

CONTOUR DETECTION BY MULTIREOLUTION SURROUND INHIBITION

Giuseppe Papari*, Patrizio Campisi**, Nicolai Petkov*, Alessandro Neri**

*Institute of Mathematics and Computing Science
University of Groningen
P. O. Box 800, 9700 AV Groningen, The Netherlands
G.Papari@cs.rug.nl, petkov@cs.rug.nl

**Dipartimento di Elettronica Applicata
Università degli Studi di Roma "Roma Tre",
Via della Vasca Navale 84, 00146 Roma, Italy
campisi@uniroma3.it, neri@uniroma3.it

ABSTRACT

In natural images, luminance changes occur both on object contours and on textures. Often, the latter are stronger than the former, thus standard edge detectors fail in isolating object contours from texture. To overcome this problem, we propose a multiresolution contour detector motivated by biological principles. At each scale, texture is suppressed by using a bipolar surround inhibition process. The binary contour map is obtained by a contour selection criterion that is more effective than the classical hysteresis thresholding. Robustness to noise is achieved by Bayesian gradient estimation.

Keywords: edge, context, contour, surround suppression, texture

1. INTRODUCTION

Edge and contour detection, an important task in computer vision, is a fertile field of ongoing research (see [1] for a survey). Standard edge detectors react to all non-negligible luminance changes in an image, irrespective whether they are originated by object contours or by texture (grass, foliage, waves, etc.). Moreover, luminance changes due to texture are often stronger than ones due to contours. Our goal is to isolate objects in a scene; therefore, some further process is needed with respect to general purpose edge detectors. Specifically, we use some principles deployed in the Human Visual System (HVS). Psychophysical studies show that the perception of an oriented stimulus can be influenced by other similar stimuli in the surroundings [2]. Neurophysiological researches show that surround modulation is due to a specific neural mechanism. In [3] it has been suggested that surround suppression effectively enhances contours in natural images rich in textures. Other psychological experiments show that the retinal image is decomposed through band-pass filters. Low-pass filters are responsible for the so called *pre-attentive* stage of vision, corresponding to the first 0.1 ÷ 0.3 s of the image persistency on the retina, where only the general morphology is perceived [4]. High-pass filters deliver information for the subsequent *attentive* stage where details are recognized. A multiresolution approach to contour detection has been proposed in [5]. In the current work we extend our previous studies [6] in combining a multiresolution approach and surround inhibition. We propose a method that detects contours at different resolutions and combines them by a contour-oriented selection algorithm. At each scale, noise is reduced by optimal Bayesian Minimum Mean Square Estimation (MMSE) of the gradient in additive noise and texture is suppressed by a biologically motivated surround inhibition process.

2. SCALE DEPENDANT CONTOUR DETECTOR

The proposed single scale contour detector is depicted in Fig. 1, where $I^w(x,y)$ is a noisy version of a given image $I(x,y)$ corrupted by additive independent noise $w(x,y)$. First, the gradient of the in-

put image I^w is evaluated by convolution with the gradient of a Gaussian kernel $g_\sigma(x,y)$ [8]. The gradient estimation depends on the parameter σ , which we will call *scale*, or *resolution*.

$$\nabla_\sigma I^w = \nabla \{ I^w * g_\sigma \} = \{ I^w * \nabla g_\sigma \}, \quad g_\sigma(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$



Fig. 1 Scale dependant contour detector.

Then, Bayesian denoising, described in Section 2.1, is applied on the noisy image gradient. Surround inhibition is performed as detailed in Section 2.2.

2.1 Bayesian denoising

Our goal is to find the optimal estimator $\nabla_\sigma \tilde{I} = \hat{\mathbf{a}}(\mathbf{z})$ of the unknown vector $\mathbf{a} = \nabla_\sigma I$, when a noisy version $\mathbf{z} = \nabla_\sigma I^w = \nabla_\sigma I + \nabla_\sigma w$ is observed. As well known from the Bayesian estimation theory, the optimal MMSE estimator is given by:

$$\hat{\mathbf{a}}(\mathbf{z}) = \frac{p_{\mathbf{z}|\mathbf{a}}(\mathbf{z}|\mathbf{a})p_{\mathbf{a}}(\mathbf{a})}{\int p_{\mathbf{z}|\mathbf{a}}(\mathbf{z}|\mathbf{a})p_{\mathbf{a}}(\mathbf{a})d\mathbf{a}} \quad (2)$$

According to recent statistical studies on natural images [7], both $p_{\mathbf{a}}(\mathbf{a})$ and $p_{\mathbf{z}|\mathbf{a}}(\mathbf{z}|\mathbf{a})$ are assumed Gaussian Scale Mixture (GSM), with covariance matrices \mathbf{A}_i and \mathbf{N}_k respectively:

$$\begin{cases} p_{\mathbf{a}}(\mathbf{a}) = \sum_{i=1}^{K_1} \lambda_i \mathcal{N}_2(\mathbf{a}, \mathbf{0}, \mathbf{A}_i) \\ p_{\mathbf{z}|\mathbf{a}}(\mathbf{z}|\mathbf{a}) = \sum_{k=1}^{K_2} \beta_k \mathcal{N}_2(\mathbf{z}, \mathbf{a}, \mathbf{N}_k) \end{cases}, \quad \sum_{i=1}^{K_1} \lambda_i = \sum_{k=1}^{K_2} \beta_k = 1 \quad (3)$$

where:

$$\mathcal{N}_2(\boldsymbol{\xi}, \boldsymbol{\mu}, \mathbf{R}) = \frac{1}{2\pi \det \mathbf{R}} \exp \left[-\frac{1}{2} (\boldsymbol{\xi} - \boldsymbol{\mu})^T \mathbf{R}^{-1} (\boldsymbol{\xi} - \boldsymbol{\mu}) \right] \quad (4)$$

By substituting eq. (3) in eq. (2), we can find the following closed expression for the optimal MMSE estimator:

$$\hat{\mathbf{a}}(\mathbf{z}) = \frac{\sum_{i,k} \lambda_i \beta_k \mathcal{N}_2(\mathbf{z}, \mathbf{0}, \mathbf{A}_i + \mathbf{N}_k) \mathbf{A}_i^{-1} (\mathbf{A}_i + \mathbf{N}_k)^{-1} \mathbf{z}}{\sum_{i,k} \mathcal{N}_2(\mathbf{z}, \mathbf{0}, \mathbf{A}_i + \mathbf{N}_k)}. \quad (5)$$

The nonlinearity defined by eq. (5), applied to each pixel of the gradient $\nabla_\sigma I^w$, gives the best estimation of $\nabla_\sigma I$.

2.2 Surround inhibition

Next, a surround inhibition operator taking into account the context

influence of the surrounding of each point is deployed. In [3] an inhibition term $T_\sigma(x, y)$ is introduced as the local average of the gradient magnitude $M_\sigma(x, y) = |\nabla_\sigma \tilde{I}(x, y)|$ on a ring r around each pixel. $T_\sigma(x, y)$, large in textured areas and small for isolated edges, is then subtracted from $M_\sigma(x, y)$, thus reducing the response to texture. With this type of inhibition, there is a certain auto-inhibition of isolated edges. Moreover, edges at texture borders are considerably inhibited as well, which is not desirable with respect to the detection of region boundaries. In order to overcome these problems, we propose a bipolar inhibition term and a two-level inhibition process.

A. Bipolar inhibition term

The ring r is split along the edge direction $\varrho_\sigma(x, y)$ in two halves r^+ and r^- , on which two local averages $T_\sigma^\pm(x, y)$ are evaluated (see Fig. 2) and the smaller value is taken as an inhibition term. Specifically, let us consider the following pairs of orientation dependent filters $w_{\sigma, \phi}^\pm(x, y)$, which define two half-rings oriented along an angle $\phi \in [0, \pi)$:

$$W_{\sigma, \phi}^\pm(x, y) = |g_\sigma(x, y) - g_{4\sigma}(x, y)|^+ \cdot U[\pm(x \cos \phi + y \sin \phi - a)]$$

$$w_{\sigma, \phi}^\pm(x, y) = \frac{W_{\sigma, \phi}^\pm(x, y)}{\iint_{R^2} W_{\sigma, \phi}^\pm(x, y) dx dy} \quad (6)$$

where

$$|\xi|^\pm = \begin{cases} \xi, & \xi \geq 0 \\ 0, & \xi < 0 \end{cases}, \quad U(\xi) = \begin{cases} 1, & \xi \geq 0 \\ 0, & \xi < 0 \end{cases} \quad (7)$$

The weighted local averages are defined by the following convolutions and, for each pixel, the minimum response is taken:

$$\begin{cases} T_\sigma^\pm(x, y) = \{M_\sigma * w_{\sigma, \phi}^\pm\}(x, y)|_{\phi=\varrho_\sigma(x, y)}, \\ T_\sigma(x, y) = \min\{T_\sigma^+(x, y), T_\sigma^-(x, y)\} \end{cases} \quad (8)$$

The convolutions are computed for a discrete set of orientations $\{\phi_i\}_{i=1}^{N_\phi}$, $\phi_i = \pi(i-1)/N_\phi$ and, for each pixel, the result obtained for the angle ϕ_i that is closest to the gradient orientation $\varrho_\sigma(x, y)$, is used (see Fig. 3).

On isolated edges (Fig. 2a), the local averages on both sides are very low, ideally zero. Consequently, $T_\sigma(x, y)$ is low and contours are not inhibited. In textured areas (Fig. 2b), the local averages on both sides are high and so will be the inhibition term. Borders of textured areas are not inhibited, since $T_\sigma(x, y)$ is low on such points (Fig. 2c). These modifications lead to a considerable improvement of the surround inhibition effect obtained in [3].

B. Two-level inhibition

The proposed inhibition scheme is shown in Fig. 4, where the inhibition term $T_\sigma(x, y)$ is computed as specified in Section 2.2.A. Two inhibited gradient fields $c_{\sigma, \alpha_1}(x, y)$ and $c_{\sigma, \alpha_2}(x, y)$ are evaluated, corresponding to strong and weak inhibition, respectively;

$$c_{\sigma, \alpha}(x, y) = |M_\sigma(x, y) - \alpha T_\sigma(x, y)|^+, \quad \alpha = \alpha_1, \alpha_2, \quad \alpha_1 > \alpha_2 \quad (9)$$

The strongly inhibited gradient $c_{\sigma, \alpha_1}(x, y)$ contains only little texture, but some weak contours are broken (Fig. 5a). On the contrary, the weakly inhibited gradient $c_{\sigma, \alpha_2}(x, y)$ contains all contours, but texture is still present (Fig. 5b).

In order to get the advantages of both inhibition levels, we combine c_{σ, α_1} and c_{σ, α_2} as follows:

First, we apply non-maxima suppression and consider the sets S_1 and S_2 of the non-zero pixels of c_{σ, α_1} and c_{σ, α_2} .

$$S_i = \left\{ (x, y) \mid c_{\sigma, \alpha_i}(x, y) \neq 0, \quad \frac{\partial c_{\sigma, \alpha_i}}{\partial u} = 0, \quad \frac{\partial^2 c_{\sigma, \alpha_i}}{\partial u^2} < 0 \right\}, \quad i=1, 2$$

where u is the direction of the gradient $\nabla_\sigma \tilde{I}$. Note that $S_1 \subseteq S_2$, since $\alpha_1 > \alpha_2$. Then, the output binary map b_σ is defined as the set of pixels of S_2 which are connected to at least one point of S_1 . In this way, the broken contours of S_1 are restored and most of the texture present in S_2 is removed (Fig. 5c).

The sets S_1 and S_2 have been obtained without thresholding the

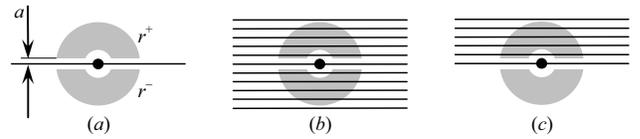


Fig. 2 Half-rings on which $M_\sigma(x, y)$ is averaged, for: isolated edges (a), textured areas (b), borders of textured areas (c).

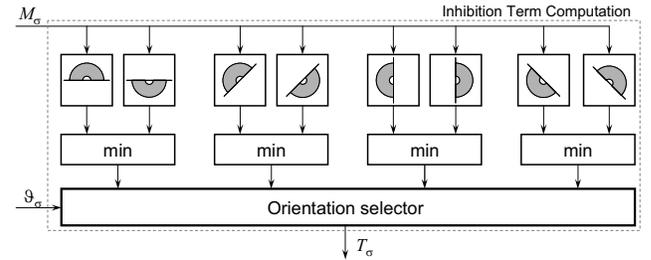


Fig. 3 Computation scheme of the inhibition term $T_\sigma(x, y)$.

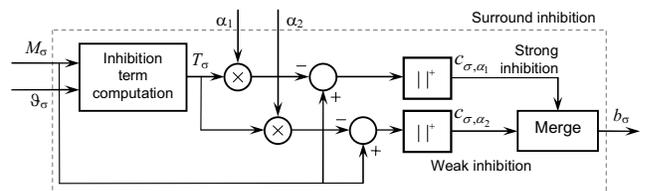


Fig. 4 Surround inhibition block (see Fig. 1).

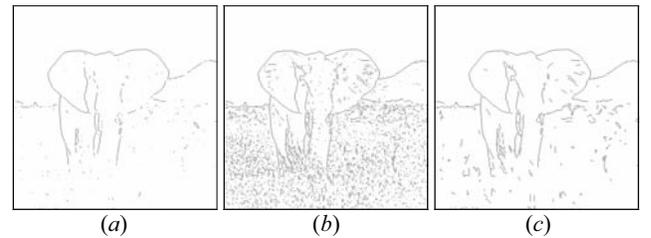


Fig. 5 Sets S_1 (a) and S_2 (b) of the nonzero pixels of respectively c_{σ, α_1} and c_{σ, α_2} , with strong and weak inhibition. (c) Output map b_σ in which contours are restored and texture is suppressed.

gradient magnitude in order to preserve the weak contours, thus some undesired edges are still present in Fig. 5c. In the common situation where some object contours are weaker than texture, the standard hysteresis thresholding techniques will fail. In the next section, the information at more resolutions are combined to select the desired contours and suppress even more texture.

3. MULTISCALE CONTOUR COMBINATION

As well known from multiresolution analysis, coarse scales contain the general morphology and almost no texture. On the other hand, contours detected at coarse scales are smoothed, shifted [9], and the non-maxima suppression destroys the junctions [10]. Thus, information at more scales can be combined in order to obtain contours being as detailed as at the fine scales, but without the texture that does not appear at coarse scales [5, 6, 11].

In [6], this has been achieved through the *pixel by pixel* logic AND combination of binary contour maps obtained at different scales. A morphological dilation at coarse scales is applied, in order to compensate the shifting and restore the junctions (see [6] for details). However, this approach has the drawbacks shown in Fig. 6a: part of the object contour $C_1^{(n-1)}$, detected at the scale $n-1$, falls outside its counterpart $\tilde{C}_1^{(n)}$ at coarser scale n , therefore it is destroyed by the logic AND. Conversely, the accidental superposition of some residual texture in $C_2^{(n-1)}$ and $\tilde{C}_2^{(n)}$ is maintained.

To overcome these problems, we propose the multiscale contour detector shown in Fig. 7. First, the binary contour maps b_1, \dots, b_N at the scales $\sigma_1 < \sigma_2 < \dots < \sigma_N$ are computed using the contour detector introduced in Section 2. Given two binary contour maps b_{n-1} and b_n , detected at two adjacent scales, a new map $b_{n-1}^{out} = CS(b_{n-1}, b_n)$ is obtained as detailed in Fig. 8, by means of the mathematical operator CS (Contour Selector). It selects from b_{n-1} (finer scale), the pieces of contours having a good overlap with the contours b_n (coarser scale) and forming long chains of non-zero pixels. The operator CS is applied iteratively from the coarsest scale up to the finest one.

In more detail, the block CS operates as follows: first, the morphological dilation $\tilde{b}_n = b_n \oplus D_2$ is computed, with a disk D_2 of radius 2 pixel as a structuring element. The connected components of b_{n-1} and \tilde{b}_n are denoted by $C_k^{(n-1)}$ and $\tilde{C}_i^{(n)}$, respectively:

$$b_{n-1} = \bigcup_k C_k^{(n-1)}, \quad \tilde{b}_n = \bigcup_i \tilde{C}_i^{(n)}. \quad (10)$$

We measure the degree of overlapping between $C_k^{(n-1)}$ and $\tilde{C}_i^{(n)}$ by the following quantity:

$$\chi_k^{(n-1)} = \frac{\text{card}\{C_k^{(n-1)} \cap \tilde{C}_i^{(n)}\}}{\text{card}\{C_k^{(n-1)}\}} \quad (11)$$

where $\text{card}\{X\}$ indicates the cardinality of the set X . All the components $C_k^{(n-1)}$ such that $\chi_k^{(n-1)}$ is below a threshold T_χ are removed. Thus the component $C_1^{(n-1)}$ in Fig. 6a will not be broken and the undesired component $C_2^{(n-1)}$ will be completely removed.

The second step consists in labeling the components $C_k^{(n-1)}$, which are included in the same component $\tilde{C}_i^{(n)}$, as part of the same ob-

ject $O_i^{(n)}$ (Fig. 6b). For each object $O_i^{(n)}$, we define its weight $R_i^{(n-1)}$ as the sum of the values of $M_\sigma(x, y)$ over $O_i^{(n-1)}$:

$$O_i^{(n-1)} = \bigcup_{\substack{\chi_k^{(n-1)} > T_\chi \\ C_k^{(n-1)} \cap \tilde{C}_i^{(n)} \neq \emptyset}} C_k^{(n-1)}; \quad R_i = \sum_{(x,y) \in O_i^{(n-1)}} M_\sigma(x, y) \quad (12)$$

The final contour map b_{n-1}^{out} is given by the union of all the objects $O_i^{(n-1)}$ whose weight is above a given threshold T_R :

$$b_{n-1}^{out} = CS(b_{n-1}, b_n) = \bigcup_{R_i^{(n-1)} > T_R} O_i^{(n-1)} \quad (13)$$

In this way, we exploit the fact that object contours form long chains of non-zero pixels (high weight), whereas textures only forms short rods (low weight). Such approach is more effective than thresholding the *local* gradient magnitude, because often luminance changes due to texture are stronger than the ones due to object contours.

4. EXPERIMENTAL RESULTS AND COMPARISON

The performance of the proposed contour detector has been compared with four other existing algorithms. The results are presented in Figs. 9 and 10 for a test image, without and with additive noise (SNR = 13 dB). Other results are available on the webpage www.cs.rug.nl/~papari/resultsICIP06.htm. As it can be seen, our approach gives the best results in terms of texture suppressed, cleanness of the detected contour, and robustness to noise.

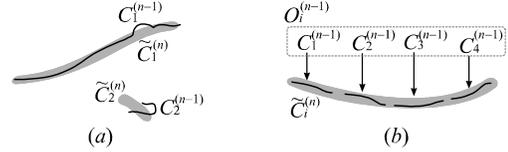


Fig 6. (a) A pixel based superposition would destroy part of the object contour $C_1^{(n-1)}$, and would keep part of the spurious texture $C_2^{(n-1)}$. (b) The object $O_i^{(n-1)}$, given by the union of the four connected components $C_1^{(n-1)} - C_4^{(n-1)}$ of b_1 , belongs to the connected component $\tilde{C}_i^{(n)}$ of \tilde{b}_n .

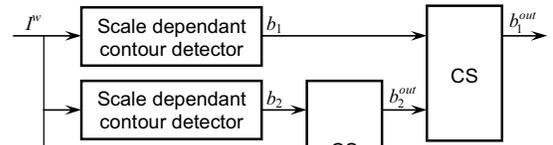


Fig. 7 Proposed multiscale contour detector.

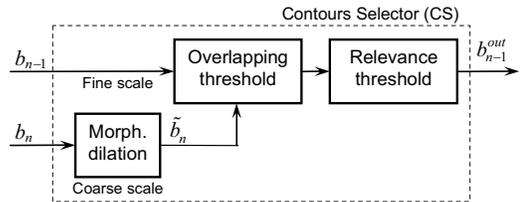


Fig. 8. Block «CS» of the scheme in Fig. 7.

With respect to the Canny edge detector (Figs. 9c, 10c), the benefits of the multiscale analysis [5], without surround inhibition, are shown in Figs. 9d, 10d: some texture is removed and noise is reduced. Comparable texture suppression is achieved with the single scale surround inhibition algorithm proposed in [3], (see Figs. 9e, 10e). The combination of multiscale analysis and surround inhibition [6] gives the much better results shown in Figs. 9f, 10f. The improvement proposed here leads to the even better result shown in Figs. 9b, 10b.

A ground-truth based performance evaluation has also been performed. The similarity ρ between each Algorithmic Result (AR) and the Ground Truth (GT) has been computed as follows:

$$\rho = \frac{EC}{EC + MC + FP} \quad (14)$$

where EC (Exact Contours) indicates the number pixels present both in the AR and in the GT. MC (Missing Contours) indicate the number of pixel present in the GT but not in the AR. FP (false Positive) indicates the number of pixel present in the AR but not in the GT, and measure the amount of not suppressed texture. ρ is always between 0 and 1, with $\rho = 1$ iff AR = GT.

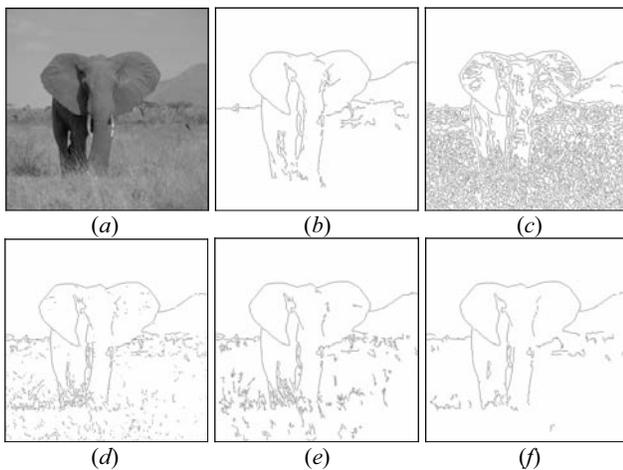


Fig. 9 Input image (a) and contours detected with: (b) the proposed approach, (c) the Canny edge detector, (d) the multiscale edge detector CARTOON without surround inhibition [5], (e) single [3] and (f) multi scale surround inhibition [6].

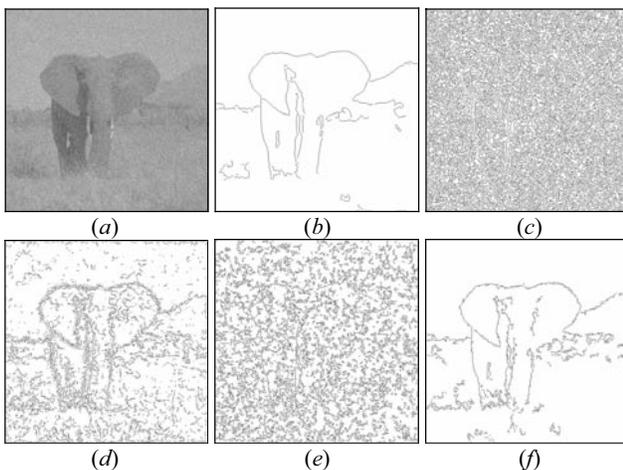


Fig. 10. Contours detected on a noisy image test (SNR = 13dB).

For each studied algorithm, the values of ρ have been computed on a set of 24 images. The average value $\bar{\rho}$ and the standard deviation σ_ρ are shown in Fig. 11, both for noiseless and noisy images.

5. SUMMARY AND CONCLUSION

The proposed multiscale contour detector exploits important aspects of the HVS, in order to isolate object contours from texture. Edges surrounded by other edges are inhibited, since the HVS perceives them as texture rather than object contours. The bipolar mechanism introduced here avoids the auto-inhibition of the weak contours. The two-level inhibition process operates a strong inhibition on textured areas and a weak one on object contours. Similarly to the HVS, contours are detected at more resolutions. All the contours' parts having a low overlapping degree χ_k with respect to the adjacent coarser scale, and a low weight R_k are removed. Thresholding *global* quantities as χ_k and R_k is more effective than thresholding the *local* gradient magnitude $M_g(x,y)$. Robustness to noise, for the general non-Gaussian case, is achieved by using a Bayesian estimator. GSM models are employed for both the image and the noise and a closed form of the estimator has been provided. As shown by experimental results and performance evaluations, our algorithm outperforms both standard and more sophisticated approaches, based on single and multiscale surround inhibition.

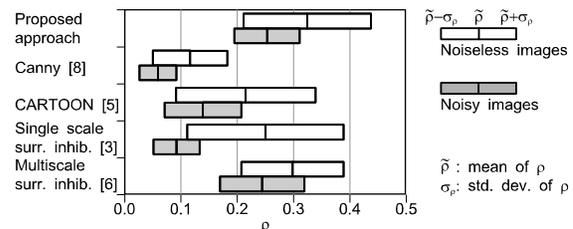


Fig. 11. Quantitative performance comparison.

6. REFERENCES

- [1] M. Basu, "Gaussian-based edge-detection methods: A Survey", *IEEE SMC-C* **3** (32) (2002) 252-260.
- [2] D.J. Field, A. Hayes, R.F. Hess, "Contour integration by the human visual system: Evidence for a local association field", *Vision Research* **33** (2) (1993) 173-193.
- [3] C. Grigorescu, N. Petkov and M.A. Westenberg: "Contour detection based on non-classical receptive field inhibition", *IEEE Trans. on Image Processing*, **12** (7) (2003) 729-739.
- [4] B. Julesz, "Visual Pattern Discrimination", *IRE Transactions on Information Theory*, **8** (1962) 84-92.
- [5] W. Richards, H.K. Nishihara, B. Dawson, "CARTOON: A biologically motivated edge detection algorithm", *MIT A.I. Memo No. 668* (1982).
- [6] G. Papari, P. Campisi, N. Petkov, A. Neri, "A multiscale approach to contour detection by texture suppression", *SPIE Image Proc.: Alg. and Syst.* (2006), Vol. 6064A, San Jose, CA.
- [7] J. Portilla and E.P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients" *Int. J. Comput. Vis.*, **40** (1) (2000) 49-71.
- [8] J.F. Canny, "A computational approach to edge detection", *IEEE PAMI* **8** (6) (1986) 679-698.
- [9] K.H. Liang, T. Tjahjadi and Y.H. Yang. "Bounded diffusion for multiscale edge detection using regularized cubic B-spline fitting" *IEEE SMC-B* **29** (2) (1999) 291-297.
- [10] A. Ding, A. Goshtasby, "On the Canny edge detector", *Pattern Recognition*, **34** (3) (2001) 721-725.
- [11] F. Bergholm, "Edge focusing", *IEEE PAMI* **9** (6) (1987) 726-741