DSP&CAES Conf., July 16-16, 1993 Nicosia Cyprus.
pp. 229-237

Gabor phase in boundary tracking and region segregation

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

Abstract

An eye movement mechanism with the task of tracking texture boundaries has been simulated. Gabor phase information has been dynamically used in order to track region boundaries. The tracking has been initiated manually from a point nearby a texture boundary with an initial boundary direction estimate. A filter selection strategy depending on the azimuth angles of the receptive fields, and a control mechanism relying on phase discontinuites which are computed through means of magnitudes and magnitude of means of the complex filter responses have been devised. The experiments suggest that letting the Gabor phase play a dynamic role in boundary tracking results in accurate tracking at pixel resolution even when the textures differ only in phase. As a by product tracking segregates textured regions.

1 Introduction

Description of image contents via local spectrum, e.g. Gabor spectrum, has been getting increased importance, [3, 2, 8, 4, 11, 7], in image analysis. However, most attention has been given to the magnitude of the spectrum (power spectrum) which has been extensively used in order to model repetitively occuring patterns, i.e. texture. Image regions which differ in their statistics of the local power spectra can be detected by using local power spectrum as it is or its various derivatives, [3].

We propose Gabor phase discontinuities in appropriately chosen frequency bands to define region boundaries. We will argue that with relevant orientation selection mechanisms the phase is able to define region boundaries more precisely than the power spectrum alone. This is particularly true when the neighboring regions have similar power spectra. We will present a control rule for an "eye", below called hypereye, following phase discontinuities, automatically. We will present experiments supporting the idea that the Gabor phase is sufficiently powerfull to track boundaries even when two neighborhood regions have identical power spectra, e.g. the upper two patches in Figure 3. This gives also new sense to texture since such regions are considered as identical by the human observer when they are seen independently. In contrast to this, when these regions are spatially adjacent, the texture boundary is eminent. Unless the local orientations of the boundary points, or equivalently which filters are the most significant at each boundary point, are known, the Gabor power spectrum can not be utilized in order to track the texture boundaries. But when the regions themselves consist of oriented textures, it is not evident how this orientation can be screened off from the local orientation of the boundary. In contrast to this, given suitable selection and tracking mechanisms, the Gabor phase discontinuites seem to cope with boundaries which pose serious difficulties to modelling with power spectrum alone. We will however not argue that the proposed hyper eye, the selection mechanism, and the control rule are ultimate solutions to the texture discrimination prob-

lem. They should merely be seen as demonstrators for the thesis that the Gabor phase information can play an active role in boundary detection and refinement in biological as well as machine vision.

2 The vision system and the filter selection

We assume that the feature extracting processors are directly attached to sub-regions i.e. *receptive fields,* of a camera which can be moved by means of a motor system. The camera-processor system will be called *hyper eye.* The spatial orientation angle of a receptive field will be called the *azimuth* of that field. The motion of the hyper eye causes different parts of a scene to be exposed at the receptive fields of the eye. The center of the eye will be referred to as the *focus of attention.* A feature extracting processor will sub serve only a given receptive field and will simulate the operational equivalent of a *hyper column, [1, 6, 9].* Each processor computes the filter responses of a full range of orientations. Figure 2 illustrates a layer of the receptive fieslds corresponding to the same frequency. Different frequency layers exist and overlap each other. In Figure 1, for the sake of illustration, these layers are taken apart vertically. In the biological systems, to each receptive field, a full range of orientation tuned cells corresponds. These cells are sensitive to a a range of frequencies, [5, 10, 9]. However, there is very little known about the feedback mechanisms of the wirings of the neurons in the biological visual systems. In this study we propose a wiring in order to cope with a very specific vision task: texture boundary tracking. Here the azimuth of a receptive field will determine the selected filter orientation, as will be motivated below.

Suppose that the input texture is approximated to a reasonable degree of accuracy by P sinusoids

$$f = \sum_{j=1}^{P} A_j \cos(r^t \omega_j + c_j) \qquad (1)$$

with ω_j corresponding to the wave vector determined by the frequency and orientation tuning of the Gabor filter h_j. This filter can be assumed to have the impulse response

$$h_j = g_j(r) \exp(i\omega_j^t r) \qquad (2)$$

with g_j being a Gaussian like function attaining its maximum at ω_j. Thus the response of h_m to f is

$$x_m = h_m * f = \sum_j \alpha_j g_m(\omega_j) \exp(ir^t \omega_j) \qquad (3)$$

where $\alpha_j = A_j \exp(ic_j)$, is complex valued. When $|x_m| >> 0$, the $\arg(x_m)$ is well defined and a support for an orientation and frequency corresponding to ω_m in the receptive field exists. Since we are dealing with finite amplitude inputs and the filter responses to frequencies other than ω_m are small the inequality

$$|x_m - \alpha_m g_m(\omega_m) \exp(ir^t \omega_m)| \leq \sum_{j \neq m} A_j g_m(\omega_j) = M\epsilon \qquad (4)$$

with ϵ being small, holds. The direction of the maximum change of $\arg(x_m)$ can thus be written as

$$\nabla(\arg x_m) = i\omega_m \qquad (5)$$

or equivalently, whenever $x_m >> 0$, the derivative of $\arg x_m$ along the direction orthogonal to ω_m is small. The phase variation along the mentioned direction is zero when there is one orientation in the receptive field.

To detect phase variations i.e. discontinuities we propose to use the triangle inequality

$$|\sum_i z_i| \leq \sum_i |z_i| \qquad (6)$$

where $\{z_i\}$ are a finite set of complex numbers. Equality holds if and only if all z_i have the same argument (co-linear). The larger the difference between the left hand and the right hand, the larger the angular variance is within the set $\{z_i\}$. The angular variance is zero when the statement above holds with equality. Dividing both sides with the number of elements in the set yields:

$$|\bar{z}_i| \leq \overline{|z_i|} \qquad (7)$$

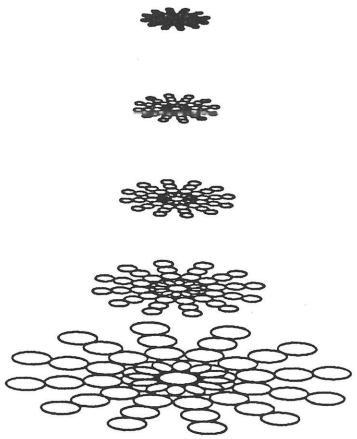

Figure 1: The receptive fields corresponding to different resolutions. The pyramidal decrease in size illustrates the size and position changes of the receptive fields.

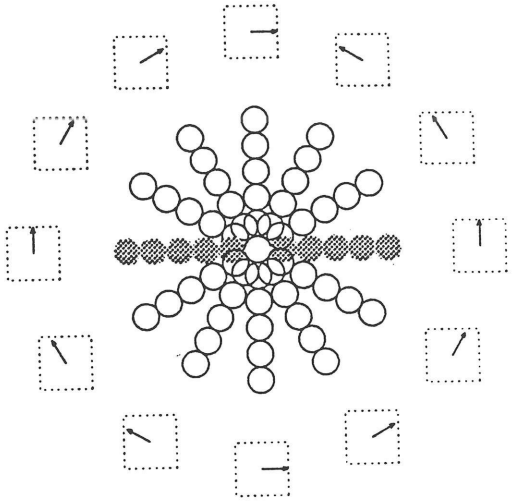

Figure 2: The configuration of the receptive fields corresponding to a particular frequency. The shaded fields represent a radius to which averaging is applied. The arrows indicate the frequency domain orientation of the selected Gabor filters for a given spatial orientation.

where the bar represents the mean. Thus the magnitude of the means and the means of the magnitudes can be utilized to detect phase discontinuities in the set $\{z_i\}$.

From the previous argumentation it is clear that every component of the Gabor spectrum has a spatial direction and a spatial distance in which the phase variations can be expected to be small. We will use this observation in order to choose subsets of the Gabor spectrum for the purpose of phase discontinuity detection. This will avoid the computations of averaging in all directions and frequencies.

We assume that in our hypereye we have receptive fields with corresponding hypercolumns specialized to different frequencies and orientations (Gabor spectrum). We will only use one orientation out of all possible in a particular azimuth. The vectors in Figure 2 illustrate the wave vectors of the selected filters for each azimuth. The Receptive field configuration for all frequencies, i.e. resolutions, is illustrated by the pyramid in Figure 1. At each receptive field in this pyramid we select a filter response, i.e. the response of the filter with orientation tuning which conforms with the azimuth angle of the receptive field and with frequency tuning corresponding to the resolution level of the field. For a given resolution, k, and orientation, l, we compute the magnitude of the mean

$$w_{kl} = |\bar{z}_i| \tag{8}$$

along a *diameter* , e.g. the fields with the same shading in Figure 2. We note that l represents a diameter as well as a filter orientation due to the selection mechanism. We repeat this for all resolutions and directions yielding a matrix w.

Similarly, the mean of magnitudes along a diameter yields a matrix c

$$c_{kl} = \overline{|z_i|} \tag{9}$$

The matrices w, and c are attributed to the the focus of attention. They have the same size as the set of the original gabor filter responses. However, these matrices are real, non-negative, and represent the local phase variations of the Gabor filter responses along the diameters. We define the norm of a matrix as

$$\|u\| = (\sum_{kl} u_{kl}^2)^{1/2} \tag{10}$$

Let r represent the focus of attention which is assumed to be a boundary point. Let $w(r)$ and $c(r)$ be computed as before. Then

$$\sigma_B(r) = \|c(r) - w(r)\| \tag{11}$$

represents the between phase-variation at the boundary point. The larger σ_B is, the higher the certainty of r being a phase discontinuity point.

3 The control rule and experimental results

Apart from a phase discontinuity measure, tracking a boundary with a hypereye requires a control rule in order to predict the next point of the boundary. However, before we mention our control strategy, in order to confirm the usefullness of the different matrices proposed, we will list some of them for a point in the middle of the upper left texture patch, and a point at the boundary of the two upper patches of Figure 3. The input image had the size of 128 x 128. In the Gabor decomposition, we had 5 octave frequencies and 6 orientations. The columns of the matrices below, represent $\frac{\pi}{2}$, $\frac{\pi}{6}+\frac{\pi}{2}$, $2\frac{\pi}{6}+\frac{\pi}{2}$, π, $4\frac{\pi}{6}+\frac{\pi}{2}$, and $5\frac{\pi}{6}+\frac{\pi}{2}$ azimuth angles. The rows represent resolutions, i.e. the first row represents the lowest resolution while the last row represents the highest resolution. Each radius of the hyper eye contained 5 *equidistant*receptive fields. The distance between the receptive fields at the lowest resolution (first rows of the matrices) was approximately 7 pixels which succesively decreased to 1 pixel at the highest resolution (last rows of the matrices) in an approximately octave fashion. The current direction was vertical.

The w and c matrices for a point at the center of the upper left patch, are

w =

1	1	3	3	3	4
2	2	3	79	13	3
0	0	3	19	2	1
0	0	0	0	0	0
0	0	0	0	0	0

c =

2	1	3	3	3	5
3	3	8	79	15	5
0	0	3	19	2	2
0	0	0	0	0	0
0	0	0	0	0	0

We can see that for the largest elements, e.g. the direction π, $c_{kl}-w_{kl}$ are very small suggesting that the current point is not a boundary point. The corresponding matrices computed at a boundary point in the middle of the upper part of the image are:

w =

2	1	4	2	2	6
3	4	5	7	20	5
1	1	31	10	4	14
15	18	5	2	2	3
7	25	4	5	6	4

c =

3	2	4	2	2	6
4	4	17	60	23	11
8	9	32	16	3	15
15	25	8	3	2	4
8	29	7	5	7	5

We observe that for the largest elements of c, $c_{kl}-w_{kl}$ is very large while along a direction close to the boundary tangent and at higher frequencies the corresponding differences, $c_{kl}-v_{kl}$, are very small, e.g. rows 3,4,5.

The control strategy consists of computing a direction $d^{(2)}$ which is used in order to move the eye to the next boundary point. The magnitude of the displacement, $\|d^{(2)}\|$=D, is chosen a priori and is constant. If the previous displacement is $d^{(1)}$ the next displacement as well as the next boundary positions are chosen as

$$d^{(2)} = \max_j \sigma_B(r+d_j) \qquad (12)$$

$$r^{(2)} = r^{(1)} + d^{(2)} \qquad (13)$$

where $\{d_i\}$ is a set of displacements corresponding to the vectors represented by the complex numbers

$$d_j = D\exp(i(j\frac{\pi}{2N_h} + \arg(d^{(1)}))) \qquad (14)$$

with j being integer and $j \in [-\frac{N_h-1}{2}, \frac{N_h-1}{2}]$. These displacements correspond to N_h equidistant angles ahead of the current direction so that $d_0 = d^{(1)}$. In the experiments we used $N_h = 3$ since it gave more robust tracking than larger values. This behaviour may be explained by the nonlinearity in the control rule. Another experimental finding was about the frequencies which constitute the rows of the matrices c and w. When we let all frequencies enter the computations the tracking was less stable compared to only including the highest two frequencies. This can be explained by the high frequency nature of the considered boundaries. The experiments indicate that it is possible to determine the limiting frequency automatically, i.e. the highest frequency at which there is a significant phase discontinuity and a significant share of the total energy. Consequently, only the highest two resolutions, e.g. the last two rows of the matrices displayed above, have taken part in σ_B computations.

We have chosen the first boundary point as well as its tangent direction manually and let $D = 2$ pixels. The result of such trackings are illustrated by Figure 4. The starting points are labeled with 1, 2, and 3. The corresponding starting directions were $\frac{\pi}{2}$, $\frac{\pi}{2}$, and π. The white points represent the traces of the focus of attention. The trace 1, had wrong initial direction but correct starting point while trace 3 had the oposite situation. In both cases the rapid convergence to the texture boundaries show the stability with respect to bad initial condition estimations. Also note the boundary accuracy at the regions with labels A and B where local intensity variations vanish. At the instant when all 3 trackings were near the central junction point of the four patches, the hypereye continued upwards. It can be noted that, except at the junction region the traces are very accurate and confirmative, i.e. the traces at the upper part coincide despite their different origins. In the experiments we have changed the initial conditions. This has resulted in that the eye continued from left to right or from down to right, ...etc., but has always followed the correct boundaries. The minor non-coincidence at the junction can be explained by the fact that there are more than one direction along which there are strong phase discontinuities. Since our simple control rule does not use the texture parameters, i.e. the Gabor power spectrum, on one side of the boundary, the tracking can not guarantee closed boundaries in a single tracking. With texture characteristics on one side of the boundary at hand, it is evident that closed boundaries can be guaranteed even in a single tracking.

The trackings of the hypereye are interrupted close to the image boundaries because otherwise some receptive fields of it would be outside of the image which here simulates the scene. Of course, in such a situation, in the biological systems the *head* would move in order to bring the scene into the desired part of the visual field.

4 Conclusions

The results we have obtained on a textured scene indicate that the the Gabor phase can account for boundary detections and accurate boundary estimations. This approach allows for a texture boundary notion even within the same texture. Considering, that the Gabor phase has not been given a clear role in texture discrimination tasks in neither biological nor machine vision systems, the hypereye approach in this study can be seen as a demonstrator as well as an indicator showing possible gains when the Gabor phase is actively exploited. Thanks to the fact that only a limited number of receptive fields around the focus of attention are utilized, and all computations are parallel and distributable to fixed regions of the visual scene, the proposed mechanism or its alikes can be made responsible for the movement of a head-eye system following texture or object boundaries.

Acknowledgement

This research has been supported by Thomson CSF.

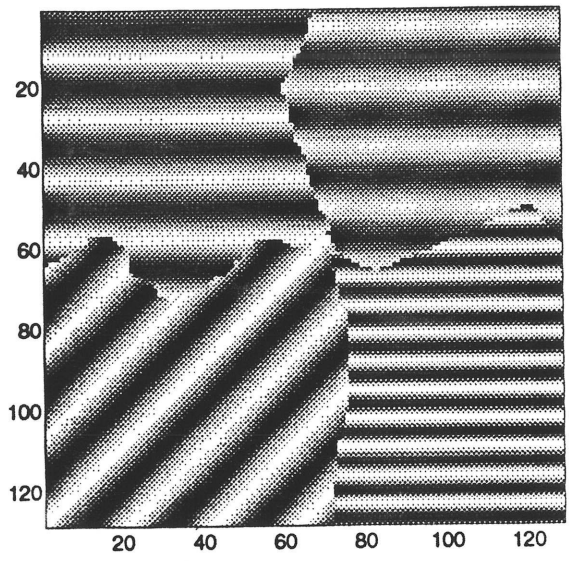

Figure 3: The original image which represents the scene on which the hyper eye is guided.

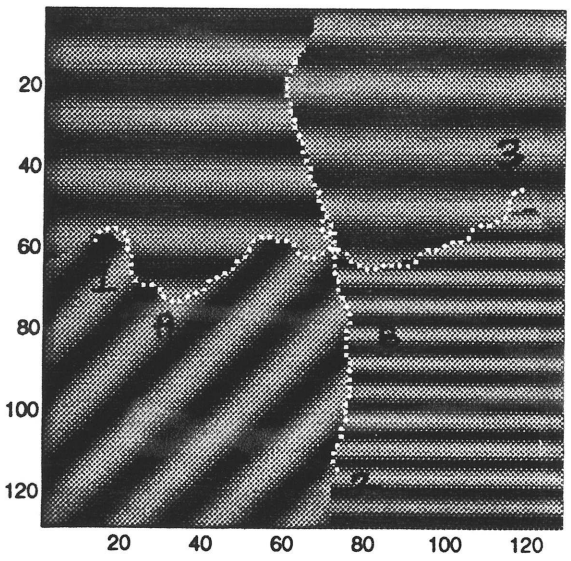

Figure 4: The traces of 3 boundary trackings using Gabor phase discontinuities.

References

[1] K. Albus *A quantitative study of the projection area of the central and the paracentral visul field in area 17 of the cat. II. The spatial organization of the orientation domain.* Exp Brain Res, Vol. 24, pp. 181-202, (1975)

[2] J. Bigün *Frequency and orientation sensitive texture measures using linear symmetry* Signal Processing, vol. 29, pp. 1-16, (1992)

[3] J. Bigün and J.M.H. du Buf *N-folded symmetries by complex moments in Gabor space.* In press IEEE-PAMI (1993)

[4] A. C. Bovik, M. Clark, and W. S. Geisler *Multichannel texture analysis using localized spatial filters* IEEE-PAMI vol. 12, No. 1, pp. 55-73, (1990)

[5] D.H. Hubel and T.N. Wiesel *Receptive fields of single neurones in the cat's striate cortex* J. physiology (London), No. 148, pp. 574-591 (1959)

[6] A. K. Jain, and F. Farrokhnia *Unsupervised texture segmentation using Gabor filters* Pattern Recognition, Vol. 24, No. 12, pp. 1167-1186, (1991)

[7] M. Andersson *Controllable multidimensional filters and models in low level computer vision* PhD thesis no. 282, Linköping University, (1992).

[8] G. Orban *Neuronal operations in the visual cortex* Springer, Berlin, (1984)

[9] D. A. Pollen, and S. F. Ronner *Phase relationships between adjacent simple cells in the visual cortex* Science, Vol. 212, pp. 1409-1411, (1981)

[10] M. Porat and Y. Y. Zeevi *Localized texture processing in vision: analysis and synthesis in the Gaborian space* IEEE Bio Medical Eng. BME-36 pp. 115-129 (1989)