Contour Detection Based On the Property of Orientation Selective Inhibition of Nonclassical Receptive Field

Li Long, Yongjie Li
School of Life Science and Technology
University of Electronic Science and Technology of China
Chengdu, 610054, China
liyj@uestc.edu.cn

***Abstract*—The majority of neurons in primary visual cortex (V1) of brain are orientation-selective. Both the classical receptive field (CRF) and the non-classical receptive field (NCRF), which modulates the CRF and mainly yields inhibition, could be orientation-selective and may obtain different tune. For a single neuron, the response is determined by the interaction of its CRF and NCRF. And the horizontal connections play an important role when forming inhibition. Inspired by those visual cortical mechanisms, a modified inhibition model, called orientation selectivity of NCRF, is introduced to improve the performance of contour detectors. The orientation saliency is determined by the energy, the output of a Gabor Energy filter, in each direction. And the inhibition term is taken based on the saliency of the orientation. This proposed method could selectively retain the object contours and suppress texture edges effectively, which is demonstrated by the processing of several natural images.**

***Keywords*—contour detection, orientation selectivity, non-classical receptive field, inhibition, texture, Gabor energy, horizontal connecting.**

I. Introduction

Contour detection is one of the most important task in object recognition in computer vision and is a fundamental operation in image processing. For the past two decades, a large number of edge detection algorithms have been proposed, some of which are quite remarkable [1]. And research on edge detection is still a fertile activity. But the traditional operators can not distinguish the edges generated from texture and objects. Since a contour, unlike edges which can be generated by texture, represents a line delimiting the object in a scene. It makes the detection more difficult. However, Human visual system was the mechanisms to extract this feature rapidly and effectively.

Evidence have been found by psychophysicists and neurophysiologists [2] that: Visual neurons receive signal inputs from their receptive fields (RF or Classical RF), The area beyond receptive fields (integration field, IF or NCRF), although being unresponsive to visual stimulation, exerts modulatory effects on the cell activities elicited by RF stimulation. Most of the cortical IFs are inhibitory, showing selectivity to stimulus parameters. The interaction between RF and IF provides the neuronal basis for the detection of various texture contrast, the perception of relative speed between moving object and background and the generation of visual illusions. Li's work [4] shows that the extent of the IFs is most frequently 2-5 times of the size of RFs. The tuning properties of IFs, for most cells, are similar to, but broader than, the tuning of RF. When stimulus features are similar within the RF and IFs, the inhibition goes to the maxima. But when stimulus pattern differs in RF and IF, the inhibition decreases or even disappears. This phenomena suggests the IFs associated with RFs, not just affect the excitatory of RF, but to form a feature detector. The horizontal connections of the neurons play an import role in forming the inhibition of NCRF, and the effect of horizontal input is state-dependent, with the size and sign of the laterally evoked response changing according to the balance of converging inputs [6], [7].

A recent model proposed by Grigorescu et al. [9] utilized the non-classical receptive field inhibition to reduce texture edges and to extract isolated contours in real images. However, the model did not explain the source of NCRF inhibition. In this paper, we took the horizontal connections of V1 neurons as the derivation of inhibition, and introduced a modified inhibition model, in which isotropic and anisotropic inhibition are both taken into account in this model.

The rest of the paper is organized as follows: Section II describes the computational model of V1 neurons. Several parameters that determine the inhibition strength of NCRF are also discussed. In section III we verified the model with natural images and evaluated the performance of our model by comparing to other detectors. Finally, section IV summarizes the main conclusions of this work and discusses further improvements.

II. Contour Detector

A. Excitatory Response

Gabor functions are used to simulate the V1 neurons in a lot of works [8], [10], [11]. Gabor energy, the response modulus of orthogonal pairs of Gabor filters, can capture typically fundamental characteristics of complex cells [12]. Thus, in this paper, Gabor energy filter is chosen to simulate the response of complex cells. Two-dimensional Gobor filter is given as [9]:

978–1–4244–1676–9/08/$25.00 © 2008 IEEE

$$g(x,y;\theta) = e^{-\frac{\tilde{x}^2+\gamma^2\tilde{y}^2}{2\sigma^2}}\cos(2\pi\frac{\tilde{x}}{\lambda}+\varphi)\,. \tag{1}$$

where $\tilde{x} = x\cos\theta + y\sin\theta, \tilde{y} = -x\sin\theta + y\cos\theta$. φ determines the ellipticity of the CRF. CRF's size is determined by σ, the standard deviation of Gaussian. Parameter λ determines the wavelength, and σ/λ determines the spatial frequency bandwidth. θ is an angle parameter to determine the prefer orientation of a single V1 neuron, $\theta \in [0,\pi)$. φ is a phase offset, determining the symmetric of $g(x,y;\theta)$, $\varphi = 0$ and $\varphi = \pi$ refer to symmetric while $\varphi = -\frac{\pi}{2}$ and $\varphi = \frac{\pi}{2}$ refer to asymmetric.

The Gabor filter response $e(x,y;\theta)$, also mentioned as simple cell response [10], is computed by convoluting $g(x,y;\theta)$ with $I(x,y)$ at the point (x,y), where $I(x,y)$ denotes the input image.

$$e(x,y;\theta) = I(x,y) * g(x,y;\theta)\,. \tag{2}$$

Then the complex cell's response,Gabor energy, is given as:

$$E(x,y;\theta) = \sqrt{e_0^2(x,y;\theta) + e_{-\frac{\pi}{2}}^2(x,y;\theta)}\,. \tag{3}$$

where $e_0(x,y;\theta)$ and $e_{-\frac{\pi}{2}}(x,y;\theta)$ are the output of symmetric and asymmetric Gabor filters, respectively.

Thus a group of orientation selective complex cells can simulate by a group of Gabor energy filter $E(x,y;\theta)$ with different orientation θ_i.

$$\theta_i = \frac{i\pi}{N_\theta}, \quad i = 0,1,\ldots,N_\theta - 1. \tag{4}$$

where N_θ is the orientation number.

B. Inhibition Terms

NCRF and CRF together form a feature detector [2] and the horizontal connection of neurons compose a great part when forming inhibition [7], [16]. Horizontal connection intensity of neurons is determined by the distance and orientation contrast between NCRFs and CRFs. We hypothesize that: (a) Inhibition pattern can be divided into two types, orientation selective inhibition and disorientation selective inhibition, by the Gabor energy distribution on orientations, referred as Inhibition Orientation Saliency; (b) Both orientation and disorientation inhibitions interact within neuron response. Thus we compose the inhibition terms as follow.

1) Distance Influence: Intrinsic connections exist between neurons and others around them and the intensities are distance related. Difference of Gaussian function (DOG), used to simulate the receptive field [14], can describe neuron connection strengths in distance. As the connection strengths can not be negative, The non-negative DOG function $DOG^+(x,y)$ is given as:

$$DOG^+_{\sigma_1,\sigma_2}(x,y) = \left| \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{x^2+y^2}{2\sigma_2^2}} - \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{x^2+y^2}{2\sigma_1^2}} \right|^+ . \tag{5}$$

where σ_2 and σ_1 are the standard divisions of two Gaussian functions, denoting the size of NCRF and CRF, respectively. Since NCRF's size is 2-5 times bigger than CRF, we set $\sigma_2 = 4\sigma_1$ and $\sigma_1 = \sigma$. $|x|^+$ denotes an operation that: $|x|^+ = x$ while $x > 0$ and $|x|^+ = 0$ while $x \leqslant 0$. The weight function based on distance can be written as:

$$w_d(x,y) = \frac{DOG^+_{\sigma_1,\sigma_2}(x,y)}{\|DOG^+_{\sigma_1,\sigma_2}(x,y)\|}. \tag{6}$$

where $\|\cdot\|$ denotes the $L1$ norm. The distance based inhibition term $i_d(x,y;\theta)$ is computed as a convolution with $E(x,y;\theta)$:

$$i_d(x,y;\theta) = E(x,y;\theta) * w_d(x,y). \tag{7}$$

2) Orientation Contrast: Most of NCRFs have orientation selectivity and their prefer orientations may not the same as CRFs [13]. Thus the orientation contrast between CRF and NCRF would impact the inhibition strength [15]. We compute the orientation contrast $i_{oc}(x,y;\theta)$ as:

$$i_{oc}(x,y;\theta) = \cos[\Delta\theta]^+. \tag{8}$$

where $\Delta\theta$ is the difference of the prefer orientations of CRF, $\theta_c(x,y)$, and NCRF ,$\theta_n(x,y)$. $[\Delta\theta]^+$ denotes an operation:

$$[\Delta\theta]^+ = [\theta_c - \theta_n]^+ = \begin{cases} \Delta\theta, & |\Delta\theta| < \pi/2; \\ \pi - \Delta\theta, & |\Delta\theta| \geqslant \pi/2. \end{cases} \tag{9}$$

while the prefer orientations θ_c and θ_n are determined by the maxima Gabor energy of all orientations as follow:

$$\begin{aligned} \theta_c(x,y) &= arg\max\{E_{\lambda,\sigma,\theta}(x,y)\,|\,\theta = 0,\pi/N_\theta,\ldots,(N_\theta-1)\pi/N_\theta\}; \\ \theta_n(x,y) &= arg\max\{i_{d,\theta}(x,y)\,|\,\theta = 0,\pi/N_\theta,\ldots,(N_\theta-1)\pi/N_\theta\}. \end{aligned} \tag{10}$$

3) inhibition Orientation Saliency: We define the inhibition orientation saliency as a ratio of the maxima intrinsic inhibition in all directions divided by the mean energy. Two threshold are set to determine whether the inhibition should be isotropic or anisotropic. The function is given as:

$$i_s^*(x,y) = \frac{\max\{i_d(x,y;\theta)\}}{(\sum i_d(x,y;\theta))/N_\theta}. \tag{11}$$

And the saliency coefficient $i_s(x,y)$ is determined as:

$$i_s(x,y) = \begin{cases} 1 & i_s^*(x,y) \geqslant t_h; \\ \frac{i_s^*(x,y)-t_l}{t_h-t_l} & t_l \geqslant i_s^*(x,y) < t_h; \\ 0 & i_s^*(x,y) < t_l. \end{cases} \tag{12}$$

Where t_h is the upper threshold that determine whether the inhibition pattern is orientation selective or not. And t_l is the lower threshold that determine whether the inhibition pattern is disorientation selective or not. As can be seen the $i_s(x,y)$ is fixed into $i_s(x,y) \in [0,1]$. Associated with the orientation saliency coefficient, the orientation saliency $i_{os}(x,y)$ based inhibition intensity is determined by:

$$i_{os}(x,y) = \begin{cases} \max\{i_d(x,y;\theta)\} & i_s(x,y) > 0\,; \\ (\sum i_d(x,y;\theta))/N_\theta & i_s(x,y) = 0\,. \end{cases} \tag{13}$$

Equation (13) means that if inhibition pattern is orientation selective, then the inhibition input is the energy on the preferred orientation. When the pattern is disorientation selective, the inhibition input is an average of energies on all orientation.

4) Interaction of CRF and NCRF: The response of a complex cell $r^*(x, y)$, after the orientation selective interaction with NCRF, is defined as below:

$$r^*(x, y; \theta) = E(x, y; \theta) - \alpha_1 \cdot i_{os}(x, y) \cdot i_s(x, y) \cdot i_{oc}(x, y; \theta). \quad (14)$$

Since the disorientation selective inhibitions are ubiquitous in V1 cortex [17], We add an disorientation selective inhibition term after the orientation selective inhibition. And the output becomes:

$$r(x, y; \theta) = |r^*(x, y; \theta) - \alpha_2 \cdot (1 - i_s(x, y)) \cdot i_{os}(x, y)|^+. \quad (15)$$

Where $|\cdot|^+$ is defined in (5), α_1 and α_2 denote the intrinsic connection strengths. Operator $|\cdot|^+$ is used here, because the neuron response can not be negative. We give the final output of the detector as:

$$r(x, y) = \max\{r(x, y; \theta) | \theta \in [0, \pi)\}. \quad (16)$$

This operator will respond to lines and bars that have different orientations with the texture. The suppression intensity is co-determined by the orientation saliency of NCRF, and the orientation contrast between CRF and NCRF. As the textures are thought to retain the similar orientations and object's contour presents, mostly, in different orientations with texture edge, the texture edge would be strongly suppressed.

We don't use the curve saliency operator to wipe off non-contour elements [18]. Thus, the results will contain some short edges and small pieces. Non-max suppressing and hysteresis thresholding [1] are used to generate binary maps.

III. Experiment Results

To verify our contour detection model, several natural images are employed (some images with ground truth contours are download from website: *http://www.cs.rug.nl/~image*).In order to get some comparison we employed Canny [1] detector and bar detector [9] in the experiments.

A. Contour Extraction

We set the parameters of our detector as: $N_\theta = 12$, $\sigma = [1.8, 2.0]$, $\varphi = 0.5$, $\sigma/\lambda = 0.56$, $\alpha_1 = [1.2, 1.6, 2.0, 2.4]$, $\alpha_2 = [0.6, 1.0, 1.4]\alpha_1$, $t_h = [2.5, 3.0, 3.5, 4.0]$, $t_l = [1.5, 2.0]$. The percentage of candidate edge pixels p is set to $p = [0.1, 0.2]$. For the Canny edge detector and bar detector we use the same scales set, and the inhibition strength of bar detector is set to $\alpha = [1.0, 1.2]$ as same as the value in paper [9].

Figure 1(a) shows a picture of a chair, backgrounded with grass, and the results of canny edge detector, bar detector and our model are shown in Fig. 1(b), Fig. 1(c) and Fig. 1(d), respectively. We can see that the Canny edge detector does not distinguish between contour edges and texture edges, thus the output of Canny detector obtained a lot texture noise and with a large parameter, $\sigma = 2.2$, the contours are more smooth, which makes the result hard to distinguish. Bar detector and our detector suppress textures, but in our model more textures are suppressed while keeping similar object contours. Fig. 2 is a test of a gazelle image with the ground truth given in Fig. 2(b). The result of canny detector, as discussed above, is

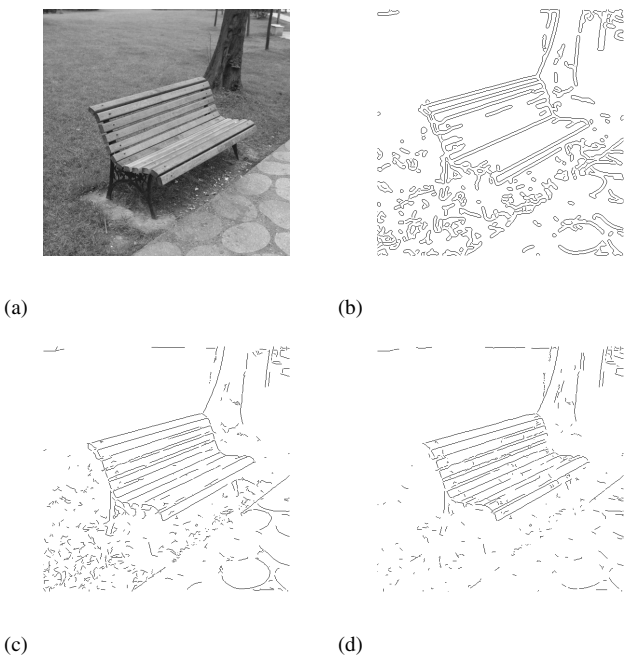

Fig. 1. A test of Chair where: (a) is the original image, (b) is the result of canny edge detector; (c) is the output of bar detector; The output of our model is presented in (d). All the detectors use the same parameter $\sigma = 2.2$. And for the bar detector, the parameter α is set to $\alpha = 1.0$. For our detector, the rest parameters are set as: $\alpha_1 = 1.2, \alpha_2 = 2.2\alpha_1, p = 0.1$ and the orientation saliency parameter are set to $t_h = 3.0, t_l = 1.5$.

hard to recognise the object while bar selector and our model's outputs easily exhibit the object of gazelle. Meanwhile, the grass edges, at the bottom of the picture, are more suppressed by our detector, which can be seen in Fig. 2(e). Some more examples with similar results are shown in Fig. 3 and Fig. 4.

B. Performance Evaluation

We adopt the method proposed by Grigorescu et al. to evaluate the performance of this model [9]. Let E_{GT} and B_{GT} be the set of edge pixels and background pixels of the ground truth edge image, respectively, and E_D and B_D be the set of edge pixels and background pixels of the operator-detected edge image, respectively. The set of correctly detected edge pixels is $E = E_D \cap E_{GT}$, false negatives are given by the set $E_{FN} = E_{GT} \cap B_D$, and the false positives are given by the set $E_{FP} = E_D \cap B_{GT}$. The performance of edge detectors are defined as:

$$P = \frac{card(E)}{card(E) + card(E_{FP}) + card(E_{FN})}. \quad (17)$$

where $card(x)$ denotes the number of elements of set x. p is the percentages of correctly detected edge pixels, $e_{fn} = card(E_{FN})/card(GT)$ and $e_{fp} = card(E_{FP})/card(E)$ denote the percentages of false negatives and false positives, respectively. Human-marked images are used as ground truth. Comparison of Canny, Bar detector and out model are shown in Table I.

From those data in Table I we can see that our detector can obtain better result by suppressing texture edges more

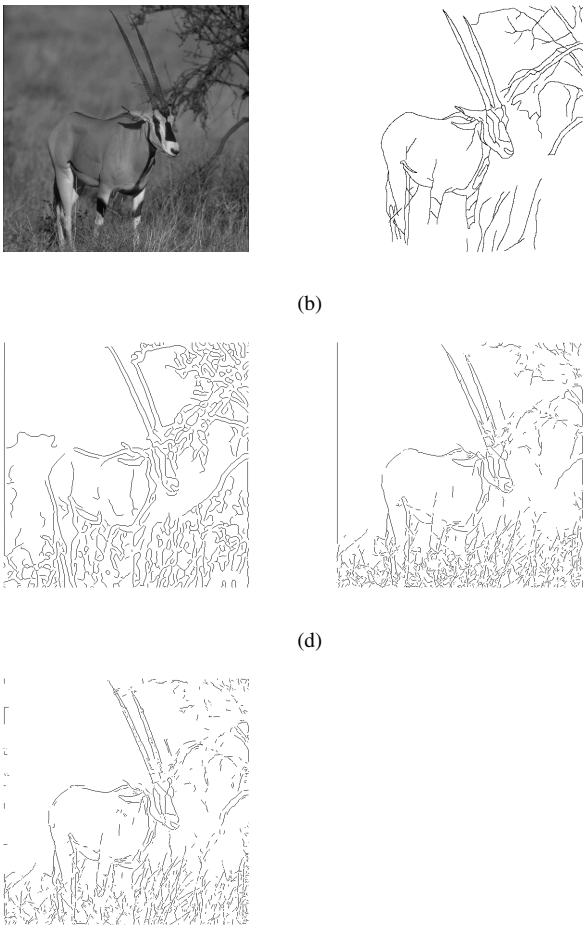

Fig. 2. A test of Gazelle where: (a) is the original image, (b) a hand-draw ground truth of object's contours; (c) is the result of canny edge detector; (d) is the output of bar detector; The output of out model is presented in (e). All the detectors use the same parameter $\sigma = 2.2$. And for the bar detector, the parameter α is set to $\alpha = 1.2$. For our detector, the rest parameters are set as: $\alpha_1 = 2.2, \alpha_2 = 1\alpha_1, p = 0.2$ and the orientation saliency parameter are set to $t_h = 3.0, t_l = 1.5$.

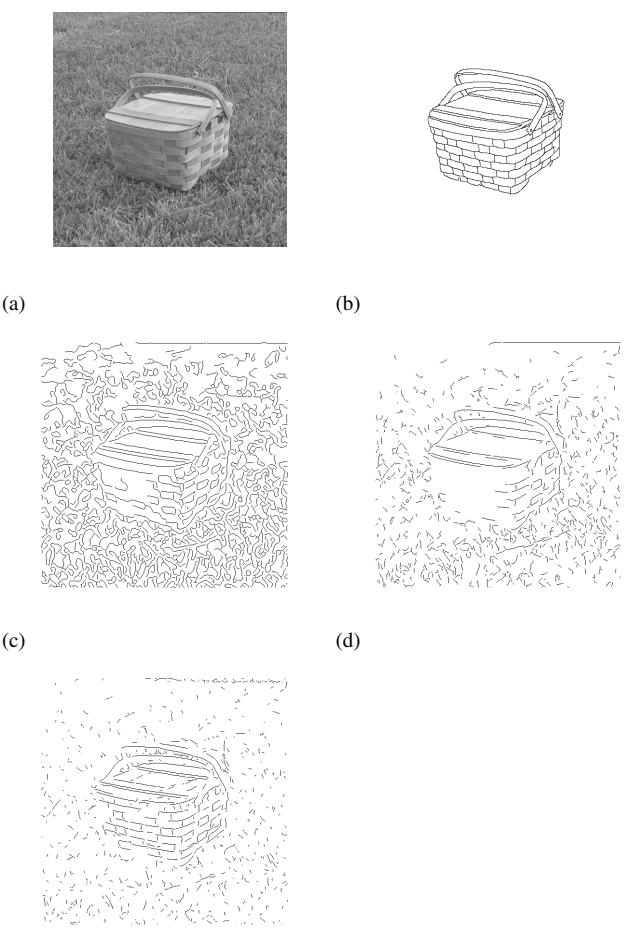

Fig. 3. A test of Basket where: (a) is the original image, (b) a hand-draw ground truth of object's contours; (c) is the result of canny edge detector; (d) is the output of bar detector; The output of our model is presented in (e). All the outputs are the best ones by traversing the parameters set.

effectively. Our detector contributes a smaller e_{fp} or e_{fn} and a higher P, comparing to canny edge detector and bar detector.

IV. Conclusion

Human visual system can process vision information rapidly and correctly. Contour extraction is one of the most important operation in low-level vision procession. A better contour detection will supply more useful and correct information to the postprocess. Thus, it plays a key role in information compressing and data coding in computer vision system. The model developed in this paper will response to lines and bars that in different orientations. Edges with the same or similar orientation will be treated as texture and be suppressed. The result of the non-meaningful information of an image is reduced, and it's a preparation for further process, such as coding, pattern recognition et al..

In this paper, we proposed a new inhibition model, which is mostly based on orientation selectivity, to improve the performance of contour detector. Compared to the bar detector, we put more attention to the orientation information and built some dynamic connection strength mechanism based on the horizontal connections that co-determine the inhibitions. And we combined the anisotropic and isotropic inhibition. The results show that these modifications make a saliency improvement.

However, this model still leaves future work to be done: (1) The vertical connection of V1 neurons play an important part in forming the NCRF, and the modulations are somehow facilitation. This mechanism suggests that when processing vision information, human visual system are trying to interfuse the similar features. When doing the contour or edge detection, this mechanism can be adopted to link edges or distinguish edges from texture and objects; (2) Our model in this paper was several parameters. To gain a better output, the optional parameters need to be found. But human visual system can do the work effectively. Thus, some self-adaptation mechanisms

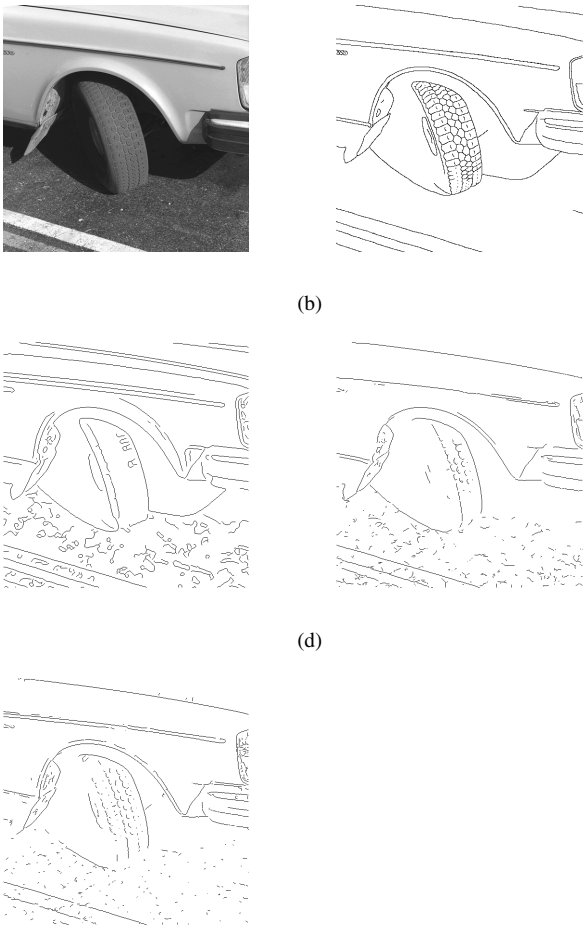

Fig. 4. A test of Tire where: (a) is the original image, (b) a hand-draw ground truth of object's contours; (c) is the result of canny edge detector; (d) is the output of bar detector; The output of our model is presented in (e). All the outputs are the best ones by traversing the parameters set.

should be introduced to determine the parameters automatically.

Acknowledgment

This work is supported by the grants from NSFC of China (Grant No.30500140 and 30730036) and a grant from the 973 Project of China (Grant No. 2007CB311001).

References

[1] J. F. Canny, "*A computational approach to edge detection*," IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679-1494, 1986.

[2] W. Li and C. Y. Li , "*Intergration field beyond the classical visual receptive field*," Chinese Journal of Neuroscience, vol. 1, no.1, pp. 1-6, 1994.

[3] M. K. Kapadia, G. Westheimer and C. D. Gilbert, "*Spatial distributioin of contextual interactions in primary visual cortex and in visual perception*," Journal of Neurophysiology, vol. 84, no. 4, pp. 2408-2062, 2000.

TABLE I
Performance Evaluation

	σ	e_{fp}	e_{fn}	P
Gazelle				
canny	2.2	0.62	0.22	0.41
bar detector	2.2	0.63	0.23	0.44
our model	2.2	0.49	0.23	0.50
Basket				
canny	2.2	0.82	0.24	0.21
bar detector	2.2	0.76	0.43	0.26
our model	2.2	0.49	0.24	0.49
Tire				
canny	2.2	0.53	0.33	0.41
bar detector	2.0	0.41	0.43	0.45
our model	1.8	0.18	0.46	0.64
Gnu				
canny	2.2	0.68	0.13	0.34
bar detector	2.4	0.41	0.23	0.58
our model	2.0	0.26	0.42	0.62

[4] C. Y. Li and W. Li, "*Extensive integration field beyond the classical receptive field of cat's striate cortical neurons and tuning properties*," Vision Res., vol. 34, no.18, pp. 2337-2355, 1994.

[5] D. H. Hubel and T. N. Wiesel, "*Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex*," Physiology, vol. 160, pp. 106-154, 1962.

[6] J. A. Hirsch and C. D. Gillbert, "*Synaptic physiology of horizontal connections in the cat's visual cortex*," J. Neurosic., vol. 11, pp. 1800-1809, 1991.

[7] J. Bolz and C. D. Gillbert, "*Generation of end-inhibition in the visual cortex via interlaminar connections*," Natures, vol. 320, pp. 262, 1986.

[8] R. Mehrotra, K. R. Namuduri and N. Ranganathan, "*Gabor filter-based edge detection*," Pattern Recogit., vol. 25, no. 12, pp. 1479-1494, 1992.

[9] C. Grigorescu. N. petkov and M. A. Westenberg, "*Contour detection based on nonclassical receptive field inhibition*," IEEE trans. Image Process. vol. 12, no. 17, pp. 729-739, 2003.

[10] J. G. Daugman, "*Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two demensional visual cortical filters*," Opt. Soc .Am. A, vol. 2, no. 7, pp.1600-1169, 1985.

[11] M. C. Morrone and D. C. Burr, "*Feature detection in human vision: a phase-dependent energy model*," Proc. R. Soc. London (Biol.), vol. 235, pp. 221-245, 1988.

[12] W. Chan and G. Coghill, "*Text analysis using local energy*," Pattern Recognition, vol. 34, pp. 2523-2532, 2001.

[13] Z. C. Qiu, Z. Li, F. J. GU and T. D. Shou, "*A new computational model of retainal ganglion cell receptive fields II. modeling center-surround interactioins in orientation selectivity of a ganglion cell receptive field with extended disinhibitory area*," Acta Biophysica Sinial, vol. 16, no. 2, pp. 296-302, Jun., 2000.

[14] R. W. Rodieck, "*Quantitative analysis of cat retinal ganglion cell response to visual stimuli*," Vision Res., vol. 5, no. 11, pp. 583-601, Dec., 1965.

[15] C. Blakemore and E. A. Tobin, "*Lateral inhibition between orientation detectors in the cat's visual cortex*," Exp. Brain Res., vol. 15, pp. 439-440, 1972.

[16] J. H. Chisum and D. Fitzpatrick, "*The contribution of vertical and horizontal connection to the receptive field center and surround in V1*," Neuron Networks, vol. 17, no. 5-6, pp. 681-691, Jun./Jul., 2004.

[17] D. L. Ringach, R. M. Shapley and M. J. Hawken, "*Orientation selectivity in macaque V1: diversity and laminar dependence*," The Journal of Neuronscience, vol. 22, no. 13, pp. 5639-5651, Jul., 2002.

[18] Q. L. Tang, N. Sang and T. X. Zhang, "*Extraction of salient contours from cluttered scenes*," Pattern Recognition, vol. 40, no. 11, pp. 3100-3109, Nov., 2007.