A contour detection method based on some knowledge of the visual system mechanisms

Kamel Belkacem-Boussaid
3367 Beckman Institute. University of Illinois at Urbana-Champaign. 405 N. Mathews Ave.
Urbana, IL 61801. USA.
belkacem@staff.uiuc.edu

Azeddine Beghdadi
Institut Galilee. Universite Paris Nord.
Avenue J.B. Clement. 93 430 Villetaneuse
France. bab@lpmtm.univ-paris13.fr

Abstract

We propose a contour detection method based on the mechanisms from biological visual perception. The temporal analysis of image is the basis of the model. The temporal notion means that the static image is transformed into a data flow. Each element of the flow is treated independently from the others. Our aim is image segmentation through contour detection. The model is composed of a succession of five stages: noise reduction, asynchronous processing, isotropic filtering and adaptive smoothing, dynamic thresholding, and temporal integration. To evaluate the proposed approach, objective and subjective analyses are performed on synthetic and actual images.

INTRODUCTION

The usefulness of contour detection in a great number of applications has been well established and demonstrated. Indeed, this operation is of great help for further image analysis and scene understanding. There are many kinds of edge detectors [18-19]. Most of them are based on derivative operators that give a high response at the contour points and a low response in homogeneous areas. The oldest and simplest edge detector is undoubtedly the digital gradient operator. However, its usefulness in image processing is limited, since it is not an isotropic operator. Given the bidimensionality of the image signal, one has to ensure rotation-invariance in the design of a derivative operator. The gradient magnitude fulfills this requirement but the gradient vector does not. This limitation has led to the development of many directional gradient filters [13]. In contrast, the laplacian filter is an orientation-invariant derivative operator. However, because of the second-order derivative, this operator is more sensitive to noise than the gradient operator. One can reduce the effects of noise by smoothing the image before applying the laplacian. Marr and Hildreth adopted this strategy for designing the well-known LOG (Laplacian of gaussian) filter as a contour detection operator [14]. Other edge detectors, such as the Sobel one, are based on the same principle. However, now, no universal contour detection method has emerged. Furthermore, as compared to the great number of investigations in this field, a little effort has been spent on the design of a quantitative evaluation of the contour detection, except the pioneer work by Abdou and Pratt [22], the one by Canny [15], and other few similar works [23]. In Canny's approach, three performance criteria for good detection are defined. However, there is an inevitable trade-off between contour localization accuracy and detectability, as noticed by Perona and Malik [5]. These authors propose a novel method inspired from Koenderink's work [16], using an anisotropic diffusion process, and based on three criteria:causality, immediate localization and piecewise smoothing. The idea of using anisotropic diffusion has led many other researchers to develop similar approaches for image filtering or image enhancement [6]. However, most of these sophisticated methods, which have been proved efficient for given images, are criterion oriented. Some requirements, irrespective of the image content, must be satisfied for the design of the oriented image treatment. Moreover, some parameters are chosen empirically. For instance, in the method of Perona and Malik, the edge sharpening depends on the diffusion function, which is selected in a heuristic manner. Furthermore, no convergence criterion has been specified in their iterative method, as noted by Gerig et al [8]. Another important point, which has not been widely discussed in almost all the edge detection methods, is the influence of the visual perception criteria on the design of the contour detection method. This observation has motivated us to propose another approach, which takes into account some knowledge of human perceptual mechanisms. Not only does the proposed method take into account spatial information content, as in almost all the known methods, but it also considers the temporal information as the human visual system does. The introduction of the time parameter in the treatment is justified by some psychovisual and physiological experiments as reported in [2,3]. It should be noticed that the time parameter is also introduced in the anisotropic diffusion equation used in image enhancement [17], in contour detection [5] and in image filtering [6]. However, in all these methods, the time variable has not its actual significance. For example, in the Perona and Malik's method and in Gerig et al's one, the t parameter plays the role of the iteration number. Moreover, the cited computational approaches for contour detection consider the image as static information: the image pixels are taken as a whole set at a given instant of observation. Obviously, this is of a great advantage when the image is processed by a parallel architecture machine. In fact, some physiological and psychovisual experiments show that when an image is captured by the human visual system, all the image elements are not processed at the same time. These results have been used as the basic guideline in the design of architecture for asynchronous visual indexing system [24]. Following this idea, an asynchronous contour detection scheme is proposed in the present contribution. As in [1], the asynchronous model, presented in section II is based on the transformation of a static image into a temporal data flow. One of the novel features of the present model is that it considers only the innovation in the image temporal analysis. We call this phenomenon ″ exclusive asynchrony.″ Another specificity of the proposed approach rests upon the treatment of this data flow for edge detection. Instead of using edge operators in each temporal stage, we use the method of Kundu and Pal [7], which is based on the well-known Fechner-Weber law. Indeed, one way to quantify our ability to resolve two stimuli, which are the same except for their intensities or luminances, is to measure the just-noticeable difference (J.N.D). The JND is used as a basic guideline in Kundu and Pal's method in order to detect the meaningful edges in the observed scene. Thus, this paper deals with the processing of discrete pictures based on these visual properties. We believe that the decomposition of the image into a data flow allows an easier treatment and makes the scene interpretation less ambiguous. Indeed, in real-world vision, the images often contain partial information of several different types in each part of the image. The aim of this paper is not to compare the proposed method with all the existing contour detection algorithms. Our main purpose is to demonstrate the efficiency of a model based on visual perception mechanisms for contour detection and to improve the methodology such as proposed by Kundu and Pal.

THEORY OF THE MODEL

The proposed model consists of five stages, as done approximately in the visual system. This contour detection method is computationally expansive but it is the price to pay for good contour localization.

II.1 The noise reduction step

The first stage of the model (fig 1) concerns noise reduction.

The noise is generally supposed to be additive, uncorrelated to the image, and generally localized at high frequencies. To reduce the noise effect, we apply to the image signal a low filter, with a gaussian impulse response. The choice of this filter is motivated by the separability that it allows within the variables, which is time saving, and by the fact that it offers a good trade-off between good noise reduction and good preservation of the details in the image. The application of this linear filter, the impulse response of which is h (x,y), to the image signal $I(x,y)$ provides a signal A (x, y), given by the following convolution:

$$A(x,y) = I(x,y) * h(x,y) \qquad (1)$$

II.2 Exclusive asynchrony

In the second stage, we introduce the notion of asynchrony. Asynchrony means that the various levels of the image are not processed at the same instant [1]. The experience of Hess [2], shows that if one considers two neighbor objects, a clear one and a darker one, together in movement in an orthogonal direction to the line joining them, the clear object is perceived in advance of the darker one, even when they have the same speed. The pendular experience of Pulfrich confirms this phenomenon. A clear static object is perceived more rapidly than a dark object [3]. Thus, a maximum of intensity or luminance I_{max} is associated with weak response delay τ_{min} and reciprocally, to a minimum of luminance I_{min} corresponds a high response delay τ_{max}. Delays τ_{min} and τ_{max} are supposed equal to 2 ms and 22ms respectively. The interval between such values appears to be physiologically a plausible one [4]. In the visual system, the relationship between the luminance and the time necessary to the transfer of the signal (time delay) is a non-linear law. The delay decreases rapidly for strong luminances and remains almost constant for weak ones. In the model, as in [2,3] the luminance obeys the rule:

$$I(t) = \frac{\delta}{t^3} \qquad (2)$$

Consequently, the image is transformed into a data flow according to this law.

The coefficient δ is given by: $\delta = A_{max}.\tau_{min}^3$ (or $A_{min}.\tau_{max}^3$).

A_{max} represents the maximum level in the perceptible luminance range of the actual image. At the instant t within the interval [τ_{min}, τ_{max}], all the points of the image are processed if the associated luminance verifies the condition:

$$A(x,y) \geq \frac{A_{max}(\tau_{min})^3}{t^3} \qquad (3)$$

In the present method, only the innovation at a given instant, as compared to the preceding instants, is taken into account, which means that, at a given instant t, only new intensities are considered. That leads to what we call, ″exclusive asynchrony″. Thus, if B (x, y, t) indicates the innovation at a given instant t, one has:

$$A(x,y,t) < B(x,y,t) \leq A(x,y,t+dt) \qquad (4),$$

where B (x, y, t) represents the innovation. As it will be seen in section III, the choice of the time step dt is determinant for the results.

II.3 Isotropic linear filtering and adaptive smoothing

A band pass isotropic filter is applied to each element of the data flow. This filter is used to smooth homogeneous regions and to enhance edges. The resulting signal C (x, y, t) is linked to B (x, y, t) by the following convolution:

$$C(x,y,t) = \left(\rho.\exp-\frac{x^2+y^2}{2.\sigma_2^2} - \zeta.\exp-\frac{x^2+y^2}{\sigma_3^2}\right) * B(x,y,t), \qquad (5)$$

where ρ and ξ are normalization constants, such as $\rho.\sigma_2^2 \geq \zeta.\sigma_3^2$ [21]. This operation is followed by a diffusion process, which controls the appearance of edges in regions of strong contrasts and to homogenize regions with small variations of luminance. We perform one iteration of the adaptive smoothing used in [6,8] by convoluting B (x, y, t) with a decreasing function g(x,y,t). The function g(x,y,t) is similar to the one used in [5] and depends on the difference signal C(x,y,t. Then the obtained signal is:

$$D(x,y,t)=B(x,y,t)*g(x,y,t)$$
$$g(x,y,t)=\frac{1}{1+\left(\frac{C(x,y,t)}{K}\right)^2} \quad (6).$$

Where K is an adaptation constant [5].

II.4 Dynamic thresholding

This operation is the stage of decision in our model. The well-known Weber-Fechner law is performed to choice the contour detection thresholding in each image element of the data flow. This rule, for a given object, provides with the just noticeable difference ΔL of luminance (J.N.D), or threshold of luminance, according to the background luminance L [25]:

$$\Delta L = K_1 L \quad (7)$$

Where, K_1 is the Weber constant, and its order is somewhat of 2%. This law is used as a basic criterion to detect significant edges.

a° Local contrast

One calculates at every pixel of the image obtained from the precedent stage, a local contrast. This contrast is given by:

$$E(x,y,t)=\frac{\Delta D(x,y,t)}{\overline{D(x,y,t)}} \quad (8)$$

The numerator of the equation (8) is nothing more than the gradient gray level. It is given, by the average of the different responses obtained through the application of a series of directional Gabor's filters. The choice of the Gabor's filter is justified referring to the visual cortex behavior [9]. If one calls $G^k(x,y)$ the response of the Gabor's filter in the θ_k direction, and $F^k(x,y)$ the response associated with this filter, then one obtains:

$$F^k(x,y,t)=D(x,y,t)*G^k(x,y) \quad (9).$$

The mean value of the gradient, for N directions, is:

$$\Delta D(x,y,t)=\frac{1}{N}\sum_k F^k(x,y,t) \quad (10).$$

The expression of $G^k(x,y)$ is the one proposed by Daugman [9]:

$$G^k(x,y)=\psi.\exp-\left(\frac{(x.\cos\theta_k+y.\sin\theta_k)^2}{2.\sigma_4^2}+\frac{(-x.\sin\theta_k+y.\cos\theta_k)^2}{2.\sigma_5^2}\right) \quad (11)$$
$$*\cos 2\pi F(x.\cos\theta_k+y.\sin\theta_k)$$

Where ψ is a factor of normalization. σ_4 and σ_5 define the size of the filter. They are such that $\frac{\sigma_5}{\sigma_4}=1.88$ as reported in [10] (for simple cells of the visual cortex V1 associated with the X ganglion cells of acute vision [20]). The denominator in equation (8), $\overline{D(x,y,t)}$, represents the average level of the close neighbors of the considered pixel in a window of size w=nxm. The associated expression, as used by Pal and Kundu [7], adapted to our case, is

$$\overline{D(x,y,t)}=\frac{1}{2}.\left\{D(x,y,t)+\frac{1}{w-1}.\sum_{p,q}\frac{D(x-p,y-q)}{\sqrt{(x-p)^2+(y-q)^2}}\right\} \quad (12).$$

The gray levels of neighbor pixels are weighted by the inverse of the distance that separates them from the center of the window. In accordance with the idea of Kundu and Pal, a point of the image at the instant t is said to belong to the contour if its contrast, defined by equation [8], satisfies the condition:

$$E(x,y,t)\geq E_{min} \quad (13).$$

Where E_{min} represents the minimum contrast related to the maximum contrast E_{max} by :

$$E_{min}=K_1.E_{max} \quad (14)$$

b° Decision stage

The maximum contrast is computed by analyzing the whole image. The result of the step is a binary one and it is provided by:

$$T(x,y,t)=\begin{cases}255 & \text{if } E(x,y,t)\geq E_{min}\\ 0 & \text{elsewhere}\end{cases} \quad (15).$$

II.5 Temporal integration

The different images, that are provided by the preceding stage are superimposed to constitute the final image S (x, y) [11, 12]:

$$S(x,y)=\sum_{\tau_{min}}^{t} T(x,y,t). \quad (16).$$

III Experimental results and discussion

In this section, we present some results obtained by our model. We also discuss on the choice of the different parameters used for the model. As previously said, the use of the temporal processing can avoid all saturation of the segmented image. It allows a better recognition of the object, by gradually presenting the object characteristics. The most prominent information in the image appears at first sight.

In the preceding relationship between luminance and time (equation (2)), the luminance decreases rapidly. There is instant t_s at which the saturation starts. In order to limit this saturation, we stop the image processing at t_s. The instant t_s can be written: $t_s=\tau_{min}+\gamma.dt$, where dt is the time step, depending on the characteristics of the image, and γ, the number of iterations. The various parameters that have been cited in the theoretical part of the paper are associated with the following values:

$$\sigma_2=0.37, \sigma_3=1.14\ \rho=100, \xi=6 \text{ (isotropic filtering).}$$
$$\sigma_4=0.66, \sigma_5=1.24; F=0.5, \psi=100 \text{ (Gabor filters)}$$
$$K=10, K_1=4\%, N=4.$$

The standard deviation of the gaussian filter is assumed 0.25. This method is applied to synthetic and actual images.

III.1° Synthetic image

In order to test the reliability of our algorithm, we have synthesized a test image. The image is constituted by two regions. A background of luminance value L_b and a central object of luminance L_0, and such that $L_b > L_0$ (figure 2a). In order to simulate a real situation, a gaussian noise is added to the image obtained after a smoothing process. The luminance of the object is 64, and the one of the background, 128. The result obtained trough the application of the proposed model on the test image is presented in figure 2b. The contour lines are clearly perceived and the central object is well separated from the background. Furthermore, a substantial reduction of the noise is noticed. Nevertheless, if one increases the number of iterations γ, an amplification of the noise occurs, and the contours become less visible (figure 2c). The time step is equal to 0.25 in both cases. For these two cases, it could be noticed that the squared borders are well localized. This result is an objective criterion to judge of the efficiency of the proposed method.

III.2° Biomedical application

The biological area could be an interest field for the application of the method. One knows that a medical image often presents a weak contrast that makes ineffective the operators of classic differentiation. Figure 3b presents the result that we have obtained on a cross section of the brain (figure 3a). The various structures in this section are clearly visible. As compared to Kundu and Pal’s approach (Figure 3c), our approach seems to bring more information. In the considered case, the time step is equal to 0.5 and the number of iterations, equal to 6. The parameters used in the comparison correspond to the optimization of both methods.

CONCLUSION

We have tried to show that algorithms based on human vision can improve the detection of contours in the image. The ″exclusive asynchrony″ processing that we apply can simplify the recognition of objects. Of course, controlling the diffusion of prominent elements of the data flow is supposed. The method allows a good location of contour lines and a reduction of the noise.

Exclusive asynchrony could be used in several areas and applications, such as image compression.

BIBLIOGRAPHY

[1] P-Y. Burgi and T. Pun ″ Asynchrony in image analysis: using the luminance-to-response latency relationship to improve segmentation″, J.Opt.Soc.Am.A, vol. 11, 1720-1726,1994.

[2] J.M. Williams and A. Lit, ″Luminance-dependent visual latency for the Hess effects, the Pulfrich effect, and simple reaction time, ″Vision Research, vol. 23,171179, 1983.

[3] R.J.W. Mansfield and J.G. Daugman ″ Retinal mechanisms of visual latency″. Vision Research, vol. 18, 1247-1260,1978.

[4] J.H.R. Maunsel and J.R. Gibson ″ Visual response latencies in striate cortex of macaque monkey ″. J.of Neurophysiology, vol. 68, 1332-1334,1992.

[5] P. Perona and J. Malik ″ Scale-Space and Edge Detection Using Anisotropic Distribution ″, IEEE Trans. on Pattern Analysis and Intelligence Machine, vol. 12, 629-639, 1990.

[6] P. Saint-Marc, J.S. Chen, and G. Medioni, ″ Adaptive Smoothing: A General Tool for Early Vision ″. IEEE Trans. on Pattern Analysis and Intelligence Machine, vol. 13, 514-529, 1991.

[7] M.K. Kundu and S.K. Pal, ″ Thresholding for edge detection using human psychovisual phenomena″. Pattern Recognition Letters, vol. 4, 433-441, 1986.

[8] G. Gerig, O. Kubler, R. Kikinis and.F.A. Jolesz, ″Nonlinear Anisotropic Filtering of MRI Dated″. IEEE Trans. on Medical Imaging, vol. 11, 221-232, 1992.

[9] J.G. Daugman, ″Uncertainty relationship for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters″. J.Opt.Soc.Am.A, vol. 2, 1160-1169, 1985.

[10] S.J. Anderson, D.C. Burr and M.C. Morrone, ″Spatial two-dimensional and spatial-frequency selectivity of motion-sensory mechanisms in human vision″. J.Opt.Soc.America., vol. 8,1340-135, 1991.

[11] J.J. Hopfield, ″ Neurons, Dynamics and Computing″. Physics Today, 6, 40-46,1994.

[12] P-Y. Burgi, ″Analysis and understanding of neuronal structures″ Neurocomputing, vol. 4, 137-166,1992.

[13] J.M.S. Prewitt, ″ Object Enhancement and extraction ″, Picture Processing and Psychpictorics, Academic Press, 1970.

[14] D.C. Marr and H. Hildreth, ″ Theory of edge detection ″, Proc.Roy.Soc. London, vol. B 204, 187-217, 1980.

[15] J. Canny, ″ A Computational Approach to Edge detection ″, IEEE Trans. Pattern Analysis and Machine Intelligence, 679-698, 1986.

[16] J.J. Koenderink and A.J. Van Doorn, ″ Visual detection of spatial contrast; influence of location in the visual field, target extent and illuminance level ″. Biol. Cybernetics, vol. 30, 157-167, 1978.

[17] L.S.G. Kovasznay and H.M. Joseph, ″ Image Processing ″, Proc.IRE, vol. 43, 560-570, 1955.

[18] V.S. Nalwa and T.O. Binford, ″ On detecting edges ″, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 8, 699-714, 1986.

[19] L.S. Davis, ″A survey of edge detection techniques″, Computer Graphics and Image Processing, vol. 4, 248-270, 1975.

[20] P. Buser and M. Imbert, ″ Neurophysiologie Fonctionnelle IV ″, Vision, Hermann Paris, 1987.

[21] C. Koch and I. Segev ″ Methods in Neuronal Modeling ″, 335-360, the MIT Press, 1989.

[22] I.E. Abdou and W.K. Pratt, ″ Quantitative design and evaluation of enhancement / thresholding edge detectors ″, IEEE 67, 753-763, 1979.

[23] M. Petrou and J. Kittler, ″ Optimal edge detectors for ramp edges ″, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 13, 483-491, 1991.

[24] R. Milanese, J.M. Bost and T. Pun, ″ A bottom-up attention system for active vision ″, ECAI 92. 10th European Conference on Artificial Intelligence, B.Neuman Ed.

[25] T.N. Cornsweet, ″ Visual Perception ″, Academic Press 1970.

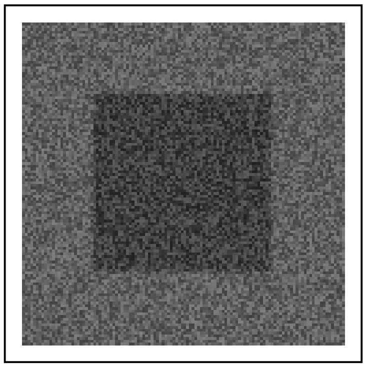

Figure 2a
Test image

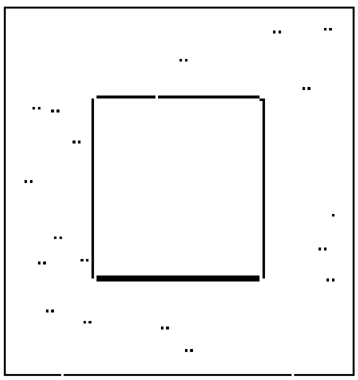

Figure 2b
dt=0.25, γ=1

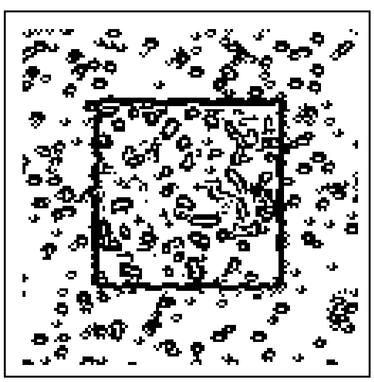

Figure 2c
dt=0.25, γ=3

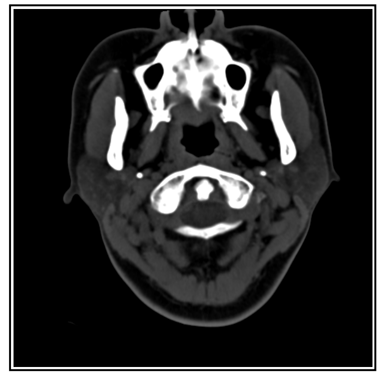

Figure 3a
Cross section of
the brain

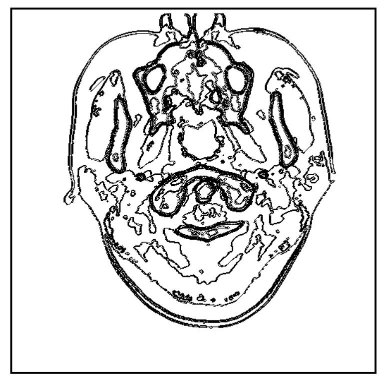

Figure 3b
Our Model
dt =0.25, γ=6

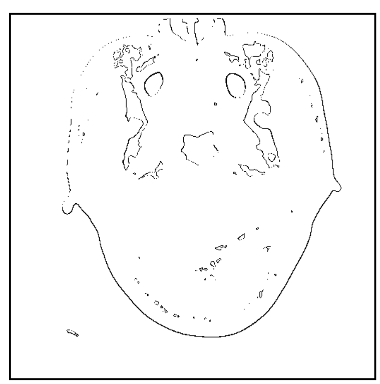

Figure 3c
Kundu and Pal model
α_1=0,α_2=0.1,α_3=0.9,β=4%

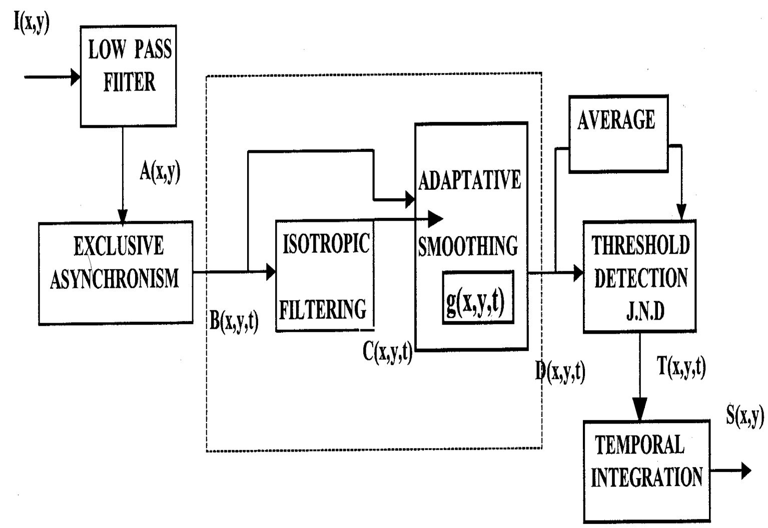

FIGURE 1 : DIAGRAM OF THE MODEL