Speed, frequency, and orientation tuned 3-D Gabor filter banks and their design*

Josef Bigün
Signal Processing Laboratory
Swiss Federal Institute of Technology
Ecublens (DE) CH-1015 Lausanne Switzerland

Abstract

Design of 3-D filters for image sequences is addressed. A family of filter banks which is capable of mimmicking low level human motion perception i.e. being sensitive to low spatial frequencies moving fast and high spatial frequencies moving slowly, is proposed. The sensitivity is modelled quantitatively as existance of many filters with small band widths in a range of measurements. The speed, orientation, and spatial frequency tunings as well as the corresponding sensitivities are controllable in an explicit fashion. Implementation results are discussed.

1 Introduction

Design of linear 1-D and 2-D filters have been at the center of attention for many applications which involve pattern analysis and synthesis, [4, 13, 3]. Often, the success of a pattern recognition or a compression scheme depends crucially on whether the utilized filters effectively capture, for human, essential features of the signal or not. A functionality of the vision systems of primates, is that the organisms are capable of analysing the visual field locally, among others, by means of a vast number of linear filters, [1, 11, 9] which are tuned to given orientations of moving oriented line patterns such as moving edges or gratings. The hyper columns, which can be seen as toolboxes, of the receptive fields include cells whose optimal spatial frequencies decrease with increased eccentricities of the receptive fields, [10]. Although many of the cells in a given "toolbox" have an allpass character when responding to different speeds of the gratings, a very large population have a bandpass character, i.e. they are speed tuned. The optimal speed of such cells increase with increased eccentricity of their corresponding cells. Since the spatial frequency sensitivity of the band pass cells in any hyper column is inversely proportional to their corresponding receptive field eccentricity, [12], the linear cells in the visual cortices of primates are very sensitive to *i) gratings with high spatial frequencies moving slowly* and *ii) gratings with low spatial frequencies moving fast.*

We will propose a Gabor filter bank design technique in order to fulfill these criteria. It involves a uniform sampling of the Fourier spectrum after a non-linear coordinate transformation. The transformation can be seen as consisting of two parts, i) the transformation of the spatial frequency coordinates by means of the harmonic polar transformation, ii) a temporal mapping which is applied to the temporal, and the spatial frequency coordinates jointly. Frequency domain deformations yielding uniformly distributed orientation and log-frequency band pass filters have been used in 2-D pattern recognition, e.g. [13, 3], as well as coding, e.g. [4, 13]. For example, a common characteristics of the efficient image compression schemes is to avoid large quantization errors in the low frequency parts of the local or global spectra which is a direct consequence of non-uniform sensitivity of the human visual system to spatial frequencies.

Much of the earlier work which deal with 3-D local spatio-temporal analysis by linear filter banks concern local motion estimation, which is equivalent to the problem of plane and line fitting in 3-D, rather than characteristics which concern filter distribution and design strategies, e.g. [5, 2, 7, 8]. Although, Fleet and Jepson, [5], and Florack et. al., [6], proposed speed sensitive filters they have not discussed the filter distribution and the sensitivity issues. By contrast, we will propose a tesselation technique which will simplify the non-uniform sampling, [13], of the local spectra also with respect to the temporal direction.

*This project is supported by Thomson CSF (France).

2 Sensitivity and deformation

Consider the coordinate transformation (deformation),

$$x_1 = \exp(\xi_1)\cos(\xi_2) \quad (1)$$
$$x_2 = \exp(\xi_1)\sin(\xi_2) \quad (2)$$
$$x_3 = \exp(\xi_1)\tan(\theta(\xi_1,\xi_3)) \quad (3)$$

We will assume that x_j's correspond to the cartesian frequency coordinates while ξ_j's correspond to the new coordinates. Here x_3 is the temporal frequency coordinate. The choice of x_1 and x_2 to harmonic polar transformation has been made extensively in 2-D image analysis since it yields octave (Geometric progression) bandwidths and uniform orientation tunings independently, [3, 13, 4] and mimics the spatial characteristics of the human visual system. Therefore, and in order to simplify the problem, the harmonic polar mapping has been chosen here too. Since we wish that speed versus spatial frequency sensitivities be equally valid for all spatial orientations we have assumed that x_3 is independent of ξ_2. Furthermore, we note that the velocity, v, of a planar wave is $\tan\theta$, the behaviour of which with respect to the spatial frequency and temporal frequency coordinates ξ_1 and ξ_3 we wish to study. The angle θ controls the speed via the injective function $\tan\theta$. The functional determinant of the total maping is thus:

$$dV_{\mathbf{x}} = \exp(3\xi_1)(\tan^2(\theta)+1)\frac{\partial\theta}{\partial\xi_3}dV_{\xi} \quad (4)$$

Roughly, the desired sensitivity corresponds to cell sizes (filter supports) which are represented by Figure 1. The factor $\exp(3\xi_1)(\tan^2(\theta)+1)$ in (4), which represents the infinitesimal volume, does not behave according to Figure 1 i.e. it increases with increasing ρ and θ, along the lines (1) and (a). Thus, we must manipulate $\partial\theta/\partial\xi_3$ in order to correct for this deficiency. Since, $\exp(3\xi_1)(\tan^2(\theta)+1)$ is an increasing function in ξ_1, and θ, the overall behaviour of $\partial\theta/\partial\xi_3$ will be the overall behaviour of $dV_{\mathbf{x}}$. Heuristically, we choose the products of linear functions in order to model Figure 1.

$$\frac{\partial\theta}{\partial\xi_3} = [(1-p_1(\rho))(1-p_2(\theta)) + p_1(\rho)p_2(\theta)] + C \quad (5)$$

where $p_1 = (\rho-\rho_{min})/(\rho_{max}-\rho_{min})$ and $p_2 = \theta/\theta_{max}$ are linear functions of ρ and θ. They are chosen in such a way that both p_j's assume values only in $[0,1]$. For a small ρ, (or θ) the first term in (5), which is a decreasing function, is dominant, conversely

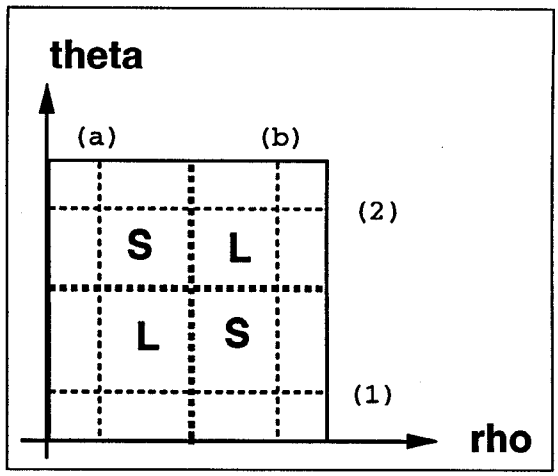

Figure 1: The distribution of the filter sizes. L = Large, and S = Small.

the second term is dominant when ρ (or θ) is large. Along a cross going through the central part of the figure, dV, practically, does not change. Without loss of generality we assume that this cross is at the center. For constant values of p, (5) is an ordinary differential equation, with the natural boundary condition $\theta(\xi_1,0)=0$ yielding:

$$\theta(\xi_3) = \begin{cases} \frac{1}{1-\alpha}[1-\alpha^{\xi_3}], & \text{if } p_1 \neq \frac{1}{2}; \\ \xi_3, & \text{if } p_1 = \frac{1}{2}. \end{cases} \quad (6)$$

where, $\alpha = (p_1+C)/(1-p_1+C)$.

Naturally, there are functions other than the heuristically chosen one in (5), but the merit of the current choice is that the resulting θ, is easily invertible with respect to ρ (and thereby with respect to ξ_1) and ξ_3 in terms of elementary functions. Consequently, the total transformation , (1-3,6), is invertible by means of elementary functions. The invertability is an atractive property, since we actually start from the $\mathbf{x}$ domain, that is we map Cartesian grid points, which constitute our frequency domain, to ξ domain in which the Gabor functions (filters) are defined as:

$$\exp(-(\xi-\xi^0(k,l,m))^t S^{-2}(\xi-\xi^0(k,l,m))/2) \quad (7)$$

with S being a diagonal non-singular matrix representing the standard deviation of the Gaussian along the coordinates $\xi = (\xi_1,\xi_2,\xi_3)^t$. Equation (7) represents Gaussian filters which only differ by their center frequencies $\xi^0(k,l,m)$ the arguments of which are indices in ξ_1,ξ_2,ξ_3 directions and are chosen as points on a uniform cubic grid. Thus, such complete filter sets constitute a Gabor filter bank.

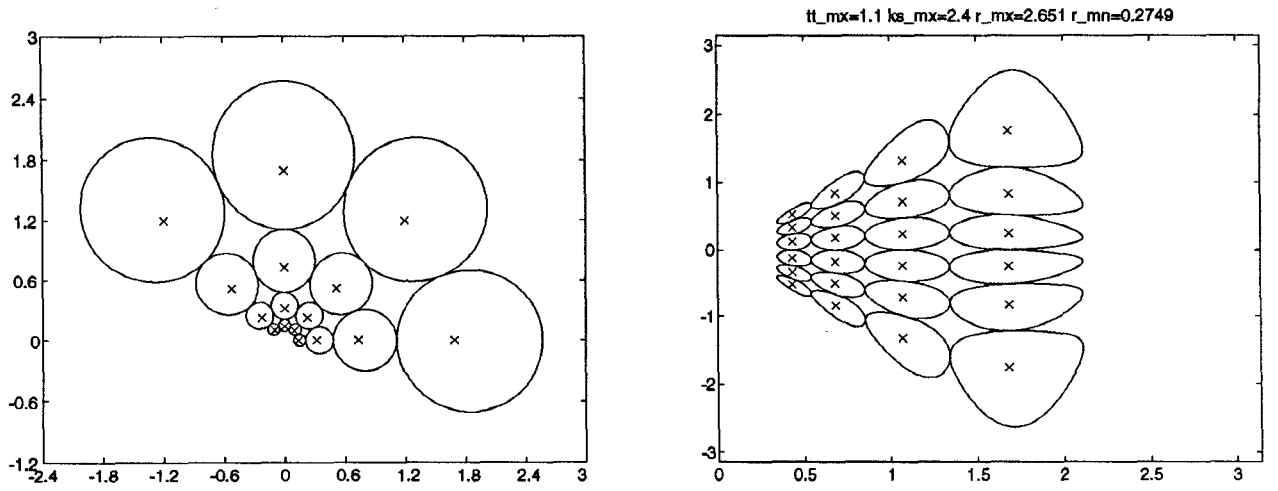

Figure 2: (Left) Spatial frequency distribution of the filters and their supports. (Right) Spatio temporal distribution of the filters. y axis is the temporal frequency axis.

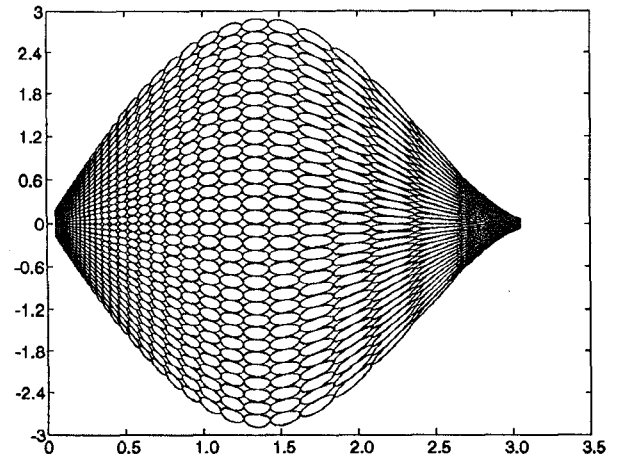

Figure 3: The spatio temporal distribution of filters when the number of filters is extremely large.

3 Implementation

The closed curves in our 2-D (and surfaces in 3-D) figures represent the iso level curves of the filters half-way between mesh points (in the ξ coordinates) visualised in the x_j coordinates. At these curves, the filter amplitudes have the value, 0.375. Consequently, at neighboring mesh points the values of the Gaussians decrease to 0.020. Figure 2 Left, displays the iso level curves of a filter bank cutting the x_1–x_2 plane in which the harmonic polar coordinates provide for center frequencies and band widths increasing geometrically in the radial direction while they provide for a uniform change, in the angular direction. The iso level curves are oval rather than perfect circles. High sensitivity to low spatial frequencies and low sensitivity to high frequencies are achieved by means of the corresponding filters have large and small band widths respectively. Similarly, we illustrate the spatio temporal dimensions of our filters in Figure 2 Right. We note that the highest spatial frequency filters have centers which have smaller elevation angles, θ, than the centers corresponding to low spatial frequency filters. Also the support of high freqency high speed filters cover a large sector of the elevation angles i.e. speed, compared to high spatial frequency low speed filters. This accounts for low sensitivity for low spatial frequencies with high speeds and vice versa. Here we observe an interesting property. In order to adjust for the speed bandwidth requirements at the same time as keeping the spatial bandwidths to increase in a geometric fashion, the filters are gradually rotated. This has the effect of increasing the speed bandwidths where it is desired, e.g. at low spatial frequency and high speeds the filter main direction is aligned with the corresponding filter coordinate vector while for high frequency high speeds it tends to be orthogonal to that direction. This is clearly visible when we increase the number of spatial frequency channels, see Figure 3. However, filtering with such a large bank is not very different than a DFT, since the frequency supports are very small. Finally, Figure 4 shows a 3-D view of a Gabor filter bank with 4 temporal channels.

References

[1] B. W. Andrews, D. A. Pollen *Relationship between spatial frequency selectivity and receptive field profile of simple cells* J. Physiol., (London), Nr. 287, p. 163-176, (1979)

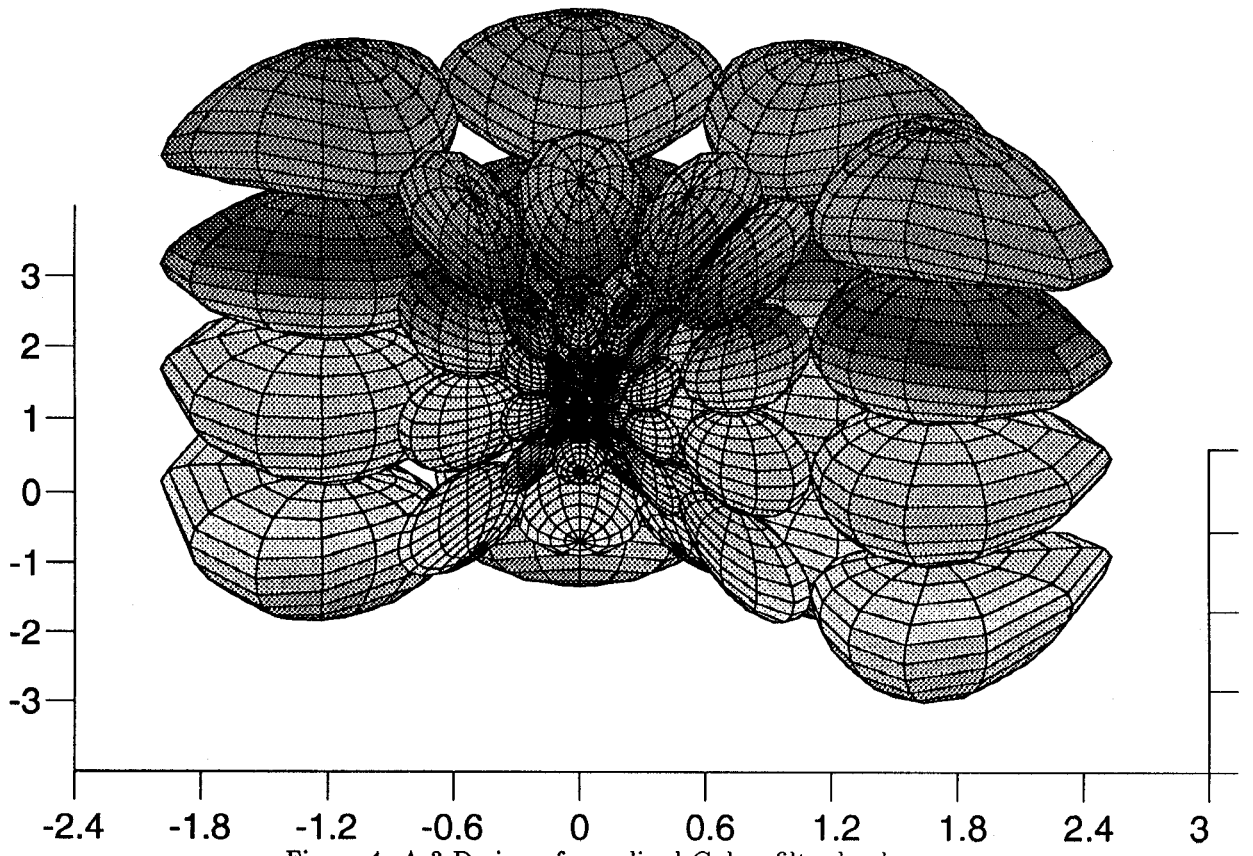

Figure 4: A 3-D view of a realized Gabor filter bank.

[2] J. Bigün, G. H. Granlund and J. Wiklund *Multidimensional orientation estimation with aplications to texture analysis and optical flow* IEEE-PAMI vol. 13, No. 8, p. 775-790 (1991).

[3] J. Bigün and J.M.H. du Buf *N-folded symmetries by complex moments in Gabor space.* IEEE-PAMI vol. 16, No.1, p. 80-87 (1993)

[4] J. G. Daugman *Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression* IEEE ASSP vol. 36, p. 1169-1179 (1988)

[5] D. J. Fleet and A. D. Jepson *Hierarchical construction of orientation and velocity selective filters* PAMI vol. 11 No. 3 (1989) p. 315-325

[6] L.M.J. Florack, B.M. ter H. Romeny, J.J. Koenderink and M.A. Viergever *Families of tuned scale-space kernels* ECCV92 p. 19-23 (1992)

[7] H. Haglund, H. Bårman, and H. Knutsson *Estimation of velocity and acceleration in time sequences* SCIA91, p. 1033-1041, (1991)

[8] D. J. Heeger *Optical flow using spatiotemporal filters* Int. J. Comp. Vision 1, p. 279-302, (1988)

[9] Kulikowski J. J. and P.O. Bishop *Fourier analysis and spatial representation in the visual cortex* Experientia, Nr. 37, p. 160-162, (1981)

[10] L. Maffei and A. Fiorentini *Spatial frequency rows in the striate visual cortex* Vision Res. Nr. 17, p. 257-264 (1977)

[11] J. A. Movshon, I. D. Thompson and D. J. Tolhurst *Spatial summation in the recptive fields of simple cells in the cat's striate cortex* J. Physiol. (London) vol. 283 p. 53-77 (1978)

[12] G. A. Orban, H. Kennedy, and H. Maes, *Response to movement of neurons in areas 17 and 18 of the cat: velocity sensitivity* J. Neurophysiol. No. 45, p. 1043-1058 (1981)

[13] M. Porat and Y. Y. Zeevi *The generalized Gabor scheme of image representation in biological and machine vision* PAMI, No. 4, p. 452-468 (1988)