (b) the proposed approach, (c) the Canny edge detector [1], (d) single-scale surround inhibition [54], (e) a modification of the multiscale edge detector CARTOON [45], and (f) a multiscale contour detector with surround inhibition [46].

convolutions, where $N\phi$ is the number of orientations, and the binarization is computable in linear time. Therefore, for each scale $2N\phi + 2$ convolutions are required.

We tested the proposed algorithm on a set of 40 images, both noiseless and corrupted by additive white noise (with SNR = 16 dB, 13 dB, 10 dB), and compared it with other four existing contour detectors. Both visual inspection of the results and quantitative comparison with weighted ground truths lead us to the main conclusion of this paper that *the proposed algorithm is superior to other known*

algorithms in terms of amount of texture suppressed, amount of detected contours and their cleanness and robustness to noise. The effective suppression of texture in the resulting binary contour maps is a very important aspect of the proposed algorithm, because modern shape recognition algorithms that use contour information (see, e.g., [68]) rely on clean contour maps that are not corrupted by texture noise [69–71].

Similar to other multiresolution contour detection approaches (see, e.g., [45, 46, 58, 72, 73]) we rely on the
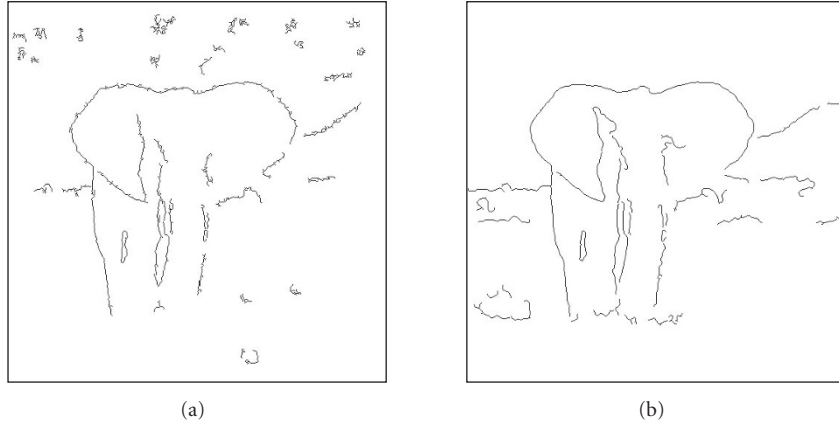
(a)

(b)

FIGURE 23: Contours detected using the proposed algorithm for the noisy elephant image (a) without and (b) with Bayesian denoising.
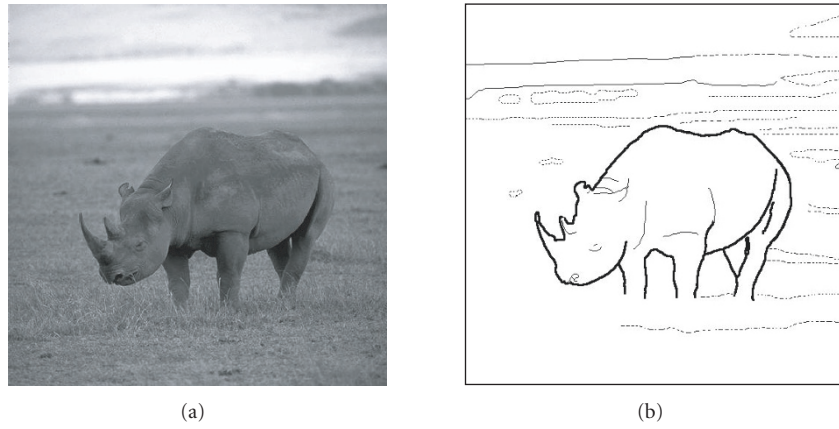

(a)

(b)

FIGURE 24: (a) An input image and (b) its weighted ground truth. The thick lines specify contours drawn by at least 5 (out of 8) observers ($\gamma(x, y) = 1$). The thin lines are contours specified by 3 or 4 observers ($\gamma(x, y) = 1/3$). The dotted lines are edges considered to be contours by less than 3 observers; the latter are not used in the performance evaluation.
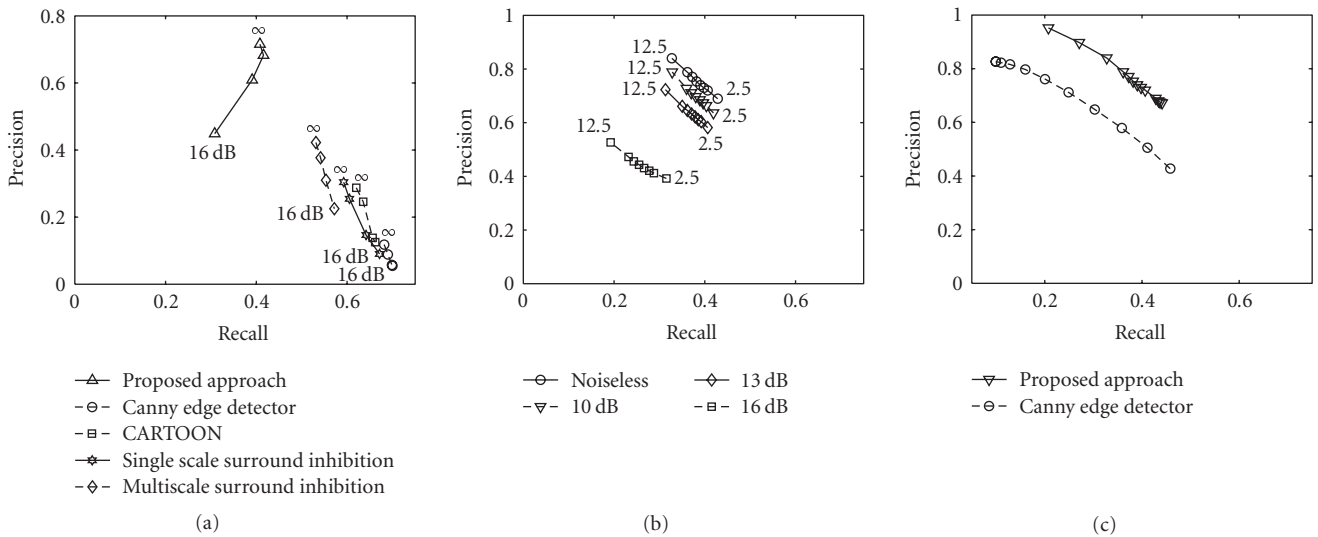

(a)

(b)

(c)

FIGURE 25: (a) Plots of precision versus recall for different algorithms and varying SNR values. (b) Precision versus recall of the proposed algorithm for different SNRs and threshold values. (c) Precision-recall plots of the Canny edge detector and the proposed algorithm for various values of their respective threshold parameters.

assumption that object contours are present in the results at all resolutions, while texture appears only at the finer scales. More precisely, the former part of this assumption holds for sharp contours only. Blurred contours are detected more effectively at coarser scales. Though the results are encouraging for a wide set of test images with different structure, contrast and texture types, such an assumption might be a limitation as it was not verified for all considered resolutions. However, such a limitation is shared by all other edge detection techniques and finding a universal optimal solution for this problem is still an open issue.

ACKNOWLEDGMENT

REFERENCES

[1] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[2] W. Frei and C.-C. Chen, "Fast boundary detection: a generalization and a new algorithm," *IEEE Transactions on Computers*, vol. 26, no. 10, pp. 988–998, 1977.

[3] E. C. Hildreth, "The detection of intensity changes by computer and biological vision systems," *Computer Vision, Graphics, and Image Processing*, vol. 22, no. 1, pp. 1–27, 1983.

[4] J.-B. Martens, "Local orientation analysis in images by means of the Hermite transform," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1103–1116, 1997.

[5] R. Nevatia and K. R. Babu, "Linear feature extraction and description," *Computer Vision, Graphics, and Image Processing*, vol. 13, no. 3, pp. 257–269, 1980.

[6] P. H. Gregson, "Using angular dispersion of gradient direction for detecting edge ribbons," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 7, pp. 682–696, 1993.

[7] O. A. Zuniga and R. M. Haralick, "Integrated directional derivative gradient operator," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 17, no. 3, pp. 508–517, 1987.

[8] G. Chen and Y. H. H. Yang, "Edge detection by regularized cubic B-spline fitting," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 4, pp. 636–643, 1995.

[9] S. Ghosal and R. Mehrotra, "Detection of composite edges," *IEEE Transactions on Image Processing*, vol. 3, no. 1, pp. 14–25, 1994.

[10] R. M. Haralick, "Digital step edges from zero crossing of second directional derivatives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 1, pp. 58–68, 1984.

[11] V. S. Nalwa and T. O. Binford, "On detecting edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 699–714, 1986.

[12] M. C. Morrone and D. C. Burr, "Feature detection in human vision: a phase-dependent energy model," *Proceedings of the Royal Society of London. Series B, Biological sciences*, vol. 235, no. 1280, pp. 221–245, 1988.

[13] T. C. Folsom and R. B. Pinter, "Primitive features by steering, quadrature, and scale," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1161–1173, 1998.

[14] F. Heitger, "Feature detection using suppression and enhancement," Tech. Rep. TR-163, Communication Technology Laboratory, Swiss Federal Institute of Technology, Zurich, Switzerland, 1995.

[15] P. Kovesi, "Image features from phase congruency," *Videre: Journal on Computer Vision Research*, vol. 1, no. 3, pp. 2–27, 1999.

[16] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, vol. 6, no. 5, pp. 303–313, 1987.

[17] Y. T. Zhou, V. Venkateswar, and R. Chellappa, "Edge detection and linear feature extraction using a 2-D random field model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 1, pp. 84–95, 1989.

[18] S. Ando, "Image field categorization and edge/corner detection from gradient covariance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 2, pp. 179–190, 2000.

[19] P. Meer and B. Georgescu, "Edge detection with embedded confidence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1351–1365, 2001.

[20] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 421–432, 1998.

[21] Y. Chen, C. A. Z. Barcelos, and B. A. Mair, "Smoothing and edge detection by time-varying coupled nonlinear diffusion equations," *Computer Vision and Image Understanding*, vol. 82, no. 2, pp. 85–100, 2001.

[22] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[23] J. Weickert, "A review of nonlinear diffusion filtering," in *Scale-Space Theory in Computer Vision*, vol. 1252 of *Lecture Notes in Computer Science*, pp. 3–28, Springer, New York, NY, USA, 1997.

[24] W.-Y. Ma and B. S. Manjunath, "EdgeFlow: a technique for boundary detection and image segmentation," *IEEE Transactions on Image Processing*, vol. 9, no. 8, pp. 1375–1388, 2000.

[25] J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 7–27, 2001.

[26] B. S. Manjunath and R. S. Chellappa, "A unified approach to boundary perception: edges, textures, and illusory contours," *IEEE Transactions on Neural Networks*, vol. 4, no. 1, pp. 96–108, 1993.

[27] C. Grigorescu, N. Petkov, and M. A. Westenberg, "Contour detection based on nonclassical receptive field inhibition," *IEEE Transactions on Image Processing*, vol. 12, no. 7, pp. 729–739, 2003.

[28] Z. Li, "Visual segmentation by contextual influences via intracortical interactions in the primary visual cortex," *Network: Computation in Neural Systems*, vol. 10, no. 2, pp. 187–212, 1999.

[29] N. Petkov and P. Kruizinga, "Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells," *Biological Cybernetics*, vol. 76, no. 2, pp. 83–96, 1997.

[30] N. Petkov and M. A. Westenberg, "Suppression of contour perception by band-limited noise and its relation to nonclassical receptive field inhibition," *Biological Cybernetics*, vol. 88, no. 3, pp. 236–246, 2003.

[31] D. Marr and E. C. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B, Biological sciences*, vol. 207, no. 1167, pp. 187–217, 1980.

[32] D. J. Field, A. Hayes, and R. F. Hess, "Contour integration by the human visual system: evidence for a local "association field"," *Vision Research*, vol. 33, no. 2, pp. 173–193, 1993.

[33] G. Kanizsa, *Organization in Vision: Essays on Gestalt Perception*, Praeger, New York, NY, USA, 1979.

[34] H. C. Nothdurft, "Texture segmentation and pop-out from orientation contrast," *Vision Research*, vol. 31, no. 6, pp. 1073–1078, 1991.

[35] J. A. Solomon and D. G. Pelli, "The visual filter mediating letter identification," *Nature*, vol. 369, no. 6479, pp. 395–397, 1994.

[36] M. K. Kapadia, G. Westheimer, and C. D. Gilbert, "Spatial distribution of contextual interactions in primary visual cortex and in visual perception," *Journal of Neurophysiology*, vol. 84, no. 4, pp. 2048–2062, 2000.

[37] J. J. Knierim and D. C. van Essen, "Neuronal responses to static texture patterns in area V1 of the alert macaque monkey," *Journal of Neurophysiology*, vol. 67, no. 4, pp. 961–980, 1992.

[38] H. C. Nothdurft, J. L. Gallant, and D. C. van Essen, "Response modulation by texture surround in primate area V1: correlates of "popout" under anesthesia," *Visual Neuroscience*, vol. 16, no. 1, pp. 15–34, 1999.

[39] H. E. Jones, K. L. Grieve, W. Wang, and A. M. Sillito, "Surround suppression in primate V1," *Journal of Neurophysiology*, vol. 86, no. 10, pp. 2011–2028, 2001.

[40] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.

[41] B. Julesz, "Visual pattern discrimination," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 84–92, 1962.

[42] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *Journal of Physiology*, vol. 197, no. 3, pp. 551–566, 1968.

[43] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 2091–2110, 1989.

[44] M. C. Morrone and D. C. Burr, "Capture and transparency in coarse quantized images," *Vision Research*, vol. 37, no. 18, pp. 2609–2629, 1997.

[45] W. Richards, H. K. Nishihara, and B. Dawson, "CARTOON: a biologically motivated edge detection algorithm," in *Natural Computation*, W. Richards, Ed., MIT A.I. Memo no. 668, chapter 4, pp. 55–69, MIT Press, Cambridge, Mass, USA, 1988.

[46] G. Papari, P. Campisi, N. Petkov, and A. Neri, "A multiscale approach to conour detection by texture suppression," in *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, vol. 6064 of *Proceedings of the SPIE*, pp. 107–118, San Jose, Calif, USA, January 2006.

[47] M. J. Wainwright, E. P. Simoncelli, and A. S. Willsky, "Random cascades on wavelet trees and their use in analyzing and modeling natural images," *Applied and Computational Harmonic Analysis*, vol. 11, no. 1, pp. 89–123, 2001, special issue on wavelet applications.

[48] E. P. Simoncelli, "Statistical modeling of photographic images," in *Handbook of Image and Video Processing*, A. Bovik, Ed., chapter 4.7, pp. 431–441, Academic Press, Boston, Mass, USA, 2nd edition, 2005.

[49] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 886–902, 1998.

[50] M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300–303, 1999.

[51] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Adaptive Wiener denoising using a Gaussian scale mixture model in the wavelet domain," in *Proceedings of the 8th IEEE International Conference on Image Processing (ICIP '01)*, vol. 2, pp. 37–40, Thessaloniki, Greece, October 2001.

[52] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–71, 2000.

[53] J. Jacovitti and A. Neri, "Anisotropic wavelet thresholding for Bayesian image denoising," in *Proceedings of the 11th European Signal Processing Conference (EUSIPCO '02)*, vol. 3, pp. 267–270, Toulouse, France, September 2002.

[54] C. Grigorescu, N. Petkov, and M. A. Westenberg, "Contour and boundary detection improved by surround suppression of texture edges," *Image and Vision Computing*, vol. 22, no. 8, pp. 609–622, 2004.

[55] D.-K. Xiao, S. Raiguel, V. Marcar, J. Koenderink, and G. A. Orban, "Spatial heterogeneity of inhibitory surrounds in the middle temporal visual area," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 24, pp. 11303–11306, 1995.

[56] H. J. A. M. Heimans, *Morphological Image Operators*, Academic Press, Boston, Mass, USA, 1994.

[57] H. J. A. M. Heijmans, "Connected morphological operators for binary images," *Computer Vision and Image Understanding*, vol. 73, no. 1, pp. 99–120, 1999.

[58] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–154, 1998.

[59] K.-H. Liang, T. Tjahjadi, and Y.-H. Yang, "Bounded diffusion for multiscale edge detection using regularized cubic B-spline fitting," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 29, no. 2, pp. 291–297, 1999.

[60] A. Ding and A. Goshtasby, "On the canny edge detector," *Pattern Recognition*, vol. 34, no. 3, pp. 721–725, 2001.

[61] C. F. Olson, "Adaptive-scale filtering and feature detection using range data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 9, pp. 983–991, 2000.

[62] D. R. Martin, C. C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, 2004.

[63] K. Bowyer, C. Kranenburg, and S. Dougherty, "Edge detector evaluation using empirical ROC curves," *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 77–103, 2001.

[64] L. Kitchen and A. Rosenfeld, "Edge evaluation using local edge coherence," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 11, no. 9, pp. 597–605, 1981.

[65] M. Shin, D. Goldgof, and K. W. Bowyer, "An objective comparison methodology of edge detection algorithms for structure from motion task," in *Empirical Evaluation Techniques in Computer Vision*, pp. 235–254, IEEE Press, New York, NY, USA, 1998.

[66] M. C. Shin, D. B. Goldgof, and K. W. Bowyer, "Comparison of edge detector performance through use in an object recognition task," *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 160–178, 2001.

[67] N. Petkov and M. A. Westenberg, "Suppression of contour perception by band-limited noise and its relation to nonclassical receptive field inhibition," *Biological Cybernetics*, vol. 88, no. 3, pp. 236–246, 2003.

[68] C. Grigorescu and N. Petkov, "Distance sets for shape filters and shape recognition," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1274–1286, 2003.

[69] A. Ghosh and N. Petkov, "Robustness of shape descriptors to incomplete contour representations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1793–1804, 2005.

[70] A. Ghosh and N. Petkov, "A cognitive evaluation procedure for contour based shape descriptors," *International Journal of Hybrid Intelligent Systems*, vol. 2, no. 4, pp. 237–252, 2005.

[71] A. Ghosh and N. Petkov, "Effect of high curvature point deletion on the performance of two contour based shape recognition algorithms," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 20, no. 6, pp. 913–924, 2006.

[72] F. Bergholm, "Edge focusing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 6, pp. 726–741, 1987.

[73] A. Goshtasby, "On edge focusing," *Image and Vision Computing*, vol. 12, no. 4, pp. 247–256, 1994.

Giuseppe Papari was born in Rome in 1979. He got his Master's degree in electrical engineering in 2003 from "Universitá degli Studi di Roma Tre" (Third University of Rome). Since June 2004, he has been Ph.D. student at the Computing Science Departement of the University of Groningen. His main research interests are contour detection, image segmentation, multiresolution analysis, clustering, pattern recognition, morphological analysis. He has published several papers about biologically motivated contour detection and perceptual grouping, one of which was invited. In 2006, he received one of the four IBM Awards on the IEEE International Conference on Image Processing.

Patrizio Campisi received the "Laurea" degree in electrical engineering (summa cum laude) from the University of Roma "La Sapienza," Roma, Italy, and the Ph.D. degree in electrical engineering from the University of Roma "Roma TRE," Italy. Currently he is an Associate Professor with the Department of Applied Electronics, Università degli Studi "Roma TRE," Roma, Italy. He is Coeditor with Karen Egiazarian of the book "*Blind Image Deconvolution: Theory and Applications*," CRC (to appear in 2007). He is also Coauthor with G. Papari et al. of a 2006 ICIP Best Student Paper Award for the paper titled "Contour detection by multiresolution surround inhibition." His research interests are in the area of digital signal and image processing with applications to multimedia. Specifically, he has been working on image deconvolution image restoration, image analysis, texture coding, texture synthesis, texture classification, watermarking (2D images, stereo images, videos), data hiding, and biometrics.

Nicolai Petkov is Professor of computer science and Head of the Research Institute of Mathematics and Computing Science of the University of Groningen. He got the Dr.sc.techn. degree in computer engineering (informationstechnik) from Dresden University of Technology. Prior to joining the University of Groningen in 1991, he held research positions at the University of Wuppertal, the University of Erlangen-Nürnberg, the Academy of Sciences in Berlin, and Dresden University of Technology. In 1989 he was awarded an Alexander von Humboldt scholarship of the Federal Republic of Germany. He is the author of two books and over 100 scientific papers. He is Member of the editorial boards of various journals. His current research focuses on computer simulations and understanding of the visual system of the brain and on the development of effective computer vision algorithms. He is also interested in using computers for artistic expression.

Alessandro Neri was born in Viterbo (Italy) in 1954. In 1977 he received the Doctoral degree in electronic engineering from the University of Rome "La Sapienza." In 1978, he joined the Research and Development Department of Contraves Italiana S.p.A. where he gained a specific expertise in the field of radar signal processing and in applied detection and estimation theory, becoming the chief of the advanced system group. In 1987, he joined the INFOCOM Department of the University of Rome "La Sapienza" as Associate Professor in signal and information theory. In November 1992, he joined the Electronic Department of the University of Rome "Roma Tre" as Associate Professor in electrical communications. Since September 2001, he is Full Professor in telecommunications at the University of Rome "Roma Tre." Since 1987, his research activity has been focused mainly on information theory, signal theory, and signal and image processing and their applications to both telecommunications systems and remote sensing.