DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 20 au 24 MAI 1985

IMAGE REPRESENTATION IN VISION USING A GENERALIZED GABOR SCHEME

Moshe Porat and Yehoshua Y. Zeevi

Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel

RESUME

SUMMARY

The model of visual information processing, in a combined frequency-position space, is investigated through image decomposition into a finite set of Gabor elementary functions. A set of the corresponding expansion coefficients represents according to the model an image. The Gabor scheme is extended to two spatial dimensions and generalized to account for the inhomogeneity, over-sampling, frequency octave relations and phase quantization. Comparison of reconstructed images with the original highlights the advantages of the generalized Gabor scheme . This work suggests that the generalized Gabor scheme, modeled after cortical signal processing, is attractive for image processing.

Image representation in vision using a generalized Gabor scheme

I. Introduction

The advancement of a model for image processing in vision based on the Gabor scheme is motivated by both biological and computational considerations [1]-[6]. Gabor functions were first introduced as a discrete set of one-dimensional elementary functions by means of which signals can be represented [7]. Their main advantage is in achieving the lowest bound of the joint entropy defined as the product of effective spatial extent and bandwidth. Hence, representation of a signal by these functions provides the best spectral information for every point along the signal variation. In cases where global analysis of the entire signal is not practical as in speech or vision, there is an advantage in using this representation over global (e.g Fourier) transforms which describe the spectrum of the entire signal as a whole.

Although Gabor functions were proposed as early as 1946, no analytic solution suitable for determination of the expansion coefficients was available until recently (though Gabor himself proposed in his original paper an approximate solution). An analytic method for the one-dimensional case, was presented first by Bastiaans [8] in 1980, paving the way for further research. Interestingly enough, at about the same time several investigators have proposed the Gabor functions as models for cortical cells' receptive field profiles [1]-[3],[6], however none of these attempted to elaborate a complete scheme as a computational model which can be further tested [9].

The purpose of this study is to develop a general scheme of image representation in a combined position-frequency space accounting for the biological position-dependent sampling rate, oversampling, octave relations between central frequencies and phase quantization; All of these in order to better understand the organizational principles of the visual system, and apply this knowledge in image processing and computer vision.

II. The Two-dimensional Gabor scheme

The Gabor scheme was originally suggested for processing and communication of one-dimensional signals [7]. Considerations of the two-dimensional case were restricted to the optimal properties of a single elementary function, but a complete two-dimensional scheme was not, to the best of our knowledge, presented. Such a generalization is obviously needed for image processing and may also be useful for other applications.

Our generalization is first presented in Cartesian and then in polar coordinate system; the latter is more appropriate for modeling of the visual system.

In the case of a (spatial) 2D scheme the coefficients' space is four-dimensional: x,y for positions and f_x,f_y for frequencies ($r,\vartheta,f_r,f_\vartheta$ in a polar coordinate system). It can be shown [10] that the 2D presentation of a signal $\Phi(x,y)$ can be expressed by:

$$\Phi(x,y) = \sum_{m_x}\sum_{n_x}\sum_{m_y}\sum_{n_y} a_{m_x n_x m_y n_y} \cdot f_{m_x n_x m_y n_y} \tag{1}$$

(Unless otherwise stated, all integrations and summations in this paper extend from $-\infty$ to $+\infty$), where the two dimensional elementary function is:

$$f_{m_x n_x m_y n_y} \triangleq g(x-m_x D_x, y-m_y D_y) \cdot \exp(i n_x W_x x + i n_y W_y y) \tag{2}$$

with $g(\cdot)$ being the separable Gaussian window function:

$$g(x,y) = g_x(x) \cdot g_y(y) = \frac{\sqrt{2}}{D}\exp\left[-\pi\left(\frac{x}{D}\right)^2\right]\cdot\exp\left[-\pi\left(\frac{y}{D}\right)^2\right] \tag{3}$$

normalized with respect to the two independent variable such that

$$\int |g_x(x)|^2\,dx = \int |g_y(y)|^2\,dy = 1 \tag{4}$$

It is not necessary that the two window functions $g_x(x)$ and $g_y(y)$ be identical (of equal spread) or in general even of the same type. It is however required that they both be normalized (of unit energy), and that the conditions of proper information cell size $W_x D_x \leq 2\pi$ and $W_y D_y \leq 2\pi$ be satisfied. Note that in case of equality (optimal cell size), the function set $\{f_{m_x n_x m_y n_y}\}$ becomes complete, with each of the functions being separable into $f_{m_x n_x} \cdot f_{m_y n_y}$ (Fig. 1).

To calculate the coefficient set $\{a_{m_x n_x m_y n_y}\}$ an auxiliary function must be employed since the elementary functions under consideration are not orthogonal. In the one-dimensional case Bastiaans determined the required bi-orthogonal function set, the so-called $\gamma(x)$ function [8]:

$$\gamma(x) = \left[\frac{1}{\sqrt{2}D}\right]^{1/2} \cdot \left[\frac{K_0}{\pi}\right]^{3/2} \cdot \exp\left[\pi\left[\frac{x}{D}\right]^2\right] \cdot \sum_{n+\frac{1}{2}\geq\frac{x}{D}} (-1)^n \cdot \exp\left[-\pi(n+\frac{1}{2})^2\right] \qquad (5)$$

where $K_0 = 1.8540746$ is a normalization factor (Fig. 2). Extending Bastiaans' work to two dimensions, we note that due to the separability of $g(x,y)$, and the duality of the $g(x,y)$ and $\gamma(x,y)$ functions, the latter is also separable into

$$\gamma(x,y) = \gamma_x(x) \cdot \gamma_y(y) \qquad (6)$$

This observation simplifies the the extension of the Gabor scheme into a two-dimensional (or higher dimensional) system. An alternative solution is afforded by the Zak transform [11]. Using the auxiliary function $\gamma(x,y)$, the coefficients $\{a_{m_x n_x m_y n_y}\}$ are calculated by:

$$a_{m_x n_x m_y n_y} = \int\int \Phi(x,y) \cdot \gamma^*(x - m_x D_x,\ y - m_y D_y) \cdot \exp(-in_x W_x x - in_y W_y y)\, dx dy \qquad (7)$$

To better model the visual system with its eccentricity-dependent features and for various technological applications, we represent the Gabor scheme also in polar coordinates (r,ϑ), where

$$r = \sqrt{x^2+y^2} \qquad x = r\cos\vartheta$$
$$\vartheta = \tan^{-1}\frac{y}{x} \qquad y = r\sin\vartheta \qquad (8)$$

An image $\Phi(r,\vartheta)$ may accordingly be expressed by:

$$\Phi(r,\vartheta) = \sum_{m_r}\sum_{n_r}\sum_{m_\vartheta}\sum_{n_\vartheta} a_{m_r n_r m_\vartheta n_\vartheta} \cdot g(r - m_r D_r, \vartheta - m_\vartheta D_\vartheta) \cdot \exp(in_r W_r r + in_\vartheta W_\vartheta \vartheta) \qquad (9)$$

with the coefficients being calculated similarly to the expression in the Cartesian coordinates:

$$a_{m_r n_r m_\vartheta n_\vartheta} = \int_{\vartheta=0}^{2\pi}\int_{r=0}^{r_{max}} \Phi(r\cos\vartheta, r\sin\vartheta) \cdot \gamma^*(r - m_r D_r, \vartheta - m_\vartheta D_\vartheta) \cdot \exp(-in_r W_r r - in_\vartheta W_\vartheta \vartheta)\, dr d\vartheta \qquad (10)$$

Although the image is given in Cartesian coordinates, the processing takes place in a polar coordinate system where position-dependent sampling can be easily incorporated along the r axis. The image is then encoded by cells representing the coefficients of the position $(m_r D_r, m_\vartheta D_\vartheta)$ and spatial frequencies $(n_r W_r, n_\vartheta W_\vartheta)$.

Physiological data indicate though a somewhat different representation, where cells are sensitive to a specific frequency , say f_p, and a specific orientations α_p instead of spatial frequencies $(n_r W_r, n_\vartheta W_\vartheta)$. The characteristics of such a two-dimensional scheme are currently under investigation.

III. Image representation

The basic features of the Gabor scheme for image representation were studied computationally. Examples such as demonstrated in Figure 3 illustrate that indeed one can decompose and reconstruct images using a finite set of Gabor elementary functions. Furthermore, the computations confirm and support various theoretical observations. In particular it becomes apparent that there exists a trade-off between the number of functions utilized along the spatial coordinates (m_x and m_y) and the number of frequency components employed per position. Thus, the dimensionality of the finite Gabor scheme determines the quality of image representation with a degree of freedom permitting various "paving" schemes. Similarly, the optimality achieved by the Gabor functions specifies the joint product uncertainty but permits the selection of either effective spatial spread or the effective frequency spread.

The finite expansion coefficient set provides a compact representation of an image. Graphically this can be better demonstrated by considering image cross-cuts and the corresponding coefficient distribution. Examples such as demonstrated in Figure 4 illustrate that a small number of coefficients can effectively represent an image. Given a certain dimensionality of the scheme, further compression can be achieved by thresholding the coefficients (quantization).

IV. The generalized Gabor scheme

To model the biological (human) visual system, the Gabor scheme has been generalized to incorporate effects of quantization, position dependent sampling rate and oversampling [10].

It is well known that phase information is most important in vision [12]. Using results obtained by Goodman and Silvestri [13], it is shown elsewhere [10] that for N quantization levels the one-dimensional signal reconstructed from a quantized (phase only, leaving the magnitude unchanged) set $\{a_{mn}\}$ can be expressed by:

$$\tilde{\Phi}(x) = \sum_m \sum_n |a_{mn}| \cdot f_{mn} \cdot \exp(i\tilde{\varphi}_{mn}) = \tag{11}$$

$$= \sum_m \sum_n |a_{mn}| \cdot f_{mn} \cdot \left\{ \sum_p sinc\left[p + \frac{1}{N}\right] \cdot \exp\left[i(pN+1)\varphi_{mn}\right] \right\}$$

These analytical results were computationally tested for 5 and 24 quantization levels. The results, (Fig. 5) indicate a good approximation for 24 levels, and even with 5 quantization levels the signal can be easily recognized.

Position-dependent sampling rate is incorporated through distortions of the independent position variable [10]. Such a scheme becomes attractive for various technological system in which a wide-field display system is implemented. Oversampling, characteristic of the biological visual system, improves image quality.

REFERENCES

[1] S.Marcelja, "Mathematical description of the responses of simple cortical cells", J.Opt.Soc.Am. Vol 70,no 11 pp. 1297-1300 (1980)

[2] D.M. Mackay, "Strife over visual cortical function", Nature, Vol. 289, pp. 117-118 (1981).

[3] Y.Y. Zeevi, J.G. Daugman, "Some Psychophysical aspects of visual processing of displayed information", Proc. Image II Conf., Phoenix, pp. 260-267 (1982).

[4] D.A. Pollen, S.F.Ronner, "Phase relationship between adjacent simple cells in the visual cortex", Science, Vol. 212, pp. 1409-1411 (1982).

[5] A.B. Watson, "Detection and recognition of simple spatial forms", NASA Technical Memo. 84353, Ames Res. Center, Moffet Field, CA., (1983).

[6] D.A.Pollen and S.F.Ronner, "Visual cortical neurons as localized spatial frequency filter", IEEE Trans. Sys. Man Cyb., Vol SMC-13,no.5,pp. 907-916 (1983).

[7] D.Gabor, "Theory of communication", J.I.E.E. Vol 93, pp. 429-459 (1946).

[8] M.J.Bastiaans, "A sampling theorem for the complex spectrogram and Gabor expansion of a signal into Gaussian elementary signals", Opt. Eng. Vol 20, no.4, pp. 594-598 (1981).

[9] Y.Y.Zeevi and M.Porat, "Combined frequency-position scheme of image representation in vision", J. Opt. Soc. Am. (A) Vol. 1, No. 12 pp. 1248 (1984)

[10] M. Porat Y.Y. Zeevi, "The generalized Gabor scheme of image representation in vision", EE Pub. No. 519, Technion, 1985.

[11] H.Bacry, A.Grossman,J.Zak,"Proof of completeness of lattice state in the kq representation", Phys. Rev. B12 (1975) pp. 1118-1120.

[12] A.V.Oppenheim and J.S.Lim, "The importance of phase in signals", Proc. IEEE 69, No. 5, pp. 529-541 (1981).

[13] J.W.Goodman and A.M.Silvestri, "Some effects of Fourier-domain phase quantization", IBM J.Res.Develop. no. 14 pp.478-484 (1970).

Image representation in vision using a generalized Gabor scheme

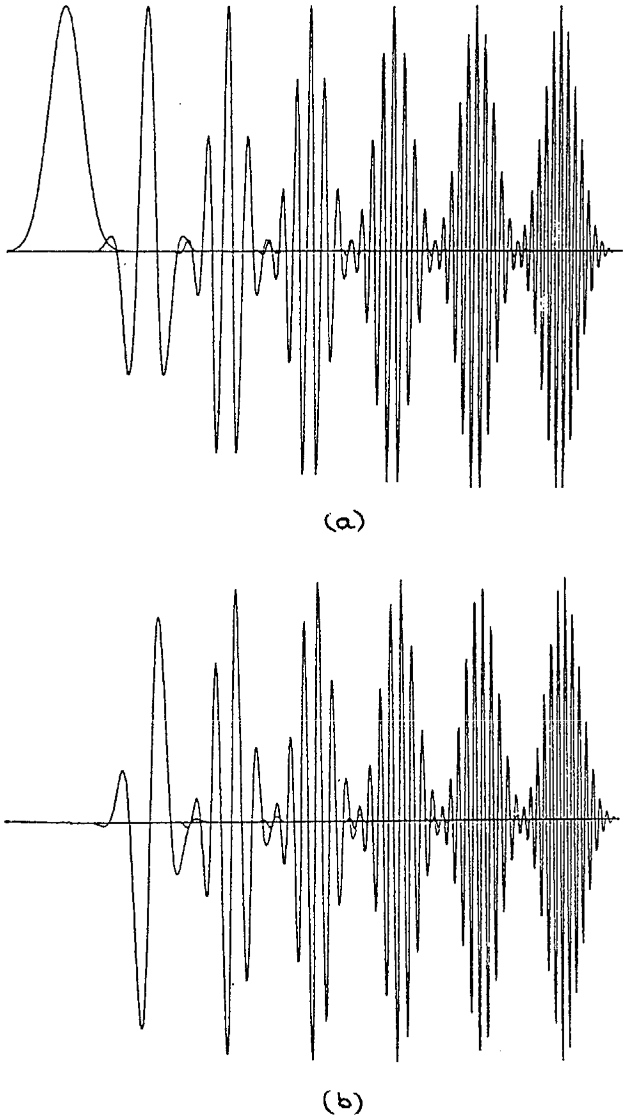

Fig. 1 - A set of the real part (a), and the imaginary part (b) of seven Gabor functions, characterized by the same spatial spread. $f_{mn} = g(x-2mD) \cdot exp(nWx)$.

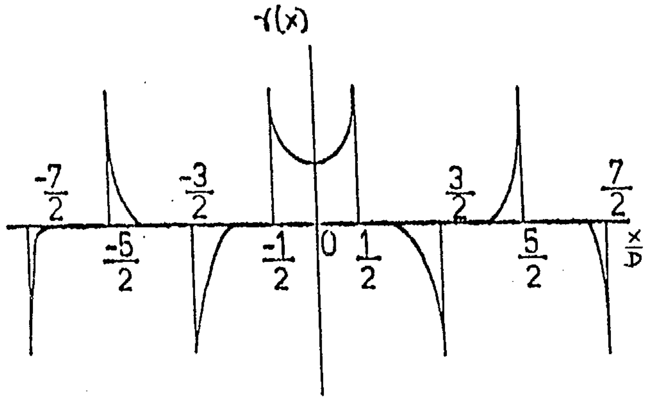

Fig. 2 - The auxiliary function $\gamma(x)$. This function is bi-orthogonal to the Gaussian window function $g(x)$.

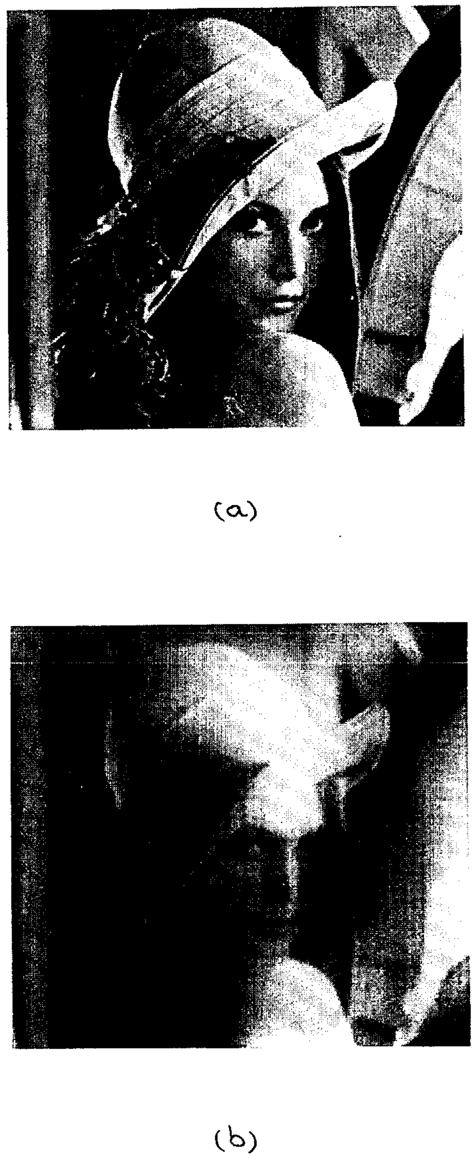

Fig. 3 - (a) Original (256x256 pixel) and (b) image reconstructed by a finite set of Gabor elementary function. This example illustrate the locality of information as captured by a confined set of frequency components.

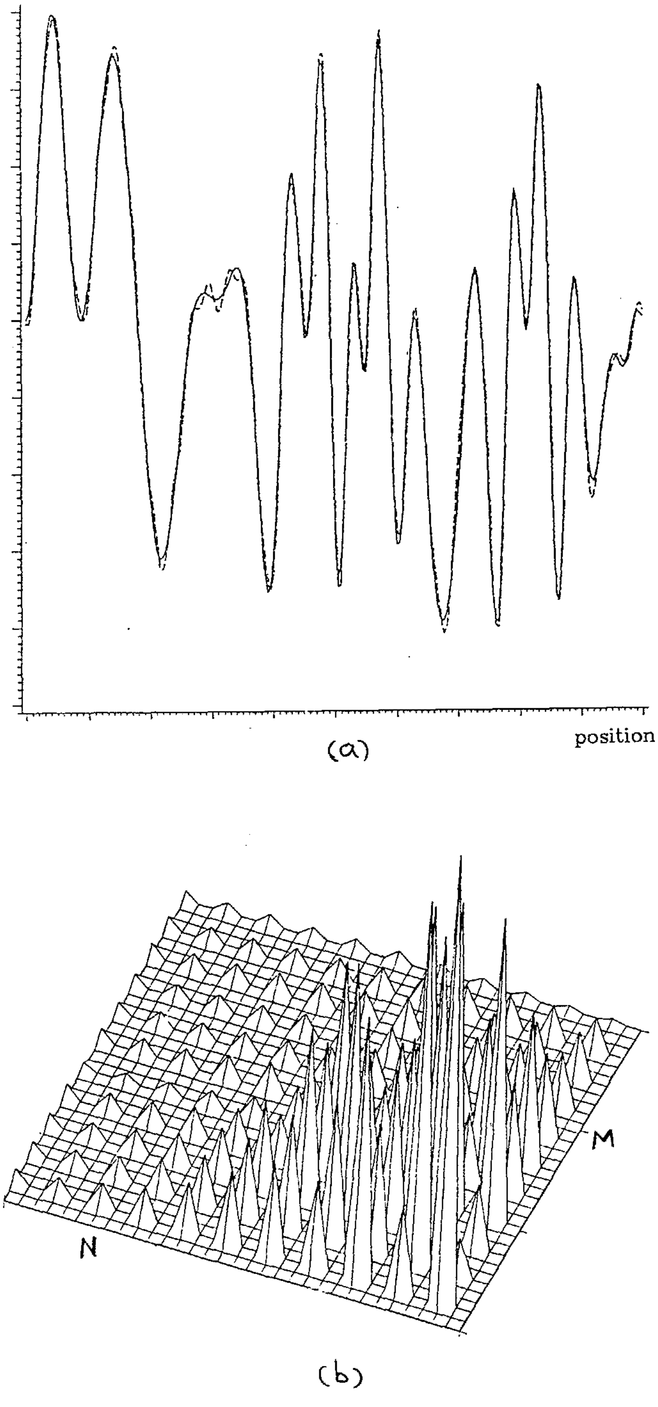

Fig. 4 - (a) Decomposition and reconstruction of a "typical" (band limited) image cross-cut. Superimposed are the original image (——), and the image reconstructed, with seven Gabor functions per position (-----).

(b) Absolute value of "Gabor cells" for the above aperiodic image cross-cut.

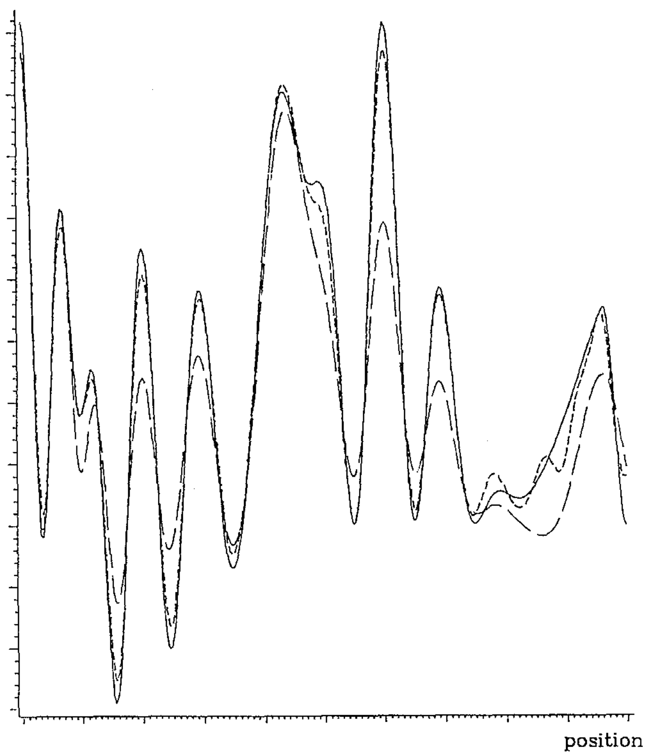

Fig. 5 - Quantization effect on image representation in a Gabor scheme. Compared are the original (——), and images reconstructed with phase step of 15 degrees (24 quantization levels ----) and phase step of 72 degrees (5 quantization levels - - -).