Fifth International Symposium on Signal Processing and its Applications, ISSPA '99, Brisbane, Australia, 22-25 August, 1999
Organised by the Signal Processing Research Centre, QUT, Brisbane, Australia

Low-level vision treatments inspired from Human Visual System

Azeddine BEGHDADI, Kamel BELKACEM-BOUSSAÏD* and Abdelouahab BOUDRAA
L2TI, Institut Galilée, Université Paris 13, Avenue J.B. Clément, 93430 Villetaneuse, France.
*Beckman Institute, University of Illinois at Urbana Champaign, 405 N. Mathews Avenue 61801, IL Urbana, USA.
E-mail:Azeddine.Beghdadi@l2ti.univ-paris13.fr

Abstract
In this paper a set of low-level vision treatments inspired from some knowledge of retino-cortical mechanisms of mammalian is proposed. The basic idea is to exploit the contrast sensitivity response of the Human Visual System (HVS) and the multichannel decomposition of the visual signal in the cortex. The scope of this work is limited to nonlinear filtering, contrast enhancement and edge detection but could be extended to other image treatments.

1. Introduction

The number of low-level vision treatments is continuously increasing and today there is no universal criteria or quality assessment measure to decide if for a given application a specific method is more adequate than others. For example, to compare noise filtering methods some objective measures such as NMSE (Normalized Mean Square Error), MAE (Mean Absolute Error) or NCC (Normalized Correlation Coefficient) are often used. In compression methods the PSNR is preferred to other objective measures. Nevertheless, very often the ultimate subjective judgment of human remains the most efficient assessment. Our aim is to demonstrate the efficiency of a model based on visual perception mechanisms for low-level vision treatments. The purpose of this paper is to propose a new approach to designing nonlinear filtering, contrast enhancement and edge detection based on some properties of the HVS. The basic idea is to exploit the contrast sensitivity response of the HVS and the multichannel decomposition of the visual signal in the cortex.

The outline of the paper is as follows. Sections 2 describes the Moon-Spencer optical contrast and its adaptation to digital images followed in section 3 by the proposed three low-level treatments. Experimental results and discussion are presented in section 4.

2. Background

Contrast perception in a natural image depends on several parameters, such as the size of the observed object, the luminance of the background, the visual angle and the spatial frequency of the observed pattern [1]. Taking all of these parameters into account using a single contrast analysis method is a rather difficult task. Thus, we will limit this study to the effect of background luminance on contrast perception. Specifically, we will deal with the Just-Noticeable Contrast (JNC) [2]. The traditional Weber-Fechner definition of contrast is useful for a uniform target embedded in a uniform background, but breaks down when either the target or the background, or both, is structured (non-uniform) [3]. Moon and Spencer have studied non-uniform surrounds and have derived a modified analytic expression for JNC using Holliday's principle [4]. The basic idea of this principle is that a non-uniform luminance distribution can be replaced by an equivalent uniform distribution called adaptation luminance.

2.1 Moon-Spencer's JNC and its adaptation to digital images

In this work we will also introduce some ideas derived from studies of primate cortical mechanisms [5]. The idea of using the JNC for low-level vision has appeared in some previous works, but in a simplified and unrealistic form [6],[7]. In a complex visual scene, the background is not uniform. It is composed of regions of different luminance and spatial frequencies. In such cases, a local band-limited contrast would be relevant for an efficient image analysis as shown by Peli [8]. To improve Kundu and Pal's method [6], we use the definition of contrast proposed by Moon and Spencer [1], and adapt it for use on digital images. We will therefore perform a local analysis of the image contrast as done in [8],[9].

Let us recall Moon and Spencer's results as reported in [1]. It was found that both the immediate surround luminance B_c (x, y) and the global luminance B_f (x, y) should be taken into account in the computation of the contrast. Another relevant quantity is the adaptation luminance B_a (x, y). This is a corrected luminance value related to both B_c and B_f. It is based on the way in which apparent background luminance is treated by the visual systems of human psychophysical subjects [1],[4], and was not employed by Weber and Fechner. The B_a (x, y) value has been determined by Moon and Spencer [1]:

$$B_a = 0.923 B_c + 0.077 B_f \qquad (1)$$

Using Holladay's principle and other previous results, Moon and Spencer proposed a new definition of JNC (C_{min}) based on the concept of adaptation luminance [1]. They considered two different formulations. Three decades later, Kretz [10] expressed C_{min} as a function of B_a (x, y) and B_c (x, y), as follows:

$$C_{\min} = \begin{cases} \dfrac{C_w}{B_c}\left(A + \sqrt{B_a}\right)^2 & \text{if } B_a \geq B_c \\ \dfrac{C_w}{B_c}\left(A + \sqrt{\dfrac{B_c^2}{B_a}}\right)^2 & \text{if } B_a < B_c. \end{cases} \qquad (2)$$

In this expression, A is an experimentally measured constant equal to 0.8 and C_w is the Weber-Fechner constant.

2.2. Our contrast

In our approach, the local contrast C (x, y) of a pixel, considered as the center of an analysis window, is computed and compared with the JNC as defined in equation (2). As in Kundu and Pal's method [6] :

$$C_W = \beta . C_{max}, \qquad (3)$$

where C_{max} is the maximum contrast, and β is a parameter, ranging from 0.02 to 0.1. Our local contrast is defined as follows:

$$C(x,y) = \frac{\Delta B(x,y)}{B_c(x,y)} \tag{4}$$

where B_c (x,y) is the average luminance of the region around the considered pixel, and $\Delta B(x,y)$ is the average response of a bank of band-pass Gabor filters applied to the original signal. Here, the classical definition of contrast is improved by the use of Gabor filters to decompose the image into components of different spatial frequencies and directions. This new definition is similar to that introduced by Peli [8] in order to take into account human frequency contrast sensitivity near threshold. Peli used an isotropic Gaussian envelope. Here, we use an anisotropic Gabor filter based on the physiological properties of visual cortical simple cells [11]. These filters are optimal for localizing a stimulus in space and spatial frequency [11]. Moreover, it has been shown that this transform achieves an efficient decomposition of the image into meaningful band-pass oriented components [5]. Let us call $G^k(x,y)$ the impulse response of such filter in the θ_k direction as defined in [11], and $F^k(x,y)$ the response associated with this filter. Then, for each k index, a band-pass filtered signal is obtained via the relation :

$$F^k(x,y) = B(x,y) * G^k(x,y). \tag{5}$$

Where B (x,y) is the original signal and * the convolution operator. In this study we use for $G^k(x,y)$ the expression proposed by Burgi and Pun in [12] :

$$\begin{aligned} X &= (x\cos\theta_k + y\sin\theta_k)^2 \\ Y &= (-x\sin\theta_k + y\cos\theta_k)^2 \\ Z &= \cos 2\pi v(x\cos\theta_k + y\sin\theta_k) \\ G^k(x,y) &= \psi\exp-\left(\frac{X}{2\sigma_1^2} + \frac{Y}{2\sigma_2^2}\right)Z \end{aligned} \tag{6}$$

Where ψ is a factor of normalization, σ_1 and σ_2, which define the size of the filter, are such that $\frac{\sigma_2}{\sigma_1} = 1.88$, as reported in [12] (for receptive fields of simple cells of primary visual cortex), and v is the spatial frequency of the lobes. In our case θ varies, but σ_1 and σ_2 are fixed. Notice that only the even-symmetric component of the Gabor filter is used as done in [13]. Once the band-pass components of the signal are captured, the average local band-pass signal along N orientations is computed, as reported in [14], using:

$$\Delta B(x,y) = \frac{1}{N}\sum_k F^k(x,y) \tag{7}$$

Here, for the sake of simplicity, we did not introduce competitive mechanisms as done in [14]. This simple summation model is of limited value, as are many models based on mammalian vision in which the channel's responses are linearly combined [13]. Here, the average orientation bandwidth of the receptive field of a cortical simple cell is assumed to be about 45°, as in [5],[15]. Thus, only the four orientations of 0°, 45°, 90° and 135° are used (i.e. N=4).

The proposed local contrast, as defined by eq.(4), is obtained by dividing ΔB by the average luminance of the surround computed in the window analysis according to the relation :

$$B_c(x,y) = \frac{\sum_{(i,j)\in\Omega^1_{xy}} B(i,j)\varphi_{xy}(i,j)}{\sum_{(i,j)\in\Omega^1_{xy}} \varphi_{xy}(i,j)} \tag{8}$$

Where $\varphi_{xy}(i,j) = \left[(i-x)^2 + (j-y)^2\right]^{\frac{1}{2}}$ and $\Omega^1_{xy} = \{(i,j)/\, d_1 < d\{(x,y),(i,j)\} \le d_1 + d_2\}$, and $d\{(x,y),(i,j)\}$ is the Euclidean distance associated with the two points. d_1 is the extent of the object and d_2 is the thickness of the surrounding region. Notice that the weighting function $\varphi_{xy}(i,j)$ gives more weight to the nearest neighboring pixels and less to more distant ones [16].

3. Low-level treatments

An image can be decomposed into two components. The first one is composed of well contrasted textures and contours. This component corresponds to well perceptible points. The second one is composed of low contrasted contours, textures, homogeneous regions and the background. This component corresponds to points which are not well perceptible. Thus, using a measure of contrast one may separate the original image into two components.

3.1 Nonlinear filtering

It is well known that noise is more annoying in homogeneous regions and low contrasted ones. Thus, only low contrasted points will be smoothed according to the following nonlinear filtering rule:

$$B_{out}(x,y) = \begin{cases} B(x,y) & C(x,y) \ge C_{min} \\ B_{new} & C(x,y) < C_{min} \end{cases} \tag{9}$$

where $B_{new} = \max(B_c, B_0)$, with $B_0(x,y)$ given by :

$$B_0(x,y) = \frac{\sum_{(i,j)\in\Omega^0_{xy}} B(i,j)\varphi_{xy}(i,j)}{\sum_{(i,j)\in\Omega^0_{xy}} \varphi_{xy}(i,j)} \tag{10}$$

with $\varphi_{xy}(i,j) = \left((x-i)^2 + (y-j)^2\right)^{-\frac{1}{2}}$ and $\Omega^0_{xy} = \{(i,j)/\, d\{(x,y),(i,j)\} \le d_1\}$.

For more details concerning the above expressions see [16]. An alternative method for computing B_{new} (x, y) might use the "min" operator instead of the "max" one, or better the linear combination of $B_0(x,y)$ and $B_c(x,y)$: $B_{new} = \alpha_C B_C + \alpha_O B_O$ where α_C and α_O are the fractions of crown and object pixels, respectively.

Unlike Kundu and Pal's method a point of noise would be eliminated using our JNC rule, since only average values are used in the computation of local contrast. Our non-linear transformation is based on the earlier

work of Weber and Fechner and of Holladay, which proposed that an object is visible only if its contrast is greater than or equal to the perceived JNC (C_{min}.). By our method, a noisy area with small gradient values is smoothed, while an area with large luminance variations remains unaffected. Well contrasted regions are preserved, and noise is smoothed out.

3.2 Edge detection

Our approach was inspired by that of Kundu and Pal [6], in which an edge thresholding method based on Weber-Fechner's law was proposed. Thus, a pixel (x,y) is identified as a contour point if its local contrast C(x,y) is greater than C_{min} .The parameter β is comparable to the scale parameter of the LOG filter. A low β value yields many contours while high β value gives just the most important ones.

3.3 Adaptive Contrast enhancement technique

The main difficulty of contrast enhancement is to enhance the visibility of relevant information in an image without amplifying noise. However, most enhancing methods suffer from noise sensitivity. Indeed, enhancing details results in unavoidable noise amplification. Here, for contrast enhancement purpose, we use the following strategy. First, the image is decomposed into two part : the first set of pixels, say S_0, consists of zones with noise amplification risk and the complementary, say S_1, is composed of zone without risk. This classification is performed on the basis of the contrast measure and the JNC. A given pixel is considered as risk element if the associated local contrast is below the JNC otherwise it belongs to S_1. Then, a risk element is filtered using the previous described nonlinear filter otherwise is is contrasted using the method described in [9]. In summary, we apply the following treatment :

$$B_{out}(x,y)=\begin{cases} B_{new} & C(x,y)\leq C_{min} \\ B_1(x,y) & C(x,y)>C_{min} \end{cases}$$

Where, B_{new} has the same meaning as before and $B_1(x,y)$ is the contrasted gray-level of the pixel.

4. Results and discussion

The proposed low-level treatments are tested on various types of images with equivalent results. Only one of these images is shown to illustrate the efficiency of these treatments. For nonlinear filtering method, a tested image (Fig. 1a) is degraded by an additive white Gaussian noise with zero mean and a variance σ^2=169 (Fig. 1b). Result after filtering is shown in Fig.1c. Qualitatively, it can be observed that our method filters out the noise without affecting the fine details (Fig. 1c). This result clearly shows the usefulness of a psychophysically based contrast measure for noise filtering purposes. The filter in the simulation operates on 7x7 window samples. Since our parameter is related to local contrast, proper filter adjustment can be automatically achieved. Note that a large value of the parameter β yields a high JNC and consequently the amount of image smoothing is very large. A low β value yields a low JNC and therefore the smoothing effect is not impressive. To show the reliability of the edge detection method, a medical image is used (Fig. 2a). It is well known that medical images, such as in computed tomography or in magnetic resonance imaging often present weak contrast that makes ineffective the operators of classic differentiation. The result of edge detection obtained with β=0.1 is shown in Fig.2b. A careful examination of this figure show that the edges are well detected and are connexe. Various medical structures are clearly visible. To show the effect of adaptive contrast enhancement a portrait test is used (Fig. 3a). Notice that the original image Lenna is of good quality and consequently does not require a contrast enhancement. However, homogeneous regions such as cheeks exhibit granular aspect. Thus, contrast enhancement such as proposed in [9] or histogram equalization increase this granular aspect. The result of adaptive contrast enhancement is illustrated in Figure 3b. Note that the image is well contrasted, fine structures and contours are well preserved and the granular aspect has not been increased.

In this paper low-level treatments derived from the mammalian visual system have been developed. Based on Moon and Spencer's contrast, the presented results show the effectiveness of the proposed method. These results reinforce the idea that biologically inspired methods may constitute a new framework for image processing treatments. However, we do not claim that the proposed methods perform well or are more effective than the classical ones. Thus, a large scale of experiments and comparisons with other methods must be performed. Furthermore, the influence of the parameters such as β must be investigated to show the robustness of these methods. We will also investigate the asynchrony, derived from some psychovisual experiments such as Pullfrich and Hess effects [12],[14], for contrast enhancement and noise filtering.

5. References

[1] P. Moon and D. E. Spencer, "Visual data applied to lighting design ", *J. Opt. Soc. Am. A*, vol. 34, pp. 230-240, 1944.

[2] G. T. Fechner, "Elemente der psychophysic", Breitkopf und Härtel, Leipzig, vol.II, ch.XVI, 1860.

[3] O. Lillesæter, "Complex contrast, a definition for structured targets and backgrounds", *J. Opt. Soc. Am. A* vol. 10, no. 12, pp. 2453-2457, 1993.

[4] P. Moon and D. E. Spencer, "The Visual effect of nonuniform surrounds", *J. Opt. Soc. Am. A*, vol. 35, pp. 233-248, 1945.

[5] R. L. DeVallois, E. W. Yund and N. Hepler, "The orientation and direction selectivity of cells in macaque visual cortex", *Vis. Res.* vol. 22, pp. 531-544, 1982.

[6] M. K. Kundu and S. K. Pal, " Thresholding for edge detection using human psychovisual phenomena", *Patt. Recogn. Lett.* , vol 4, pp. 433-441, 1986.

[7] K. Belkacem-Boussaid, A. Beghdadi and H. Dupoisot, "Edge detection using Holladay's principle", *Proc. ICIP.*, vol. 1, pp. 833-836, Lausanne,september, 1996.

[8] E. Peli, "Contrast in complex image", *J. Opt. Soc. Am. A* vol. 7, no. 10, pp. 2032-2040, 1990.

[9] A. Beghdadi and A. Le Négrate, "Contrast enhancement technique based on local detection of contrast", *CVGIP*, vol. 46, pp. 162-174,1989.

[10] F. Kretz, "Subjectively optimal quantification of pictures", *IEEE Tran. Commun.*, vol. 25, pp. 1288-1292, 1975.

[11] 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, pp. 1160-1169, 1985.

[12] P-Y. Burgi and T. Pun, "Temporal analysis of contrast and geometric selectivity in the early human visual system", *in Channels in the visual nervous system : Neurophysiology, Psychophysics and Models*, pp. 237-288, B. Blum ed. Freund, London, 1991

[13] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms", *J. Opt. Soc. Am. A* vol. 7, no. 5, pp. 923-932, 1990.

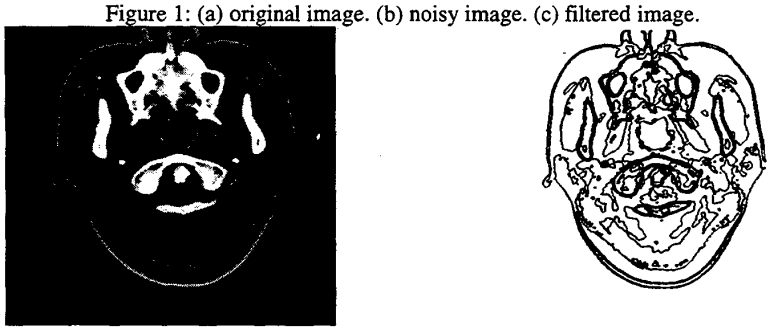
[14] P-Y. Burgi and T. Pun, "Asynchrony in image analysis : using the luminance-to-response-latency relationship to improve segmentation", *J. Opt. Soc. Am. A* vol. 11, no. 6, pp. 1720-1726, 1994.

[15] S. Marcelja, "Mathematical description of the responses of simple cortical cells", *J. Opt. Am.* vol. 70, pp. 1297-1300, 1980.

[16] K. Belkacem-Boussaïd and A. Beghdadi, "A new image smoothing method based on a simple model of spatial processing in the early stages of human vision", *IEEE. Trans Image Processing*, 1999 (in press).

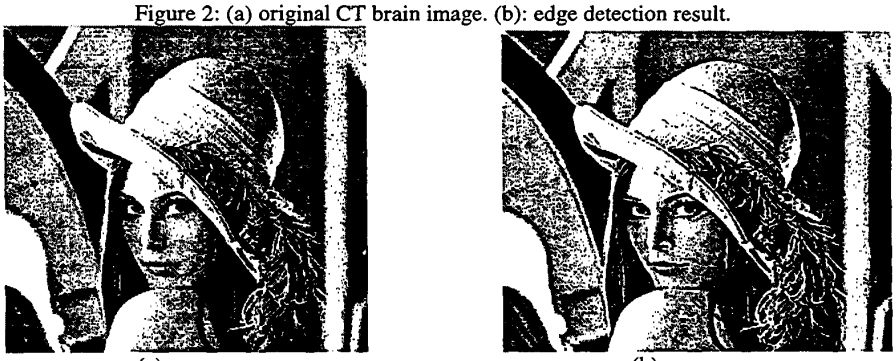
(a) (b) (c)

Figure 1: (a) original image. (b) noisy image. (c) filtered image.



(a) (b)

Figure 2: (a) original CT brain image. (b): edge detection result.



(a) (b)

Figure 3: (a) original image. (b): Adaptative contrast enhancement result.