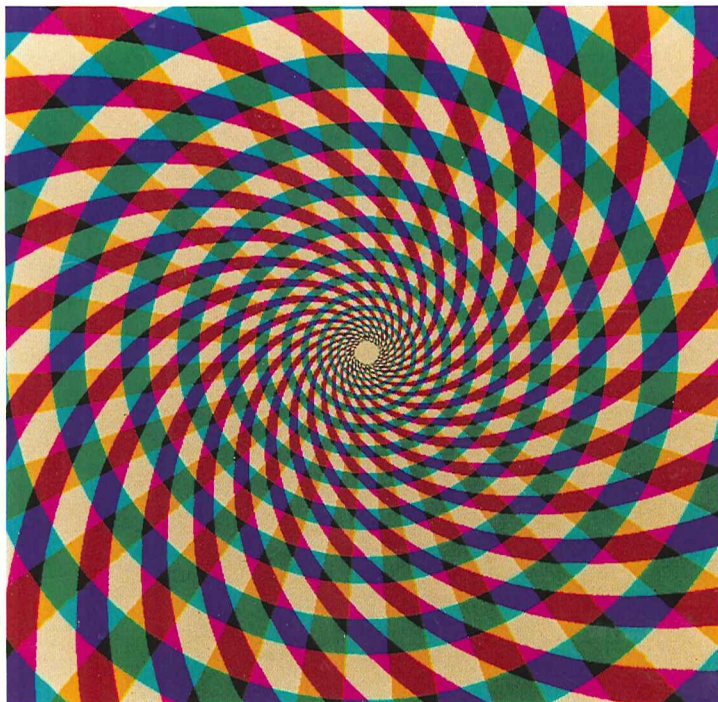


Linköping Studies in Science and Technology. Dissertations
No. 179

Local Symmetry Features in Image Processing

Josef Bigün



Computer Vision Laboratory
Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden

Linköping 1988

Linköping Studies in Science and Technology. Dissertations
No. 179

Local Symmetry Features in Image Processing

Josef Bigün



Computer Vision Laboratory
Department of Electrical Engineering
Linköping University, S-581 83 Linköping, Sweden

Linköping 1988

ISBN 91-7870-334-4
ISSN 0345-7524

Printed in Sweden by VTI-Grafiska, Vimmerby 1988

To my grandfathers:
Haçik Bigün
and
Giragos Göçerian

ERRATA

May 6, 1988

1. p. 11 after equation (7) *stands* (5). *It should be* (4).
2. p. 14 after equation (19) *stands* [8]. *It should be* [7].
3. p. 36 after equation (9) *stands* [12]. *It should be* [1].
4. p. 45 first paragraph *stands* (16). *It should be* (18).
5. p. 49 *stands* \bar{u}_{n1} . *It should be* \bar{u}_{n-1} .
6. p. 54 after equation (10) *stands* (4) and (5). *It should be* (6) and (7).
7. p. 70 in equation (31) *stands* $(J_{22} - J_{11} + \dots)$. *It should be* $(-J_{22} + J_{11} + \dots)$.
8. p. 70 after equation (32) *stands* $(J_{22} - J_{11} + \dots)$. *It should be* $(-J_{22} + J_{11} + \dots)$.
9. p. 77 in equation (14) *stands* n_d^2 . *It should be* ζ_d^2 .
10. p. 79 in equation (22) *stands* n_d^2 . *It should be* ζ_d^2 .
11. p. 80 after equation (27) *stands* [2]. *It should be* [1].
12. p. 81 after equation (31) *stands* [2]. *It should be* [4].
13. p. 82 before equation (32) *stands* Figure 5. *It should be* Figure 6.

CONTENTS

I	INTRODUCTION AND PRELIMINARIES	3
1	Introduction	4
2	Summary	4
3	General remarks	5
II	DETECTION OF CIRCULAR SYMMETRY PARAMETERS	7
1	Introduction	8
2	Description of the image by circularly symmetric isogray values	10
3	Evaluation of the energy dependent certainty	15
4	Evaluation of the energy independent certainty	18
5	Evaluation of the partial derivative image	22
6	Experimental results	22
	6.1 Noise behaviour tests	23
	6.2 Application examples	25
7	Conclusion	26
A	Some properties of the filter coefficients	34
B	Central symmetry modelling	37
III	OPTIMAL ORIENTATION DETECTION OF LINEAR SYMMETRY	47
1	Introduction	48
2	Orientation detection in n -dimensional euclidean space.	49
3	2-D implementation of finding the minimum variance axis	53
4	Experimental results	57
5	Conclusion	59
A	Overall results	63
B	The relation between the eigenvectors and the complex z plane.	70
IV	PATTERN RECOGNITION BY DETECTION OF LOCAL SYMMETRIES	71
1	Introduction	72
2	Modeling the local neighbourhoods by harmonic functions	73
3	Detection of local symmetries	79
4	Applications and experiments	80
5	Conclusion	85

V	OPTICAL FLOW BASED ON THE INERTIA MATRIX OF THE FREQUENCY DOMAIN	89
1	Introduction	90
2	Formulation of the problem	91
3	Experimental results	95
4	Conclusion	101

Chapter I

INTRODUCTION AND PRELIMINARIES

That is why a sensible man never ventures to trust his thoughts to these inadequate linguistic means, particularly not to so unchangable a form as that of written letters.

Plato

INTRODUCTION AND PRELIMINARIES

1 Introduction

The term feature extraction is widely used in computer vision. It is often used in the sense of a neighbourhood dependent mapping of a picture to a function of feature extraction. The result is a transformed picture, in which every point is associated with a point in the original picture which we will occasionally refer to as an *examined point*. The extraction of features is necessary for all aspects of processing and analysis such as classification, segmentation, enhancement and coding.

To decide whether a neighbourhood consists of a line has been subject to extensive research during the last decades. Much physiological and psychophysical evidence has indicated the importance of these structures and contributed to the extent of the research efforts. In the course of developing models to describe images, a need arises for description of more complex structures than lines. This need does not reject the importance of line structures but indicates the need to complete and utilize them in a more systematic way.

2 Summary

The concept common to all chapters is as follows. The problem of detecting certain structures is formulated in terms of a specific local coordinate transformation corresponding to these structures. Then the problem is reduced to be a line/edge detection problem, or rather the orientation of these together with the “strength” of the edges in the new coordinates. A closed form solution is proposed which is based on the principal axis analysis of the special frequency domain associated with the new coordinates. Utilizing this solution, compact computer algorithms are constructed evaluating “strength”, or rather certainties as they will be referred to, of line/edge structures together with orientation estimates. The evaluated orientation estimates are shown to be optimal in the least square sense, and the certainties have well defined meaning in terms of the least square error.

In the following chapters we will present some new methods for extraction of local symmetry features as well as experimental results and applications. It is widely believed that observation of symmetries in images is an important part of human vision. Local symmetries in images in different scales can of course originate from large as well as small neighbourhoods depending on the scale. Originally the motivation to study structures

between the obtained error and the error we would obtain if the translation was carried out in a direction perpendicular to the optimal direction. Continuing this analogy in the circular symmetry case we mention that the iso gray value curves proposed are the invariants of the infinitesimal operators corresponding to rotation and scaling with respect to a fixed point. The obtained orientation corresponds to the proportion of rotation versus scaling which changes the local image minimally. This analogy can easily be extended to other general symmetries described by iso gray value curves of the harmonic functions. To every such curve there is a corresponding infinitesimal operator which performs translation in a specific direction in the parameter space describing the curves.

The question whether these proposed symmetry structures play a critical role in human vision is more difficult to answer, not to say impossible, with our current knowledge of the mechanisms of mammal vision. We will confine ourselves to saying that the human perception of the world is perfectly able to judge its visual environment and compensate for the ego-motion, the motion of the world, as well as the confusion stereo vision might cause. To these compensation processes some symmetric patterns obviously fit well. The significance of one of these symmetries, linear symmetry, which is intimately connected to infinitesimal translations, is confirmed by neurophysiological experiments. Similarly there is support for the existence of areas in the central neural system where rotation and scaling compensation takes place, indicating the importance of circular symmetry analysis. However, we can only testify to the skill and perfection of our visual perception in innumerable difficult situations, admire it and admit that our knowledge of its mechanisms is incomplete.

Acknowledgements

I would like to express my deepest gratitude to those who have contributed to this report in various ways:

- To my supervisor, Gösta Granlund, Professor of Computer Vision at LiTH. He has been a constant source of inspiration during this work.
- To the members of the Computer Vision Laboratory at LiTH, especially to Leif Haglund, for reading the manuscripts. They have contributed with important comments and ideas for improvements.
- To our patient secretary Catharina Holmgren. She has typed parts of the manuscript and worked on the language.
- Not least to my family, especially to my sister Helen who has reminded me of the importance of housework and in fact very often did the chores during the course of this work when I lacked the time. They have proved to be indispensable with their sincere encouragement.

other than lines arose from attempts to decide whether a local image was "circular", "like a star" or "like a set of parallel lines". Such questions, and answers to these, are important when considering finger print images, X-ray images of cancer tumours, cross junctions and buildings in aerial images; on the whole, in many object/event descriptions in image processing.

To be able to give a concrete formulation to the rather vague concept of symmetry, we have utilized the iso-value curves of images. Iso gray value curves, as they will also be called, can be imagined to be the curves joining the points of a landscape map having the same altitude representing a certain gray value intensity. These curves originating from the local images are then examined with respect to how much they fit to an a priori chosen symmetry model.

- In Chapter II, circular symmetries, which originally motivated the study of other structures than lines, are considered.
- In Chapter III it is shown that the approach embraces line structures, which are also referred to as linear symmetries. Moreover, in Chapter III, a method extending the linear symmetry concept to finite, arbitrary dimensions is developed.
- Having shown the applicability of the symmetry approach to circular structures and linear structures in the previous chapters we proceed in Chapter IV to derive the fundamental conditions which extend the symmetry approach to much larger classes than circular and linear symmetries using the same unified approach. Examples are given together with experimental results.
- In Chapter V the optical flow problem is formulated in terms of linear symmetries in three dimensions and the results of the standard solution given in Chapter III are presented.

3 General remarks

One can ask what the obtained estimates of orientations based on the frequency domain of different coordinate systems refer to? Are these estimates relevant for vision?

From the following chapters it is evident that the iso gray value curves corresponding to different symmetries are invariants of infinitesimal operators with two parameters. Such an operator is possible to utilize for performing translation of a class of functions in the parameter space of iso gray value curves. To make things less complicated, we consider the well-known linearly symmetric structures. The obtained orientation for a local two dimensional image is the direction along which the neighbourhood can be translated with minimal error (in the least square error sense). In the case of an edge this error is zero for a translation parallel to the direction of the edge. That is, the local image is invariant to a certain infinitesimal operator performing a specific translation in the usual Cartesian coordinates. In the case when a neighbourhood degrades from an edge the obtained orientation corresponds to the direction along which a translation changes the local image minimally. One of the proposed certainties is the difference

Chapter II

DETECTION OF CIRCULAR SYMMETRY PARAMETERS

Don't move my circles!

Archimedes

DETECTION OF CIRCULAR SYMMETRY PARAMETERS

Josef Bigün
Computer Vision Laboratory
Linköping University
Department of Electrical Engineering
S-581 83 Linköping Sweden

Abstract

A model for circular symmetry is used to describe a local neighbourhood. A definition of circular symmetry is given which implies detection of one-dimensionality of a 2-D image after a coordinate transformation. The coordinate transformation is such that Archimedes' spirals map to straight lines. The Fourier transform of a circularly symmetric image, in these coordinates provides an energy concentration to a line in a certain direction. Local neighbourhoods consisting of one circle or several concentric circles show a concentration of energy to a line. This is also the case for lines with a common intersection point. These two types of circularly symmetric images map to two orthogonal lines in this special Fourier domain. Archimedes' spirals map continuously to lines with directions between these two orthogonal lines incorporating circles, half lines and spirals into the same model. Fitting a line in the least square sense in this special Fourier transform domain is shown to be possible to accomplish in the spatial domain as a convolution carried out on the partial derivative image. The necessary filters are derived. Two algorithms based on interpretation of the error of the fitted optimal line and its orientation are implemented. One is dependent on the energy of the variation of the local image, the other is not. Both use the same optimal estimate of the orientation of the fitted line. Experiments are carried out utilizing the implemented algorithms showing very good detection properties for spirals, circles, concentric circles, line ends and intersection point of a set of lines.

1 Introduction

There has been an increasing need to detect neighbourhoods with complex circular structures in computer vision. Extensive work has been done in low level vision, analysing statistical properties of neighbourhoods. Also structural properties of neighbourhoods, mainly in connection with edge and line detection, have been the subject of study for a long time. Patterns other than lines and edges have been modelled and detected in a higher level by either matching techniques, [1], or the generalized Hough transform, [6].

In the latter technique and many of the matching techniques, every pixel in the image is considered to be either an edge pixel or not. Based on the spatial distribution of the edge pixels, a decision process takes over their labelling. In the generalized Hough transform this decision process is voting for an edge pixel to be an edge of the modelled pattern or not. The voting process takes place in the parameter space and is followed by a search for peaks. The amount of computation is extensive but the positions of the patterns are well determined. The purpose of this paper is to present a new approach to model local circular structures in low level computer vision and give an algorithm estimating the model parameters. However, if the size of the images is reduced, the obtained parameters can be used as features to detect objects of considerable size. The motivation for this study of circular structures in image neighbourhoods originally arose in a project processing fingerprint images. The problem was to identify the few special points around which line structures in these images rotated or in which the lines ended. A similar need was observed for X-ray images of cancer tumours and cross junctions of the roads in aerial images.

The computation evaluating the proposed model parameters is carried out in two steps, each requiring only convolutions and pixelwise arithmetic operations. One of the model parameters is an angle and represents a subset of the circularly symmetric function family. By circular structures we do not only mean circles but rather a family of circles, spirals, and fan shapes, figures 6 and 7. In the next section we will more strictly define the family of patterns which is the basis of the approach. For example, if the neighbourhood consists of concentric circles then the angle or the orientation of this class is given by the angle 0 radian, while the fan shaped neighbourhoods are represented by π radians. The orientation parameter is continuous and corresponds to a subset of the circularly symmetric function family.

It is possible to represent any neighbourhood by means of the members of the circularly symmetric function family, just as any neighbourhood can be represented by means of sines and cosines. For this reason one can talk about the best fit of a subset of this function family to the local neighbourhood in the least square error sense. Two certainty measures for the least square estimate are given. The obtained result is an image in which every pixel has two real values. One is the orientation of symmetry parameters and the other is one of the certainty measures. This representation of the features is due to Granlund, [8], and has the advantage that orientation estimates with high certainties are more visible than the estimates with low certainties if viewed on a colour TV monitor. Both of the proposed certainties of the orientation estimates are based on the error obtained in the least squares fitting of the orientation parameter.

The fact that the resulting image is a continuous image and the found circular symmetry for a neighbourhood is relevant in proportion to the obtained certainty makes these images useful feature images in areas like remote sensing, texture analysis and object recognition. Some examples of the latter are demonstrated in Experimental Results. In the next section it is shown that the orientation pointing out subclasses of the circularly symmetric patterns can be considered as the usual orientation of lines and edges in another coordinate system. Moreover, it is also shown that to every

circularly symmetric pattern class represented by an orientation parameter there is a corresponding set of rotations and scalings under which these circularly symmetric patterns are invariant. This is the reason why we choose to call these patterns circularly symmetric. In Section 3 the necessary tools to implement the algorithm of finding the orientation parameter along with energy dependent certainty estimation are given. In Section 4 the energy independent certainty estimation is analysed and the necessary tools are given. Section 5 proposes a method to implement what will be defined as *the partial derivative image*, which is necessary to compute the circular symmetry features. The last section reports the results of the experiments.

2 Description of the image by circularly symmetric isogray values

Our approach in this section in defining the circular symmetries will be based on the assumption that these are invariant to rotation and scaling. Rotation and scaling with respect to the origin can be performed by the following infinitesimal operators:

$$\begin{aligned}\mathcal{L}_r &= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} \\ \mathcal{L}_s &= x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.\end{aligned}$$

For instance to rotate the image f with the infinitesimal angle $\delta\varphi$ is equivalent to

$$f' = f + \delta\varphi\mathcal{L}_r f \quad (1)$$

where f' is the image after rotation. It can be thought that the operator \mathcal{L}_r is applied successively to perform rotations of considerable amount. The same holds for the scaling operator, \mathcal{L}_s . The linear combination of these operators gives the operator $a\mathcal{L}_r + b\mathcal{L}_s$ which corresponds to rotating the image by the amount a and then scaling it by the amount b . Here a and b are assumed to be small. Successive applications of this operator would then perform a large amount of rotation followed by a scaling. However, the proportion of the amount of rotation to the amount of scaling $a : b$ is constant under successive applications. If the amount of rotation and scaling is measured in an orthogonal coordinate system, then the locus of successive operations would result in a straight line with an inclination angle defined by the direction cosine, $(a, b)^t$. What patterns are invariant to a rotation by the amount a followed by a scaling by the amount b ? The question is equivalent to solving the differential equation:

$$(a\mathcal{L}_r + b\mathcal{L}_s)f = 0 \quad (2)$$

e.g.

$$(bx - ay)\frac{\partial f}{\partial x} + (ax + by)\frac{\partial f}{\partial y} = 0. \quad (3)$$

The solution of this is any image with the isogray value curves satisfying

$$a \ln r - b\varphi = c. \quad (4)$$

That is $f(r, \varphi) = g(a \ln r - b\varphi)$, where g is a one-dimensional function and $r = \sqrt{x^2 + y^2}$. Equation (4) describes an Archimedes' spiral. When the vector $(a, b)^t$ rotates, that is, when the proportion $a : b$ is changed, the equation (4) describes continuously different curves including half lines and circles. Putting $\xi = \ln r$ and $\varphi = \tan^{-1}(x, y)$ defines the coordinate system in which (4) becomes the equation of a line.

To illustrate the concept of a coordinate system as invariants of operators, consider yet another operator pair $\mathcal{L}_x = \frac{\partial}{\partial x}$ and $\mathcal{L}_y = \frac{\partial}{\partial y}$. These infinitesimal operators correspond to translations in horizontal and vertical directions. That is, to translate the local image f with the amount δx is equivalent to

$$f' = f + \delta x \mathcal{L}_x f. \quad (5)$$

The invariants of $a\mathcal{L}_x + b\mathcal{L}_y$ define the curves

$$-bx + ay = c.$$

Putting $a = 0$ or $b = 0$ defines the curves which are the basis of the familiar Cartesian coordinate system. Without commenting on the significance of the patterns like lines and Archimedes' spirals to human vision, it should be mentioned that Hoffman, [9], has shown the existence of a Lie algebra consisting of such operators. He calls this algebra the Lie algebra of visual perception and refers to experiments performed by, among others, Hubel & Wiesel, [12], Mac Kay, [11] for support of his theory.

Having summarized the relation between the Archimedes' spirals and the rotation and scaling processes we present the idea of describing any neighbourhood by means of functions having iso-values as these spirals. It is clear that after the coordinate transformation:

$$x = \exp \xi \cos \varphi \quad (6)$$

$$y = \exp \xi \sin \varphi \quad (7)$$

a local image represented in Cartesian coordinates can be transformed to an image which is described by the real function $f(\xi, \varphi)$ in the new coordinates $(\xi, \varphi)^t$. Under this coordinate transformation the invariants of $a\mathcal{L}_r + b\mathcal{L}_s$, (5) transform to lines of $a\xi - b\varphi = c$, since $\xi = \ln r$. A linear combination of rotation and scaling operators applied to the original local image is equivalent to translating the new image in a fixed direction. Since an arbitrary local image is not symmetric in the sense that it is changed when it is rotated and scaled, we can in general not expect to find $f(\xi, \varphi)$ consisting of parallel lines. To assign an orientation to every local image $f(\xi, \varphi)$ will be the same as looking for a direction along which a translation degrades the local image minimally. In the following this degradation measure and two certainty measures based on this degradation will be defined. The purpose of the certainty measures is to provide us with the necessary information to discriminate two local images with respect to the

rotation and scaling symmetry. That is, the obtained orientation estimate may be more reliable for one local image compared to another, very much in the same way as the usual lines and edges in low level vision may be "stronger" than one another, and accordingly the reliability of the associated orientation parameters varies.

Under quite general conditions imposed on f it is possible to Fourier transform f with respect to ξ and φ :

$$F(\omega, n) = \int_{-\infty}^{\infty} \int_0^{2\pi} \exp -i(\omega\xi + n\varphi) f(\xi, \varphi) d\xi d\varphi \quad (8)$$

with n being an integer. The inverse transform becomes:

$$f(\xi, \varphi) = (2\pi)^{-2} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \exp i(\omega\xi + n\varphi) F(\omega, n) d\omega. \quad (9)$$

This formula can be seen as a description of the local image $f(\xi, \varphi)$ in terms of the functions $\exp i(\omega\xi + n\varphi)$.

Definition 1 We will use the term *circularly symmetric* for the functions $f(\bar{r})$ which can be described by means of a one-dimensional real function g as:

$$f(\bar{r}) = g(\omega'\xi + n'\varphi) \quad (10)$$

where n' is an integer and ω' a real constant.

It is straightforward to show that the circularly symmetric local images are concentrated to a line passing through the origin in the Fourier domain defined by (8) and (9). The inclination angle of this line is given by

$$\theta = \tan^{-1}(\omega', n'). \quad (11)$$

The concentration of the circularly symmetric images in the Fourier domain makes it easy to identify them in this domain. One method of examining whether there exists a circular symmetry, in analogy with [3]¹, [2], is to fit a least square error line to the special Fourier transform domain defined by (8) and (9). In this model, detection is given for a bounded area. We will lift this restriction here and develop the theory for a general case. The reason for not considering a bounded area as a neighbourhood is that the presence of a sharp boundary affects the detection negatively. Modelling of the local neighbourhood is handled by multiplying the whole image by a positive function (window), which decreases smoothly with increased distance from the center of the examined point. The error made by the least square line fitting will be large if the local image is not circularly symmetric, and will vanish if, and only if, the local image is circularly symmetric. The error or degradation for a line with inclination angle θ is defined as:

$$E(\theta) = \sum_n \int_{-\infty}^{\infty} d^2(\bar{k}, \bar{k}_\theta) |F(\bar{k})|^2 d\omega \quad (12)$$

¹See Appendix B

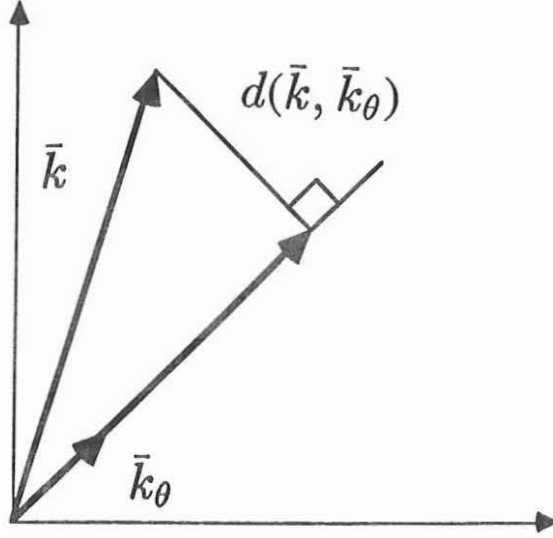


Figure 1: The figure illustrates the Euclidean distance, $d(\bar{k}, \bar{k}_\theta)$, between the point \bar{k} and the axis denoted by \bar{k}_θ .

Here $d^2(\bar{k}, \bar{k}_\theta)$ is the squared Euclidean distance between the vector $\bar{k} = (\omega, n)^t$ and $\bar{k}_\theta = (\cos \theta, \sin \theta)^t$, of which the former symbolizes the transposed coordinate vector in the special Fourier domain. By substituting

$$d^2(\bar{k}, \bar{k}_\theta) = \omega^2 + n^2 - (\omega \cos \theta + n \sin \theta)^2 \quad (13)$$

in (12) and using some familiar trigonometric relationships, Figure 1, we obtain

$$E(\theta) = \frac{1}{2}[\omega_d^2 + n_d^2 + \sqrt{(\omega_d^2 - n_d^2)^2 + 4p_d^2} \cos(2\theta + \theta_0)] \quad (14)$$

where

$$\begin{aligned} \theta_0 &= \tan^{-1}(n_d^2 - \omega_d^2, 2p_d) \\ \omega_d^2 &= \sum_n \int_{-\infty}^{\infty} \omega^2 |F(\omega, n)|^2 d\omega \\ n_d^2 &= \sum_n n^2 \int_{-\infty}^{\infty} |F(\omega, n)|^2 d\omega \\ p_d &= \sum_n \int_{-\infty}^{\infty} n\omega |F(\omega, n)|^2 d\omega. \end{aligned}$$

Consequently θ minimizing $E(\theta)$ is given by

$$2\theta_{min} = \tan^{-1}(\omega_d^2 - n_d^2, 2p_d) \quad (15)$$

since this choice of 2θ results in the cosine term attaining the value -1 . The corresponding minimum is given by:

$$E(\theta_{min}) = \frac{1}{2}[\omega_d^2 + n_d^2 - \sqrt{(\omega_d^2 - n_d^2)^2 + 4p_d^2}]. \quad (16)$$

$E(\theta_{min})$ is a measure of degradation. To get a certainty measure we propose to look at whether the obtained error of the fitted line is really small compared to the error of the line giving the maximum error and form,

$$C_{f1} = E(\theta_{maz}) - E(\theta_{min}). \quad (17)$$

Here $E(\theta_{maz})$ is the error of the line maximizing (12) and consequently this line and the line giving the minimal error are orthogonal. It is straightforward to show that

$$C_{f1} = \sqrt{(\omega_d^2 - n_d^2)^2 + 4p_d^2}. \quad (18)$$

According to (17), C_{f1} increases with increased difference between the errors obtained when the worst and best lines are fitted. Thus the magnitude of the complex number

$$z_1 = \omega_d^2 - n_d^2 + i2p_d = C_{f1} \exp(i2\theta). \quad (19)$$

will give a measure of whether there exists a circular symmetry, and the argument of it, 2θ , will give an estimation of the orientation of this circular symmetry. It should be observed that it is far more convenient to represent the inclination angle of the line by 2θ than θ , since the orientation of a line is not altered by rotating it π radians, while the addition of a π to θ would result in an angle other than θ . This ambiguity is removed by the mapping 2θ , [8].

The certainty measure C_{f1} is dependent on the energy of the local image and hence not dimensionless. Another measure of certainty is:

$$C_{f2} = \left(\frac{E(\theta_{maz}) - E(\theta_{min})}{E(\theta_{maz}) + E(\theta_{min})} \right)^c \quad (20)$$

where c is a positive constant controlling the dynamic range of C_{f2} . That is

$$C_{f2} = \left(\frac{\sqrt{(\omega_d^2 - n_d^2)^2 + 4p_d^2}}{\omega_d^2 + n_d^2} \right)^c = \left(\frac{|z_1|}{\omega_d^2 + n_d^2} \right)^c. \quad (21)$$

Since E is a non-negative function by definition, $C_{f2} = 1$ if, and only if, $E(\theta_{min}) = 0$, that is, when we have a truly circularly symmetric local image. To represent the certainty and orientation we use the same notation as before, namely we represent both simultaneously by means of a complex variable:

$$z_2 = \frac{z_1}{\omega_d^2 + n_d^2} = C_{f2} \exp(i2\theta). \quad (22)$$

The calculations of z_1 and z_2 for every neighbourhood, however, can be managed in the spatial domain, although they are defined for the special Fourier domain given previously by using the Parseval relation:

$$\omega_d^2 = (2\pi)^2 \int_{-\infty}^{\infty} \int_0^{2\pi} \left(\frac{\partial f}{\partial \xi}\right)^2 d\xi d\varphi \quad (23)$$

$$n_d^2 = (2\pi)^2 \int_{-\infty}^{\infty} \int_0^{2\pi} \left(\frac{\partial f}{\partial \varphi}\right)^2 d\xi d\varphi \quad (24)$$

$$p_d = (2\pi)^2 \int_{-\infty}^{\infty} \int_0^{2\pi} \frac{\partial f}{\partial \varphi} \frac{\partial f}{\partial \xi} d\xi d\varphi. \quad (25)$$

3 Evaluation of the energy dependent certainty

To be able to calculate ω_d^2 , n_d^2 and p_d at every point of a discretized image, it is desirable to simplify the calculations proposed by (23). By using the chain rule we have:

$$\frac{\partial f}{\partial \xi} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \xi} \quad (26)$$

$$\frac{\partial f}{\partial \varphi} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \varphi} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \varphi}. \quad (27)$$

The reason for choosing Cartesian directions x and y is that the image is stored as a set of discrete values on a rectangular mesh. Assuming f to be a bandlimited function (in the frequency domain of f with respect to x and y in the usual sense) leads us to the fact that $(\frac{\partial f}{\partial x})^2$, $(\frac{\partial f}{\partial y})^2$ and $\frac{\partial f}{\partial x} \frac{\partial f}{\partial y}$ are band limited as well and can be reconstructed from the samples, f_{xj} and f_{yj} , by:

$$\left(\frac{\partial f}{\partial x}\right)^2 = \sum_j f_{xj}^2 \mu_j(\bar{r})$$

$$\left(\frac{\partial f}{\partial y}\right)^2 = \sum_j f_{yj}^2 \mu_j(\bar{r})$$

$$\frac{\partial f}{\partial x} \frac{\partial f}{\partial y} = \sum_j f_{xj} f_{yj} \mu_j(\bar{r}).$$

Here $\mu_j(\bar{r})$ is the interpixel function governing the behaviour of the image between the image points and it will be specified later. f_{xj} and f_{yj} are the samples of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at the point \bar{r}_j . Thus ω_d^2 is given by:

$$\omega_d^2 = (2\pi)^2 \sum_j \int_{-\infty}^{\infty} \int_0^{2\pi} \mu_j(\bar{r}) [f_{xj}^2 \left(\frac{\partial x}{\partial \xi}\right)^2 + f_{yj}^2 \left(\frac{\partial y}{\partial \xi}\right)^2 + 2f_{xj} f_{yj} \frac{\partial x}{\partial \xi} \frac{\partial y}{\partial \xi}] d\xi d\varphi. \quad (28)$$

Using the relations (6) (7) (23) (28) we obtain

$$\omega_d^2 = \sum_j \int_{-\infty}^{\infty} \int_0^{2\pi} \mu_j(\bar{r}) (f_{xj}^2 x^2 + f_{yj}^2 y^2 + 2f_{xj} f_{yj} xy) d\xi d\varphi. \quad (29)$$

Similarly n_d^2 and p_d are obtained. Thus z_1 yields:

$$z_1 = (2\pi)^2 \sum_j (f_{zj} + if_{yj})^2 \int_{-\infty}^{\infty} \int_0^{2\pi} \mu_j(\bar{r})(x - iy)^2 d\xi d\varphi = \sum_j (f_{zj} + if_{yj})^2 v_j. \quad (30)$$

Obviously it is possible to evaluate z_1 as a convolution according to (30).

So far we have not specified $\mu_j(\bar{r})$. It is the analytic function obtained by inverse Fourier transforming a region function, since f is bandlimited. By region function is understood a function which is 0 outside a closed region and 1 inside it. For the sake of simplicity we put μ_j as a translated gaussian, and the examined point according to which ξ and φ are defined by (6) and (7) as origin. Although this choice of $\mu_j(\bar{r})$ does not provide us with a perfect interpixel function since it is not bandlimited in the strict sense, we will use it for a number of its properties as an approximation of the ideal interpixel function. The most important of these is its concentration in both the spatial and the frequency domain. The latter is especially useful since the resulting filters become manageably small. Thus we have

$$\mu_j(\bar{r}) = \exp(-\beta\|\bar{r} - \bar{r}_j\|^2) \exp(-\alpha\|\bar{r}\|^2). \quad (31)$$

Here β is a positive constant governing the behaviour of the continuous image between discrete image positions while α , also a positive constant but chosen smaller than β , controls the width of the neighbourhood to be considered. Thus (30) and (31) define the filter coefficients v_j :

$$\begin{aligned} v_j &= (2\pi)^2 \int_{-\infty}^{\infty} \int_0^{2\pi} \mu_j(\bar{r})(x - iy)^2 d\xi d\varphi \\ &= (2\pi)^2 \int_0^{\infty} \int_0^{2\pi} \exp(-\beta\|\bar{r} - \bar{r}_j\|^2 - \alpha r^2)(x - iy)^2 \frac{dr}{r} d\varphi \end{aligned} \quad (32)$$

since $\xi = \ln r$ and $r = \|\bar{r}\|$. Remembering $\|\bar{r}\|^2 = \bar{r}^t \bar{r}$ and $x - iy = r \exp(-i\varphi)$ we have:

$$v_j = 4\pi^3 (\alpha + \beta)^{-1} \exp\left(-\frac{\beta\alpha}{\beta + \alpha} r_j^2\right) a\left(\frac{2\beta r_j}{\sqrt{\alpha + \beta}}\right) \exp(-i2\varphi_j) \quad (33)$$

where $a(x)$, Figure 2, is the dimensionless function given by:

$$a(x) = \frac{2}{\pi} \int_0^{\infty} \int_0^{\pi} \exp\left(-r^2 + xr \cos \varphi - \frac{x^2}{4}\right) r \cos(2\varphi) dr d\varphi. \quad (34)$$

$a(x)$ is monotonous, even and

$$\lim_{x \rightarrow \infty} a(x) = 1. \quad (35)$$

Consequently (34) can be used to construct a table for $a(x)$. The values between table entries can be evaluated by using numerical interpolation techniques. Once this table is established, the advantage of this approach is that $a(x)$ can be evaluated rapidly for future kernel generations with arbitrary α and β without time-consuming numerical integrations. Since $a(x)$ is smooth, the interpolation techniques yield acceptable accuracy for points between the table entries. In the experiments below a VAX 11/750 is

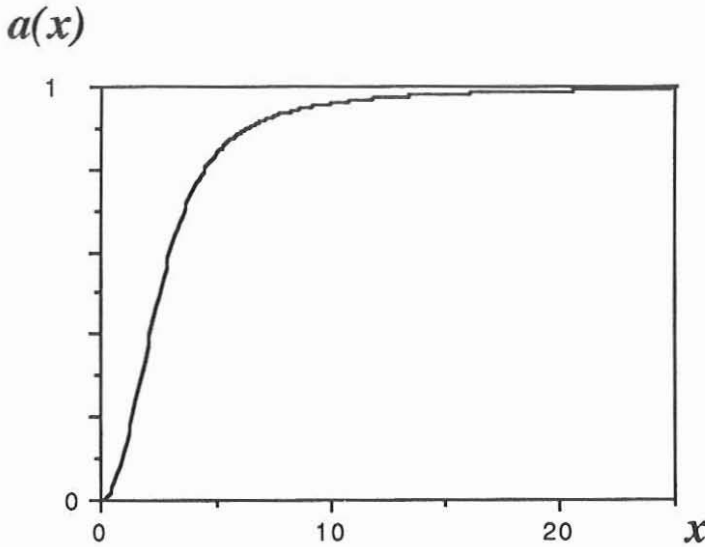


Figure 2: The function used to evaluate the kernels for circular symmetry detection purposes.

used to evaluate the table entries by means of two one-dimensional, adaptive numerical integrations chosen from the SLATEC subroutine library, and Newton's formula with 5 steps, Björck, [13], is used for the purpose of interpolation. The constructed table had entries between 0 and 25 of equal increments of 0.1.

Figure 3 illustrates v_j . Since the asymptotic behaviour of $|v_j|$ is proportional to a gaussian, the filter coefficients decrease rapidly as r_j becomes large. Filter v with coefficients v_j can be approximated by a truncated filter. In the experiments below, v is truncated at a radius, at which the magnitude of the filter is less than 1% of the maximum of the filter.

Thus evaluation of z_1 for every point in the image can be accomplished as a convolution involving multiplication and summation of complex numbers:

$$z_1 = \sum_j (f_{x_j}^2 + f_{y_j}^2) \exp(i2\varphi_{f_j}) v_j = \sum_j (f_{x_j}^2 + f_{y_j}^2) \exp(i2\varphi_{f_j}) g_j a_j \exp(-i2\varphi_j) \quad (36)$$

where

$$\begin{aligned} \varphi_{f_j} &= \arg(f_{x_j} + if_{y_j}) \\ g_j &= 4\pi^3(\alpha + \beta)^{-1} \exp\left(-\frac{\alpha\beta}{\alpha + \beta} r_j^2\right) \\ a_j &= a\left(\frac{2\beta r_j}{\sqrt{\alpha + \beta}}\right) \\ \varphi_j &= \arg(x_j + iy_j). \end{aligned}$$

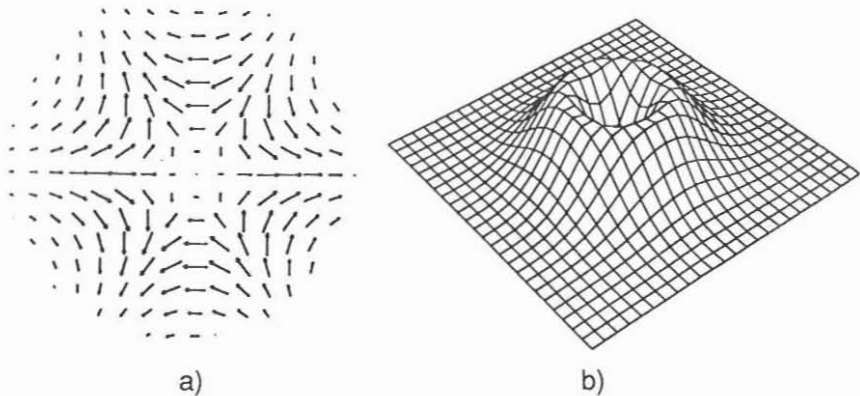


Figure 3: The figure illustrates a) the complex valued v_j b) a 3-D plot of its magnitude

Implementation of (36) is given by the flow chart in Figure 4 and yields:

1. Obtain the partial derivative image, $(f_{z_j}^2 + f_{v_j}^2) \exp(i2\varphi_{f_j})$.
2. Convolve the partial derivative image with the filter given by $g_j a_j \exp(-i2\varphi_j)$. The resulting complex valued image will contain the local orientation estimate of the scaling/rotation space at the arguments of every point, while a certainty measure for this estimate will be contained at the magnitudes. The locally orthogonal circularly symmetric neighbourhoods will be represented by vectors with opposite directions, such as circle and star-shaped neighbourhoods.

To utilize the full dynamics of the limited number of bits per pixel allocated for the magnitudes, both pictures obtained by step 1) and step 2) should be scaled by the maximum magnitudes of the obtained images respectively, if necessary. For example, if the number of available intensity levels are 256, then the maximum magnitude in the image should be mapped to maximum of these intensity levels.

4 Evaluation of the energy independent certainty

Since z_2 and z_1 are related to each other by (22), it is only necessary to derive $E(\theta_{maz}) + E(\theta_{min})$. ω_d^2 is given by (29), and a similar evaluation for n_d^2 implies

$$E(\theta_{maz}) + E(\theta_{min}) = \omega_d^2 + n_d^2 =$$

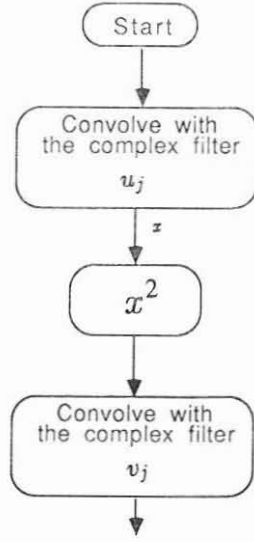


Figure 4: The flow chart of the algorithm computing energy dependent certainty together with optimal orientation estimate of circular symmetry parameters. The resulting image is complex valued with magnitudes of the pixels being certainties and arguments being orientations.

$$\begin{aligned}
 &= (2\pi)^2 \sum_j \int_{-\infty}^{\infty} \int_0^{2\pi} \mu_j(\bar{r}) (f_{z_j}^2 x^2 + f_{y_j}^2 y^2 + f_{z_j}^2 y^2 + f_{y_j}^2 x^2) d\xi d\varphi \\
 &= (2\pi)^2 \sum_j (f_{z_j}^2 + f_{y_j}^2) \int_0^{\infty} \int_0^{2\pi} \exp(-\beta \|\bar{r} - \bar{r}_j\|^2 - \alpha r^2) r dr d\varphi.
 \end{aligned}$$

But since

$$\beta \|\bar{r} - \bar{r}_j\|^2 + \alpha r^2 = (\alpha + \beta) \|\bar{r} - \frac{\beta}{\alpha + \beta} \bar{r}_j\|^2 + \frac{\alpha\beta}{\alpha + \beta} r_j^2 \quad (37)$$

we obtain

$$\omega_d^2 + n_d^2 = \sum_j (f_{z_j}^2 + f_{y_j}^2) (2\pi)^2 \exp\left(-\frac{\alpha\beta}{\alpha + \beta} r_j^2\right) \int_0^{\infty} \int_0^{2\pi} \exp\left(-(\alpha + \beta) \|\bar{r} - \frac{\beta}{\alpha + \beta} \bar{r}_j\|^2\right) r dr d\varphi \quad (38)$$

The integral term is nothing but the volume under a translated gaussian, and thus it is constant with respect to \bar{r}_j . The value of this constant is $\pi(\alpha + \beta)^{-1}$. Thus $E(\theta_{max}) + E(\theta_{min})$ is obtained as the convolution:

$$\omega_d^2 + n_d^2 = \sum_j (f_{z_j}^2 + f_{y_j}^2) \kappa_j \quad (39)$$

with

$$\kappa_j = 4\pi^3 (\alpha + \beta)^{-1} \exp\left(-\frac{\alpha\beta}{\alpha + \beta} r_j^2\right) = g_j. \quad (40)$$

Here we have assumed that the filter κ , defined by (40), is truncated at the same radius as the truncation radius of v given by (33) yielding a comparable accuracy. Thus by using (22) we obtain

$$z_2 = \left(\frac{z_1}{\sum_j (f_{x_j}^2 + f_{y_j}^2) g_j} \right)^c = \left(\frac{\sum_j (f_{x_j}^2 + f_{y_j}^2) \exp(i2\varphi_{f_j}) g_j a_j \exp(-i2\varphi_j)}{\sum_j (f_{x_j}^2 + f_{y_j}^2) g_j} \right)^c. \quad (41)$$

Since $a_j \leq 1$, the magnitude of z_2 will have the upper bound 1, and it will be attained by local images f having the highest partial derivative energies concentrated close to the boundary of the filter v . Moreover, the $\arg z_2 = \arg z_1$ for a neighbourhood giving the maximum magnitude should be $2\varphi_{f_j} = 2\varphi_j + \text{constant}$. To prove these, we apply the triangle inequality to (41).

$$|z_2| \leq \frac{\sum_j (f_{x_j}^2 + f_{y_j}^2) g_j a_j}{\sum_j (f_{x_j}^2 + f_{y_j}^2) g_j} \quad (42)$$

This is fulfilled with equality if $2\varphi_{f_j} - 2\varphi_j = \text{constant}$, since the triangle inequality holds with equality only for parallel complex numbers. Assume that the angle variation of $(f_{x_j} + if_{y_j})^2$ is such a variation and equality holds. Putting

$$P_j = \frac{(f_{x_j}^2 + f_{y_j}^2) g_j}{\sum_j (f_{x_j}^2 + f_{y_j}^2) g_j} \quad (43)$$

defines P_j as a probability distribution, since P_j 's are positive and sum up to unity. By summing up all probabilities originating from points on the same circle with the radius r_j and calling it a new probability P'_{r_j}

$$|z_2| = \sum_j P_j a_j = \sum_{r_k} a_{r_k} \sum_{z_j^2 + y_j^2 = r_k^2} P_j = \sum_{r_j} a_{r_j} P'_{r_j} \quad (44)$$

is obtained. Since $a_{r_j} \geq 0$ and monotonous, the maximum is attained at the rand, giving $P'_R = 1$, where R is the radius of the filter. This last property is in fact a result of sampling the circularly symmetric images by means of a quadratic mesh and indicates that the samples close to the origin should not be taken too seriously in circular symmetry detection, which appeals to engineering intuition. The detection of linear symmetry orientation, which is investigated for n -dimensional images earlier, [1], indicates that all samples are considered as equally important, since these images can be sampled by a quadratic mesh and reconstructed exactly from these samples. But this is not the case for circularly symmetric images, and hence an optimal weighting should be done for a given choice of interpixel function. This is achieved by $a(x)$ for the interpixel functions chosen as gaussians. But it is straightforward to calculate similar $a(x)$ for other $\mu_j(\bar{r})$, by utilizing (32).

Implementation of (41) is analogous to implementation of (36) and is given by the flow chart in Figure 5. z_2 is obtained in two steps, the first being identical to the first step of implementation of z_1 :

1. Obtain the partial derivative image, $(f_{x_j}^2 + f_{y_j}^2) \exp(i2\varphi_{f_j})$.

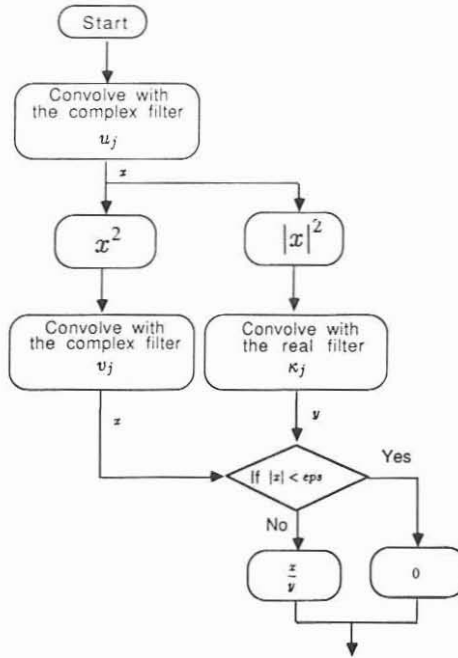


Figure 5: The flow chart of the algorithm computing energy independent certainty together with optimal orientation estimate of circular symmetry parameters. The resulting image is complex valued and ready for interpretation in terms of certainty and orientation as in the previous flow chart.

2. Perform two convolutions on the partial derivative image according to (36) and (39). It should be observed that (39) is a convolution of only the magnitude of the partial derivative image with a real valued filter. Divide the first obtained complex number with the real number obtained later and exponentiate it to a constant real number c . The last stage should be performed if the numerator has a value larger than a small threshold to ensure the numerical stability. Otherwise put the magnitude to zero, which means total uncertainty. Since the numerator may involve a large number of additions of vectors which in general are not all in the same direction, the well-known numerical error, cancellation of terms may occur. Consequently the zero level of the numerator increases. By zero level we mean the level below which the obtained values are interpreted as zero. By ensuring that the magnitude of the numerator exceeds this threshold, division by zero is avoided at the same time since $|z_2| \leq 1$, and thereby it is clear that the magnitude of the numerator does not exceed the magnitude of the denominator. The value of this threshold can either be chosen experimentally or by using numerical error analysis for the filters considered. The resulting image is ready to be interpreted as in the previous section.

5 Evaluation of the partial derivative image

In the previous sections two two-step algorithms were considered. They both deliver a complex number per pixel containing information about the circular symmetry existence around every pixel in the image. To be able to do that, *the partial derivative image*, which is also a complex valued image, is required. The technique used for estimation of $(f_{x_j} + i f_{y_j})^2$ is basically the same as utilized in the previous sections.

Reconstruct the image by means of its discrete samples, f_j , and apply $\frac{\partial}{\partial x} + i \frac{\partial}{\partial y}$ operator:

$$\left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y}\right)f(\bar{r}) = \sum_j f_j \left(\frac{\partial \mu_j(\bar{r})}{\partial x} + i \frac{\partial \mu_j(\bar{r})}{\partial y}\right). \quad (45)$$

By using the same interpixel function as before, (31) and equation (37)

$$\begin{aligned} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y}\right)f(\bar{0}) &= \sum_j -2f_j(\alpha + \beta) \left(\left(x - \frac{\beta}{\alpha + \beta}x_j\right) + i \left(y - \frac{\beta}{\alpha + \beta}y_j\right) \right) \mu_j(\bar{r}) \Big|_{\bar{r}=\bar{0}} \\ &= \sum_j f_j u_j \end{aligned} \quad (46)$$

are obtained for the origin. u_j is defined as:

$$u_j = 2\beta(x_j + iy_j) \exp(-\beta(x_j^2 + y_j^2)) \quad (47)$$

Obviously (46) can be considered as a convolution with a complex valued filter u , with the coefficients u_j given by (47). The coefficients u_j decrease rapidly as \bar{r}_j becomes large and hence can be truncated. In the experiments we have truncated u_j when $|u_j|$ has reached 1 % of its maximum. The implementation of the estimation of $(\mathcal{L}_x f + i \mathcal{L}_y f)^2$ is possible to accomplish by:

- Convolve the image f with the filter u . Assign to every pixel of the new image the square of the obtained complex value.

6 Experimental results

To verify that the algorithms described in the previous sections really work on discrete image data, they have been tested on a number of both synthetic and real images. Some of these images will be presented and discussed in detail. A GOP-300 is used for implementation of the algorithms. Figures 4 and 5 illustrate the flow charts of the implemented algorithms z_1 which represents the orientation estimate and its corresponding energy dependent certainty, and z_2 which represents the orientation estimate and its corresponding energy independent certainty. Every box in the flow chart should be considered as a unit with inputs and outputs being images. Except for the convolution boxes, all boxes consist of pointwise operating functions. Two groups of experiments are presented 1) Noise behaviour tests 2) Application examples.

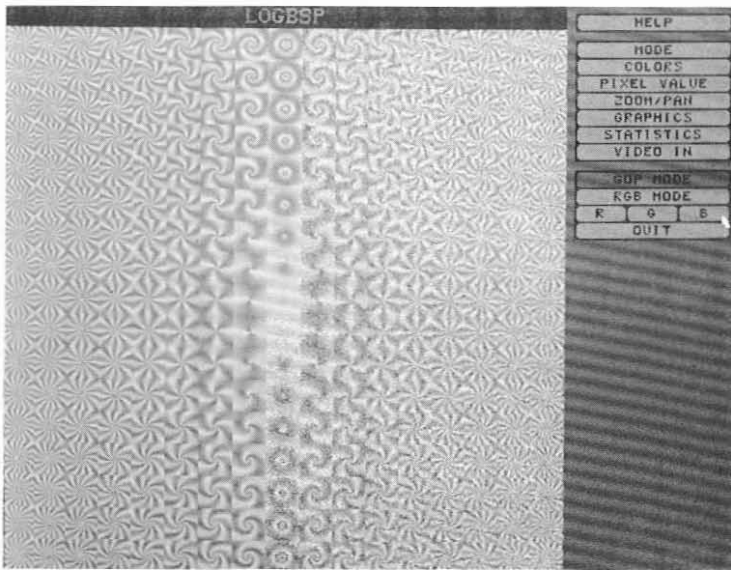


Figure 6: The test image above illustrates some circularly symmetric neighbourhoods. Gaussian white noise is added to the right half of the image. The image is of size 512x512

6.1 Noise behaviour tests

To investigate the behaviour of the algorithms in the presence of noise we have used the synthesized image shown in Figure 6. It consists of square blocks of the size 30X30 containing circularly symmetric images. The blocks in the left half of the image are generated by

$$(1 + 0.25 \cos(\omega \ln r + n\varphi))/2 \quad (48)$$

and in the right half of the image by

$$(1 + 0.25 \cos(\omega \ln r + n\varphi) + 0.75X)/2 \quad (49)$$

X is gaussian, uncorrelated noise $\sim N(0, 32)$. The values obtained by these functions are mapped to 256 gray levels. In all colour images the resolution in both the intensity and the hue consists of 256 discrete values. The peak-to-peak variation of the noise is three times the size of the cosine term in the right half of the image. Within every block w and n are constant. w changes proportionally to the y coordinates of the block centers, and n changes proportionally to the x coordinates of the block centers. The origin is the center of the picture. Thus, if a block center has the coordinate $x\hat{x} + y\hat{y}$, then the orientation of rotation/scaling at the center of the block is given by

$$2\theta = 2 \tan^{-1}(n, w) = 2 \tan^{-1}(x, y). \quad (50)$$

That is, the blocks with coordinates $x\hat{x} + y\hat{y}$ and $-x\hat{x} - y\hat{y}$ have the same orientations except for the noise. This property of the test image can be utilized to examine the noise behaviour of the algorithms.

- a) z_1 of this image is given in Figure 10. The filter used for the convolution was given in Section 3. The filter size was 25x25 with $\alpha = 1/16$. To obtain the intermediate result, the partial derivative image, a 3x3 filter is used according to the derivation in Section 5. It can clearly be seen that the block centers on a line passing through the center of the image has the same colour at the points of the right half of the image. In this representation, we have chosen green to represent circular and red to represent the star-shaped circular symmetries. All other circular symmetries with the orientations between them are mapped to colours, in such a way that this change in colours is perceived as a smooth continuous variation. This is done by utilizing I component of the IHS colour representation to represent the magnitude of z_1 and H to represent the argument of z_1 . The certainty level decreases gradually as the distance increases from the block centers until the block borders are reached, where it increases slightly and attains a local peak at the block corner. This is natural, since the neighbourhoods of the corners resemble the star-shaped circularly symmetric patterns. For this reason all the corners have barely visible red colour with an intensity equal to approximately half of the intensity of the block centers, which have the highest intensities. The certainties in the noisy part of the image are of the same order as in the left half of the image at the block centers. This property is due to the fact that z_1 is dependent on the energy of the partial derivative image, that is the energy of the variation. The right half of the image contains more energy compared to the left, which has boosted up the certainty to the level of the left part. The colour components of the block centers corresponding to each other on both sides of the image differ only by 1 or 2 degrees.
- b) Figure 9 illustrates z_2 obtained for all points of Figure 6 with the same filter parameters as before. The zero level, below which the numerator is considered as zero, here called *eps*, is equal to 1. *c*, the dynamic range controller is put to 1. The value of *eps* is chosen as 3% of maximum of $|z_1|$. The colours, e g orientations, are identical to that of Figure 10. The certainties are dimensionless and relative, contrary to the certainties in Figure 10. The certainties at the block corners are about half of the certainties at the closest block centers except for those close to the origin, the center of the image. The corresponding block centers in noisy and noiseless parts of the image show a significant difference in intensities compared to Figure 10. The noisy centers have certainty values about half the values of centers with no noise. The explanation of the phenomenon that certainties at the block borders in the vicinity of the origin are higher than their respective block centers is the cancellation of terms. To show that we have increased *eps* stepwise, the result of which forced those block corners close to the image center gradually to vanish without any interference with the block centers. Increasing *eps* had the

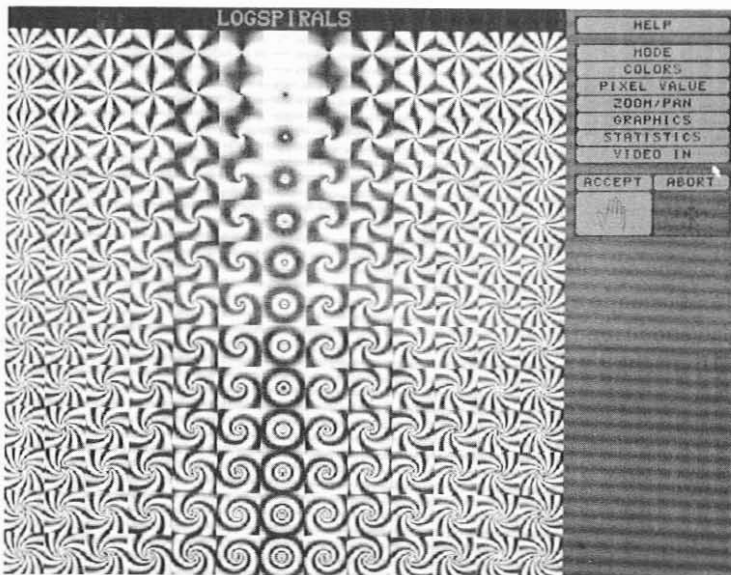


Figure 7: The image illustrates some synthetic blocks to be classified with respect to block types. The classes are right rotations, left rotations, circles, and stars.

effect of eliminating the block borders gradually, starting from the central part of the test image and moving towards its border. Figure 11 shows the obtained z_2 , with $eps = 8$, where most block borders are forced to zero, while the centers remain untouched.

6.2 Application examples

Finally, we give two examples of applications which use the resulting images of the proposed algorithms as features in pattern recognition.

- a) Figure 7 is a synthetic image to be classified with respect to “left-handedness”, “right-handedness”, “circular patterns” and “star-formed patterns”. The magnitudes and arguments of Figure 12 are utilized along with box classification to obtain the classified image in Figure 13.
- b) Figure 8 is an image of the sea bottom, where the aim is to classify it with respect to the sea anemones. Since the average size of these is much larger, about 100x100 pixels, than a practical filter size, we shrink the partial derivative image from 512x512 to 128x128, Figure 14 and Figure 15. The shrinking process is preceded by lowpass filtering to avoid aliasing. Then this image is used to obtain z_2 , resulting in Figure 16. The filter used is the same as before, with the size of 25x25. The parameter eps is put to 3, which is 9% of the maximum $|z_1|$ of the image.

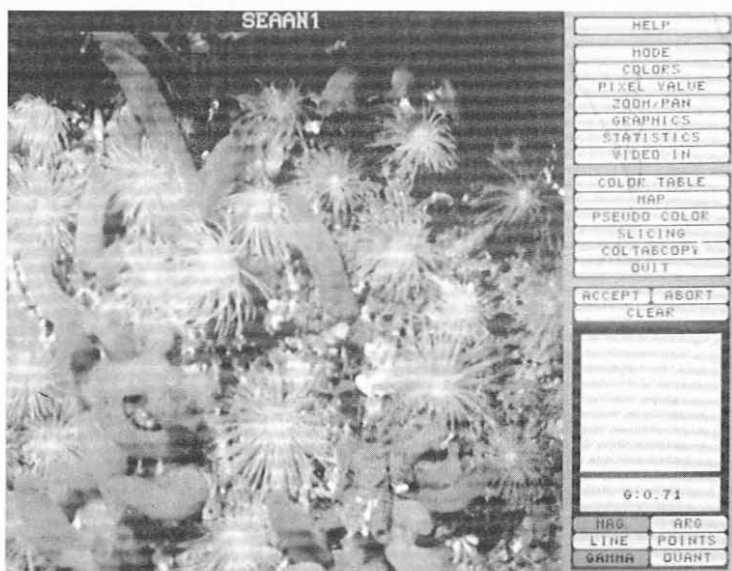


Figure 8: The image is a sea bottom photograph. The objective is to identify sea anemones.

A box classification based on Figure 8 and Figure 16 results in 90% detection of the anemones, Figure 17. Here it should be mentioned that the tune-up of the algorithms is possible and is carried out to increase the recognition rate of the sea anemones by applying the energy independent linear symmetry operator, [1], before shrinking it. But since the purpose of this paper is to demonstrate theoretical and practical aspects of circular symmetry modelling and its implementation, we will leave it out of consideration here.

7 Conclusion

In the Sections 2, 3 and 4, estimates of ω_d , n_d , p_d , $(f_x + if_y)^2$ are used to obtain an estimate of the optimal local orientation in the scaling/rotation space. The theoretically crucial point was to assume the bandlimitedness of the considered image. Although the squaring of the convolution results appearing at the first step, requires images oversampled with a factor of at least 2, experiments with practical applications have shown that most images of normal resolution fulfill this requirement. Thus it is not necessary to do anything special to eliminate errors originating from this source. In the cases where this really is a major source of error, it can easily be removed by oversampling the original image. We believe that the resulting images can be used as features in many classification applications since they describe a structural property of the neighbourhoods.

It is interesting to note that the obtained partial derivative image can be used by symmetry describing algorithms other than the circular symmetry description proposed here. An example of this is linear symmetry description, [1], which approximates the optimal local orientation of linear symmetry in the image, using the same image obtained in step 1. The only difference is the filters used for the convolutions in the second step. An extension of the circular symmetry description approach, based on the isogray value curves as proposed here, to other more general symmetries is under consideration. The preliminary results would indicate that a generalization is possible.

Acknowledgements

Credit should be given to Prof. Gösta Granlund, for suggesting the theme of this paper and being a constant source of inspiration in the course of the project; to the Computer Vision Laboratory staff, especially to Dr. Hans Knutsson, for constructive discussions; and, finally, to Swedish National Board for Technical Development for their financial support.

References

- [1] A. Rosenfeld "Survey, Picture Processing: 1984" *Computer Vision, Graphics, And Image Processing* **30**, pp. 189-242 (1985), sections (F), (G) and (H) of the paper.
- [2] R. O. Duda, P. E. Hart " Use of the Hough transformation to detect lines and curves in pictures" *Comm. ACM* **15**, 1, January 1972, 11-15.
- [3] B. H. Ballard "Generalizing the Hough transform to detect arbitrary shapes" *Pattern Recognition* **13**, 2, 1981, 111-112.
- [4] J. Bigün, "Circular symmetry models in image processing", Linköping studies in science and technology thesis No:85, 1986.
- [5] J. Bigün, G.H. Granlund "Central symmetry modelling", Third European signal processing conference, The Hague, sep 3-5 1986 pp. 883-886
- [6] J. Bigün, G.H. Granlund "Optimal orientation detection of linear symmetry", First international conf. on computer vision, London, June 1987. pp. 433-438
- [7] G.H. Granlund "In Search of a General Picture Processing Operator", *Computer Graphics and Image Processing* **8**, 155-173 (1978).
- [8] R. Jain "Image understanding for robotics applications" *Machine Intelligence and knowledge engineering for robotic applications*. Edited by A. K. C. Wong and A. Pugh, Springer-Verlag Berlin, Heidelberg 1987.
- [9] H. Knutsson "Filtering and reconstruction in image processing", Linköping studies in science and technology Dissertations No:88, 1982.
- [10] W.C. Hoffman "The lie algebra of visual perception" *J.math. Psychol.* 1966 **3** p65-98
- [11] D. M. MacKay "Interactive processes in visual perception" In W.A. Rosenblith (Ed.) *sensory communication*. New York: Wiley, 1961, pp. 339-355.
- [12] D. H. Hubel, T. N. Wiesel "Brain Mechanisms of Vision" *Scientific American*, sep. 79 Vol.241 No.3 pp. 150-162
- [13] G. Dahlquist, A. Björck "Numerical Methods." Prentice Hall, Englewood Cliffs, N.J., 1974.

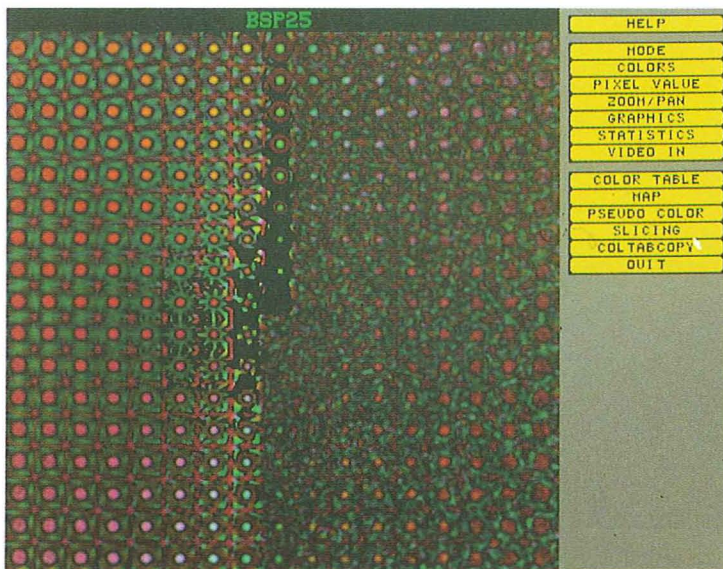


Figure 9: The image illustrates z_2 at every point of original image of Figure 6. The light intensities, i.e. the certainties, are independent of the energy. ϵ_{ps} is 1.

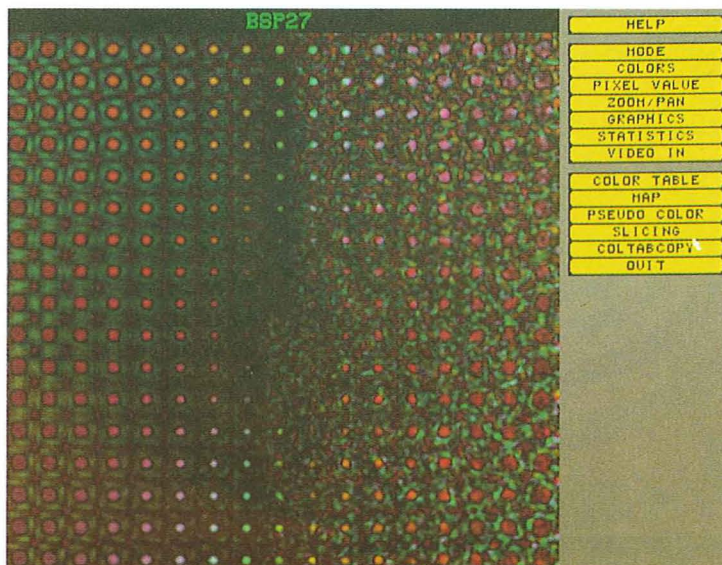


Figure 10: The image illustrates z_1 at every point of original image of Figure 6. The hue is the orientation of the neighbourhood in the rotation/scaling space. The intensity represents the certainty of this estimate.

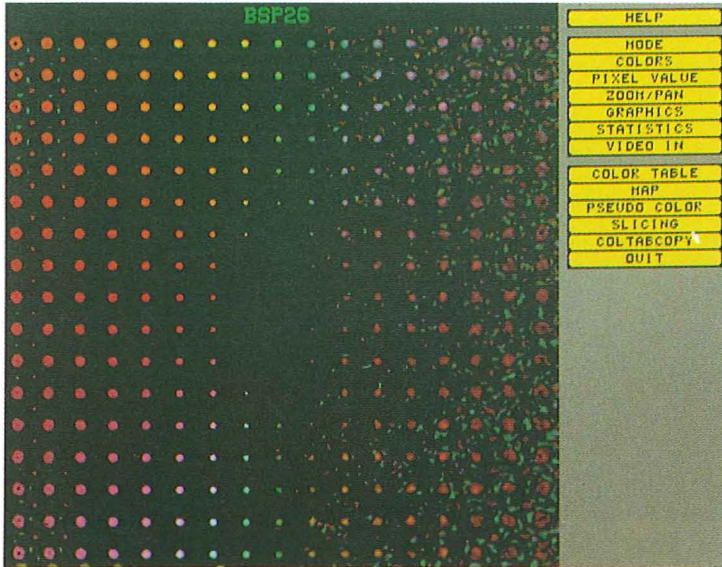


Figure 11: Same as Figure 9, but eps is chosen as 8. By changing eps one can drastically adjust the visibility of the points with circularly symmetric neighbourhoods, in comparison with points with neighbourhoods departing from this symmetry.

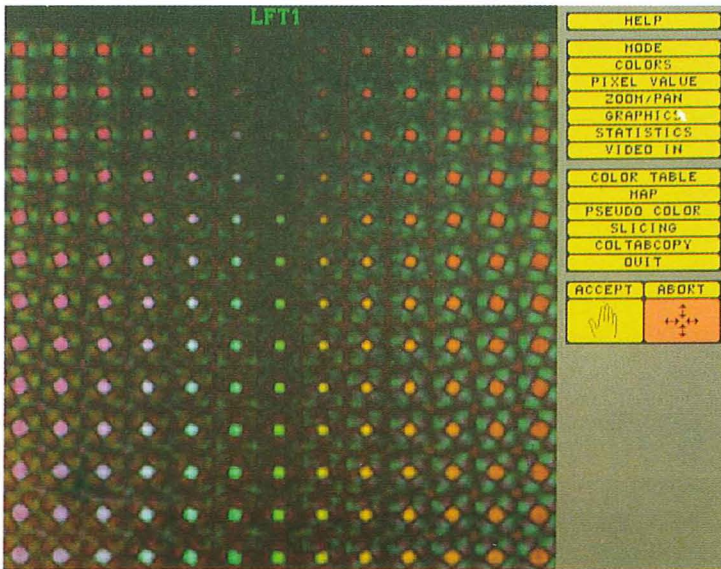


Figure 12: The image illustrates z_1 of Figure 7.

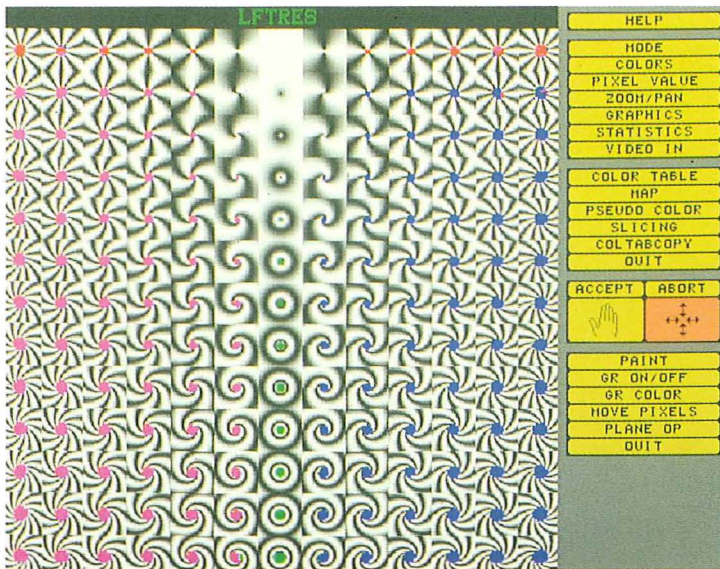


Figure 13: Box classification is used on Figure 12 to obtain this classified image.

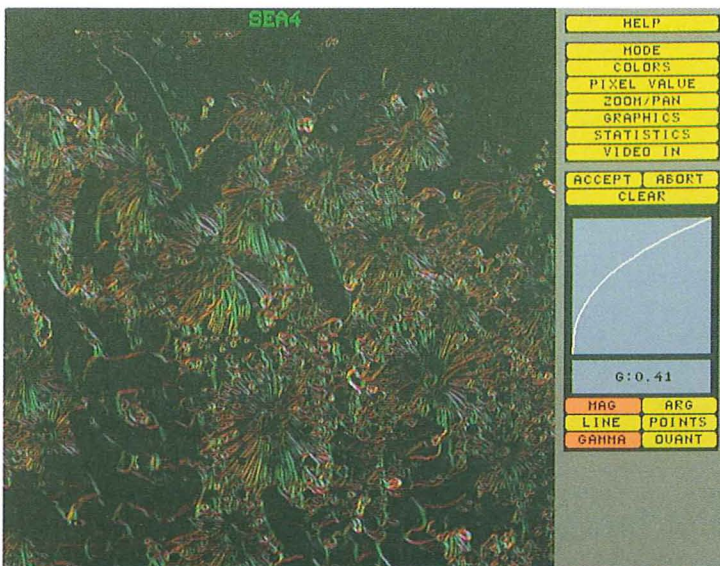


Figure 14: Partial derivatives of Figure 8. The filter size is 3x3.

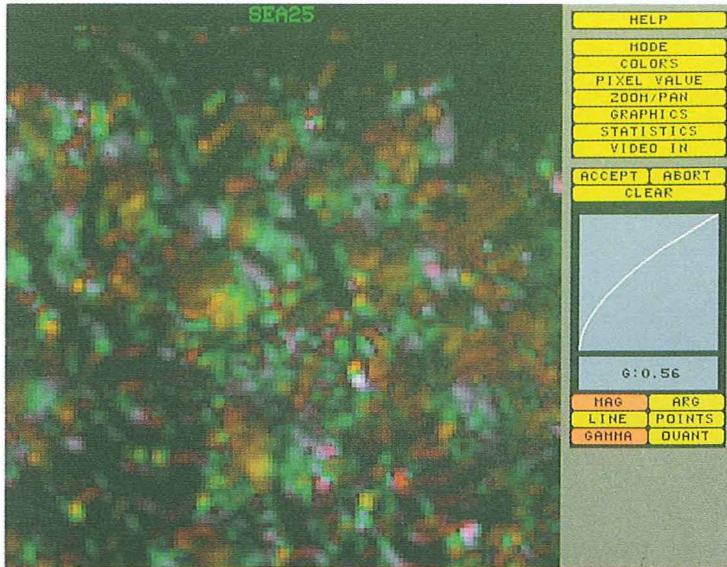


Figure 15: The image is a shrunk version of Figure 14. The image size is 128x128, but for the purpose of comparison it is photographed as a 512x512 image through repetition of pixels.

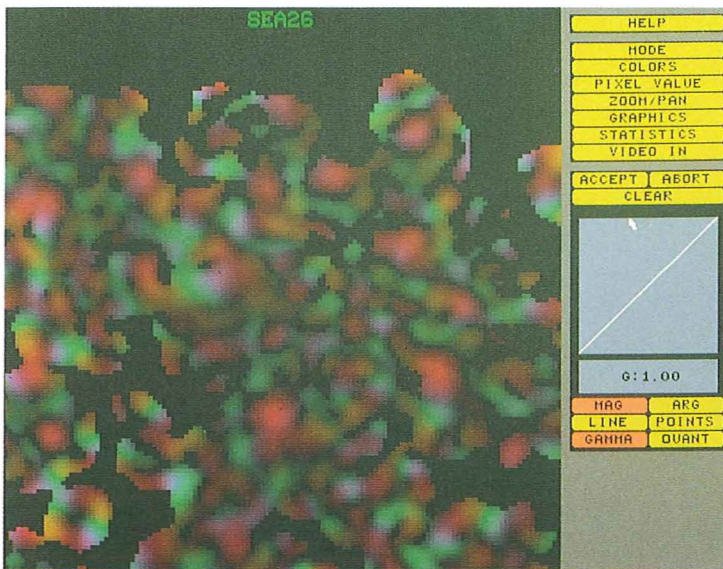


Figure 16: The image illustrates z_2 of Figure 8 based on the shrunk partial derivative image. The image size is 128x128, but it is expanded for comparison as before.

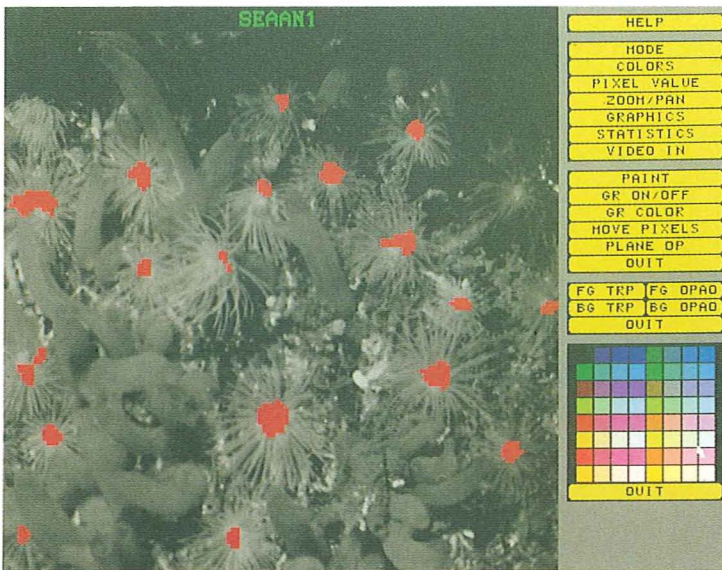


Figure 17: The image illustrates the identified sea anemones of Figure 8. Box classification is used. The used feature images were Figure 8, the intensity and hue of Figure 16. The used boxes were $[10, 255]$, $[85, 255]$, and $[113, 143]$.

A Some properties of the filter coefficients

In this appendix we will show that the filter coefficients v_j , given in this chapter involve a function which is given by

$$a(x) = \frac{2}{\pi} \int_0^\infty \int_0^\pi \exp(-r^2 + xr \cos \varphi - \frac{x^2}{4}) r \cos(2\varphi) dr d\varphi \quad (1)$$

and will prove some of its properties. By formula (32) we obtain:

$$\begin{aligned} v_j & (2\pi)^{-2} \exp(\beta r_j^2) = \\ &= \int_0^\infty \int_0^{2\pi} \exp(-i2\varphi) \exp(-(\alpha + \beta)r^2 + 2\beta r_j r \cos(\varphi - \varphi_j)) r dr d\varphi \\ &= \int_0^\infty \int_0^{2\pi} \exp(-(\alpha + \beta)r^2 + 2\beta r_j r \cos \varphi) r \exp(-i2(\varphi + \varphi_j)) dr d\varphi \\ &= \exp(-i2\varphi_j) \left[\int_0^\infty \int_0^{2\pi} \exp(-(\alpha + \beta)r^2 + 2\beta r_j r \cos \varphi) r \cos(2\varphi) dr d\varphi \right. \\ &\quad \left. - i \int_0^\infty \int_0^{2\pi} \exp(-(\alpha + \beta)r^2 + 2\beta r_j r \cos \varphi) r \sin(2\varphi) dr d\varphi \right] \end{aligned}$$

Since the sine function is odd and periodic the second integral vanishes and we obtain:

$$\begin{aligned} v_j & (2\pi)^{-2} \exp(\beta r_j^2 + i2\varphi_j) \\ &= \int_0^\infty \int_0^{2\pi} \exp(-(\alpha + \beta)r^2 + 2\beta r_j r \cos \varphi) r \cos(2\varphi) dr d\varphi \\ &= (\alpha + \beta)^{-1} \int_0^\infty \int_0^{2\pi} \exp(-r^2 + \frac{2\beta r_j}{\sqrt{\alpha + \beta}} r \cos \varphi) r \cos(2\varphi) dr d\varphi. \end{aligned}$$

Adding and subtracting $\frac{\beta^2 r_j^2}{\alpha + \beta}$ to the argument of the exponential function term in the integrand provides

$$\begin{aligned} v_j &= 4\pi^3 (\alpha + \beta)^{-1} \exp(-\frac{\beta\alpha}{\beta + \alpha} r_j^2) \exp(-i2\varphi_j) \frac{1}{\pi} \\ &\quad \int_0^\infty \int_0^{2\pi} \exp(-r^2 + \frac{2\beta r_j}{\sqrt{\alpha + \beta}} r \cos \varphi - \frac{(\frac{2\beta r_j}{\sqrt{\alpha + \beta}})^2}{4}) r \cos(2\varphi) dr d\varphi \\ &= 4\pi^3 (\alpha + \beta)^{-1} \exp(-\frac{\beta\alpha}{\beta + \alpha} r_j^2) \exp(-i2\varphi_j) a(\frac{2\beta r_j}{\sqrt{\alpha + \beta}}) \end{aligned}$$

since the cosine function is even.

Now we will show that $a(x)$ is even and

$$\lim_{x \rightarrow \infty} a(x) = 1. \quad (2)$$

$$a(-x) = \frac{2}{\pi} \int_0^\infty \int_0^\pi \exp(-r^2 - xr \cos \varphi - \frac{x^2}{4}) r \cos(2\varphi) dr d\varphi$$

$$\begin{aligned}
&= \frac{2}{\pi} \int_0^\infty \int_0^\pi \exp(-r^2 + xr \cos(\varphi - \pi) - \frac{x^2}{4}) r \cos(2\varphi) dr d\varphi \\
&= \frac{2}{\pi} \int_0^\infty \int_{-\pi}^0 \exp(-r^2 + xr \cos(\varphi) - \frac{x^2}{4}) r \cos(2\varphi) dr d\varphi \\
&= \frac{2}{\pi} \int_0^\infty \int_0^\pi \exp(-r^2 + xr \cos \varphi - \frac{x^2}{4}) r \cos(2\varphi) dr d\varphi = a(x).
\end{aligned}$$

Consider the function $f(xr)$ which we will define as one of the integrals in (1):

$$\begin{aligned}
f(xr) &= \frac{2}{\pi} \int_0^\pi \exp(xr \cos \varphi) \cos(2\varphi) d\varphi \\
&= \frac{2}{\pi} \int_0^\pi \cosh(xr \cos \varphi) \cos(2\varphi) d\varphi
\end{aligned}$$

since $f(xr)$ is even in analogy with $a(x)$. Repeated partial integration gives:

$$\begin{aligned}
f(xr) &= \frac{2}{2\pi} \int_0^\pi xr \sin \varphi \sinh(xr \cos \varphi) \sin(2\varphi) d\varphi \\
&= \frac{1}{\pi} \int_0^\pi x^2 r^2 \sin \varphi \cosh(xr \cos \varphi) \frac{2}{3} \sin^3 \varphi d\varphi \\
&= \frac{2}{3\pi} \int_{-1}^1 x^2 r^2 \cosh(xrt) (1-t^2)^{3/2} dt \\
&= \frac{(xr/2)^2}{\sqrt{\pi} \Gamma(2+1/2)} \int_{-1}^1 (1-t^2)^{2-1/2} \cosh(xrt) dt \\
&= 2I_2(xr)
\end{aligned} \tag{3}$$

where $I_2(z)$ is the modified Bessel function of the first kind of order 2.

Thus

$$a(x) = 2 \int_0^\infty \exp(-r^2 - \frac{x^2}{4}) I_2(xr) r dr \tag{4}$$

Define $q(z)$ as

$$q(z) = \frac{I_2(z)}{\exp(z)(2\pi z)^{-1/2}}. \tag{5}$$

we will utilize

$$\lim_{z \rightarrow \infty} q(z) = 1 \tag{6}$$

to prove (2). Because of (6) there exists a constant w above which $q(z)$ is bounded above, that is

$$q(z) \leq M_1 < \infty \quad \text{for } z > w. \tag{7}$$

Formula (3) reveals that $I_2(z)$ is even and strictly increases for positive values of z . The asymptotic equivalent function $\exp(z)(2\pi z)^{-1/2}$ is bounded below with a positive lower bound. Thus for non-negative z values less than w , $q(z)$ is bounded above as well. This provides us:

$$q(z) \leq M < \infty \quad \text{for } z \geq 0. \tag{8}$$

Utilizing (4) and (5)

$$\begin{aligned}
 \frac{a(x)}{2} &= \int_0^{\infty} \exp\left(-\left(r - \frac{x}{2}\right)^2\right) q(xr) r (2\pi xr)^{-1/2} dr \\
 &= (2\pi)^{-1/2} \int_{-\frac{x}{2}}^{\infty} \exp(-r^2) q\left(x\left(r + \frac{x}{2}\right)\right) \left(\frac{r + \frac{x}{2}}{x}\right)^{1/2} dr \\
 &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp(-r^2) q\left(x\left(r + \frac{x}{2}\right)\right) \sigma\left(r + \frac{x}{2}\right) \left|\sqrt{\frac{r}{x} + \frac{1}{2}}\right| dr \\
 &\leq (2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp(-r^2) (r^2 + 2) M dr < \infty
 \end{aligned}$$

is obtained for $x > 1$ where $\sigma(z)$ is given by:

$$\sigma(z) = \begin{cases} 1, & \text{if } z \geq 0; \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Thus the dominated convergence theorem [12], can be used to evaluate the limit of $a(x)$:

$$\begin{aligned}
 \lim_{z \rightarrow \infty} a(x) &= \lim_{z \rightarrow \infty} \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-r^2) q\left(x\left(r + \frac{x}{2}\right)\right) \sigma\left(r + \frac{x}{2}\right) \left|\sqrt{\frac{r}{x} + \frac{1}{2}}\right| dr \\
 &= \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \lim_{z \rightarrow \infty} \exp(-r^2) q\left(x\left(r + \frac{x}{2}\right)\right) \sigma\left(r + \frac{x}{2}\right) \left|\sqrt{\frac{r}{x} + \frac{1}{2}}\right| dr \\
 &= \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\infty} \exp(-r^2) \sqrt{\frac{1}{2}} dr = 1
 \end{aligned}$$

References

- [1] R. L. Wheeden, A Zygmund, "Measure And Integral" Marcel Dekker, Inc., Basel, 1977.

B Central symmetry modelling

This appendix considers the theoretical aspects of the case when a neighbourhood is defined by means of a circle and the local coordinate transformation is chosen as

$$\begin{aligned}x &= r \cos \varphi \\y &= r \sin \varphi\end{aligned}$$

The local coordinate transformation is not defined by a pair of harmonic functions, a concept which will be studied closer in Chapter IV. It should be observed that this does not affect the form of the theoretical solution as such. The appendix is included as a reference for this reason but also to demonstrate that the behaviour of the special Fourier domain corresponding to a new coordinate system can be studied with respect to point concentration as well as line concentration. The latter is the type of the behaviour which characterizes the symmetry definition in different coordinate systems appearing throughout this thesis. During the implementation and the experimental results, the superiority of the smooth window function and the harmonic pair given in the chapter is verified.

CENTRAL SYMMETRY MODELING[†]

Josef Bigün, Gösta H. Granlund
Computer Vision Laboratory
Linköping University
Department of Electrical Engineering
S-581 83 Linköping Sweden

Abstract

A definition of central symmetry for local neighbourhoods of 2-D images is given. A complete ON-set of centrally symmetric basis functions is proposed. The local neighbourhoods are expanded in this basis. The behaviour of coefficient spectrum obtained by this expansion is proposed as the foundation of the central symmetry parameters of the neighbourhoods. Specifically, examination of two such behaviours are proposed: point concentration and line concentration of the energy spectrum. Moreover, the study of these types of behaviors of the spectrum is shown to be possible in the spatial domain.

Introduction

There is a long list of operators that detect the existence of linear symmetry in a local neighbourhood. Most of them measure linear symmetry in the sense of lines and edges. But there is very little done to model central symmetry. Perhaps it is because images of objects in nature, are usually more irregular than circles. Nevertheless, we believe that this is one of the symmetries which human beings utilize in early vision. It seems that central symmetry should be an additional symmetry model. The fact that circularly symmetric shapes like rotating fans, diverging rays, circularly propagating water waves, etc., are observed as phosphenes when low frequency magnetic fields are applied to the temples of a subject, [1], [2], supports this belief. Moreover many manufactured objects are locally concentrated and have closed rounded boundaries. Many natural objects in low resolution images may exhibit this property, like cells seen under a microscope. Conceivable application areas are the object counting, classification as well as image coding and enhancement of certain types of images, possessing a local central symmetry property. But first we should have an intuitive feeling about what kind of patterns are called centrally symmetric in our terminology, since it is otherwise a rather vague concept.

[†]SIGNAL PROCESSING III: Theories and Applications, I.T. Young et al. (editors) Elsevier Science Publishers B.V. (North-Holland), EURASIP, 1986; pp. 883-886.

Definition 1 We will call local neighbourhoods centrally symmetric if the locus of iso-gray values constitutes parallel lines in local polar coordinates:

$$\varphi = k_1 r + k_2 \quad r \geq 0$$

for some constants k_1 and k_2 . if the locus of iso-gray values is not curves, that is when they are regions then the borders of these regions are considered as locus. We will assume that the boundary of the neighbourhood is a circle, and the origin of coordinates is the center of this circle.

Definition 2 (f, g) is the scalar product given by

$$(f, g) \triangleq \frac{1}{|\Omega|} \int_{\Omega} \frac{1}{r} f^*(\bar{r}) g(\bar{r}) d\Omega$$

with $r = |\bar{r}|$ and:

$$|\Omega| = \int_{\Omega} \frac{1}{r} d\Omega.$$

Definition 3 $\mathcal{C}(\Omega)$ is the completion of the space of complex valued functions, continuous on Ω except on a subset of Ω with zero measure, with respect to the scalar product given above. Ω is a circle with the radius R .

Consider the functions, see Figure 1),

$$\Psi_{mn}(\bar{r}) = e^{i(m\omega r + n\varphi)} \quad (1)$$

with $\omega = \frac{2\pi}{R}$ and $m, n \in \mathbb{Z}$. $\mathcal{C}(\Omega)$ is a Hilbert space with the following scalar product

$$\frac{1}{2\pi R} \int_0^{2\pi} \int_0^R \bar{f}(r, \varphi) g(r, \varphi) dr d\varphi.$$

Thus $\{\Psi_{mn}\}_{m,n \in \mathbb{Z}}$ is dense in $\mathcal{C}(\Omega)$, which follows from the Fourier series expansion theory on a rectangle, [3]. But this scalar product is the same scalar product defined earlier with Ω being a circle. Now let us consider the neighbourhood Ω , around an examined point in an image. Assume that the polar coordinates, $r = |\bar{r}|$ and $\varphi = \arg(\bar{r})$, referred to in the following are relative to the examined point, and the positive x axis from the examined point.

Let the real function $f(\bar{r})$ express the gray values in Ω , with the center at the examined point, so that \bar{r} is the local coordinate vector. Then one can expand f as

$$f(\bar{r}) = \sum_{m,n \in \mathbb{Z}} c_{mn} \Psi_{mn}(\bar{r}) \quad (2)$$

with

$$\begin{aligned} c_{mn} &= (f, \Psi_{mn}) \\ &= \frac{1}{2\pi R} \int_{\Omega} \frac{1}{r} f(\bar{r}) e^{i(m\omega r + n\varphi)} d\Omega \end{aligned} \quad (3)$$

because $\mathcal{C}(\Omega)$ is a Hilbert space and $\{\Psi_{mn}\}_{m,n \in \mathbb{Z}}$ constitutes a complete orthonormal base:

$$(\Psi_{mn}, \Psi_{m'n'}) = \delta_{mm'} \delta_{nn'} \quad (4)$$

with $\delta_{mm'}$ being the usual Kronecker delta.

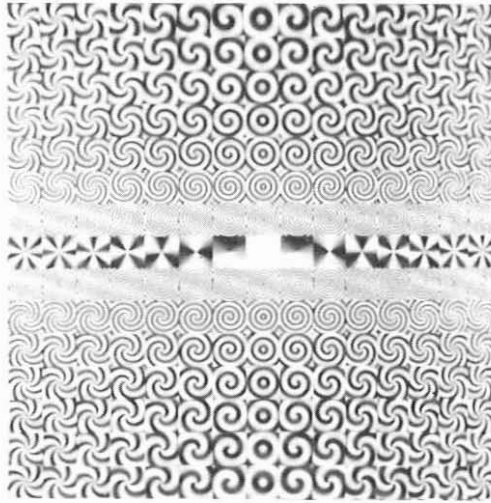


Figure 1: The image illustrates some of the basis functions Ψ_{mn} . The real parts of Ψ_{mn} are mapped linearly to the gray values of the monitor.

Point concentration

Definition 4 Let P be an operator from $C(\Omega)$ to the function set X , $X \subset C(\Omega)$. Then P is a projection from $C(\Omega)$ to X if

$$P^2 f = P f$$

for all f in $C(\Omega)$.

Our goal is to find an algorithm based on operations done in the spatial domain which still gives some indication about whether the energy is concentrated to a point in the frequency domain. The algorithm should possess the following properties:

1. Whenever the neighbourhood, $f(\bar{\tau})$, is equivalent to one of the basis functions, Ψ_{mn} , except possibly for a scale factor B , the algorithm should detect this particular basis function save a sign change of its index tuple, (n, m) . That is $(f, \Psi_{m'n'}) = 0$ for all tuples (n', m') except for a tuple (n, m) . In other cases it should give some sort of dominating tuple (n, m) . It should be noted that Ψ_{mn} is complex valued. For real neighbourhoods consisting of the real or imaginary part of a Ψ_{mn} , this condition will be enough to identify the neighbourhood except possibly a phase factor. Given the tuple (n, m) , Ψ_{mn} is unique. Call the operator of finding the tuple (n, m) , and associating the function Ψ_{mn} to that, as P then:

$$P^2 f = P f = \Psi_{mn}$$

for any $f \in C(\Omega)$. This is equivalent to saying that the sought algorithm is a projection to the countable set $\{\Psi_{mn}\}$, according to the projection definition above.

2. The projection value (or parameter) should be rotation and radial phase invariant for pure inputs of:

$$f(r, \varphi) = B\Psi_{mn}(r, \varphi)$$

with some scalar B . That is

$$Pf(r + r_0, \varphi + \varphi_0) = Pf(r, \varphi) = \Psi_{mn}$$

should be fulfilled.

3. Whenever the spectrum of the real valued local neighbourhood differs from a pattern with a point concentrated spectrum, an uncertainty parameter should reflect that. By attaining low values, for example, this parameter could indicate the relevance of the projection parameter, and conversely to suppress it if the neighbourhood differs from a central symmetric pattern.

The use of the uncertainty parameters is indicated in [4]. The uncertainty parameter and the projection parameter are combined in every point of the image to form a vector, in such a way that the magnitude of this vector becomes inversely proportional to the uncertainty parameter, (the confidence in the projection parameter) and the argument of it becomes the projection parameter. This can be visualised by allowing the magnitude to modulate the intensity of a point in a colour TV monitor and the argument of it, representing the projection parameter, to modulate the colour of the same point. The result is a colour image representing a decision in every neighbourhood of the original image. The projection parameter and the confidence parameter values evaluated in every point in the image can be thought of as two separate images influencing each other. A point with a low confidence level looks dark in the resulting image, no matter what the colour of the point is. A point with a high confidence level is emphasized by illumination, and it's colour is revealed.

The algorithm we propose consists of the computations described by (5)-(9):

$$A \triangleq \|f\| \tag{5}$$

$$m_d \triangleq \frac{\|D_r f\|}{A\omega} \tag{6}$$

$$n_d \triangleq \frac{\|D_\varphi f\|}{A} \tag{7}$$

$$C_{\Omega m}^2 \triangleq \frac{\|D_r^2 f\|^2}{A^2\omega^4} - m_d^4 \tag{8}$$

$$C_{\Omega n}^2 \triangleq \frac{\|D_\varphi^2 f\|^2}{A^2} - n_d^4 \tag{9}$$

m_d and n_d are radial respectively angular frequency measures. $C_{\Omega m}$ and $C_{\Omega n}$ are the uncertainty measures associated with m_d respectively n_d . Denote the projection parameters by the tuple (\hat{n}, \hat{m}) . That is (\hat{n}, \hat{m}) points out the location of an eventual point concentration in the spectrum. To produce \hat{m} and \hat{n} from m_d and n_d , we observe that \hat{m} and \hat{n} should be integers. Moreover they take positive as well as negative values. However since we assume real valued images, the requested point concentration will consist of two concentrations symmetrically located around the origin of the coordinates in the spectrum. This is due to the Hermitian property of the coefficient transformation. Hence we need only give the position of one of these concentrations. Thus we can assume that \hat{m} is always positive. We will simply assign to \hat{m} and \hat{n} the closest integers to m_d and n_d with the proper sign:

$$\begin{aligned}\hat{m} &= \text{round}(m_d) \\ \hat{n} &= \text{sign} \times \text{round}(n_d) \quad \text{sign} \in \{-1, 1\}\end{aligned}\quad (10)$$

sign is the sign of $\sum_{m,n \in \mathbb{Z}} mn \frac{|c_{mn}|^2}{A^2}$, the calculation of which is given in the next section. Let us see what (5)-(9) does for a neighbourhood :

$$f = \sum_{m,n \in \mathbb{Z}} c_{mn} \Psi_{mn} \quad (11)$$

We get through (5)

$$A^2 = \sum_{m,n \in \mathbb{Z}} |c_{mn}|^2$$

This is the energy of the neighbourhood in terms of the centrally symmetric function set $\{\Psi_{mn}\}$. (4) together with (6), (11) yields:

$$\begin{aligned}m_d &= \frac{\|D_r f\|}{A\omega} = \left\| \sum_{m,n \in \mathbb{Z}} i \frac{m c_{mn}}{A} \Psi_{mn} \right\| \\ &= \left(\sum_{m,n \in \mathbb{Z}} \frac{c_{mn}^2}{A^2} m^2 \right)^{\frac{1}{2}}\end{aligned}\quad (12)$$

Hence m_d is the weighted root mean square of all radial frequency measures, m . It should be observed that a particular radial frequency number, m , is weighted by the uniform sum of all angular frequency energies. The weights constitute energy distribution of the input function. The higher the energy share of Ψ_{mn} in the total energy, the closer m_d will be to m . (12) obviously fulfills the projection requirement after rounding m_d to the closest integer, \hat{m} . Similarly n_d will be the weighted mean square of all angular frequencies of different order:

$$n_d = \left(\sum_{m,n \in \mathbb{Z}} \frac{|c_{mn}|^2}{A^2} n^2 \right)^{\frac{1}{2}} \quad (13)$$

The latter is insensitive to the sign changes in n . The consequence of this is a real neighbourhood of

$$f = \Psi_{mn} + \Psi_{-m-n} + \Psi_{m-n} + \Psi_{-mn},$$

projected to a $\Psi_{|m||n|}$ input. The decision is in favour of one of the two equally strong candidates. When $f = \Psi_{mn} + \Psi_{-m-n}$ then $m_d = |m|$ and $n_d = |n|$ which in turn reflects the necessity of the variable *sign* referred to earlier (10). Uncertainty parameters C_{Ω_m} and C_{Ω_n} are proposed to be as in (8) and (9), and C_{Ω_m} yields through (4), (8), (11), (12)

$$\begin{aligned}
C_{\Omega_m}^2 &= \frac{\|D_r^2 f\|^2}{A^2 \omega^4} - m_d^4 = \sum_{m,n \in Z} \frac{|c_{mn}|^2}{A^2} m^4 \\
&- 2 \sum_{m,n \in Z} \frac{|c_{mn}|^2}{A^2} m^2 m_d^2 + m_d^4 \sum_{m,n \in Z} \frac{|c_{mn}|^2}{A^2} \\
&= \sum_{m,n \in Z} \frac{|c_{mn}|^2}{A^2} (m^2 - m_d^2)^2
\end{aligned} \tag{14}$$

which can be viewed as a weighted variance for the integers m^2 . It attains its minimum in the case when

$$\frac{|c_{mn}|^2}{A^2} (m^2 - m_d^2)^2 = 0$$

for all $n, m \in Z$. This occurs if and only if

$$\sum_n \frac{|c_{mn}|^2}{A^2} = 1$$

for some $m = m'$ since m_d^2 is constant. Thus if C_{Ω_m} is zero then there exists one unique radial frequency in the neighbourhood and it is given by the estimation, m_d . When this is the case the energy is concentrated to a horizontal line through $m = m_d$. Since C_{Ω_m} is a variance it also reveals some information about the shape of the spectral density of the neighbourhood. If C_{Ω_m} is small then it is likely to think that the neighbourhood is a degraded version of a wave with a well defined radial frequency, m_d . Conversely it is unlikely that the association of m_d to the neighbourhood will be relevant, if C_{Ω_m} is large. Interpretations of n_d and C_{Ω_n} are similar to m_d and C_{Ω_m} 's. Given m_d, n_d , and the *sign* parameters the tuple (\hat{n}, \hat{m}) is computed according to (10). We adopt the uncertainty parameters C_{Ω_n} and C_{Ω_m} for \hat{n} respectively \hat{m} . We propose C_{Ω_ν} ,

$$C_{\Omega_\nu}^2 = C_{\Omega_m}^2 + C_{\Omega_n}^2 \tag{15}$$

to be the uncertainty parameter for the tuple (\hat{n}, \hat{m}) . $C_{\Omega_\nu} = 0$ if and only if $C_{\Omega_m} = C_{\Omega_n} = 0$. But $C_{\Omega_m} = 0$ if and only if the total energy is concentrated on a horizontal line and $C_{\Omega_n} = 0$ if and only if the total energy is concentrated on a vertical line. The only possibility for the neighbourhood to fulfill these two requirements is when an input possessing total point concentration in its spectrum, with location on the intersection of the lines $m = m_d, n = n_d$.

Line concentration

We will examine whether the energy spectrum has line concentration. Let the line we look for be

$$m = \tan(\theta)n. \quad (16)$$

We assume that the line goes through the origin of the coordinates in the coefficient domain. Since the real functions coefficient transforms should be Hermitian, their energy spectra are even, forcing a possible line concentration to pass through the origin of the coordinates of the coefficient plane. A real neighbourhood f can be expanded in the basis functions as before, yielding:

$$f = \sum_{m,n \in \mathbb{Z}} c_{mn} \Psi_{mn}$$

with the Hermitian coefficients c_{mn} . The energy concentration of the neighbourhoods in general degrades from a line through the origin of the coordinates. Let us measure this degradation by $C_{\Omega\theta}$, which is the average sum of the squares of the distances of the spectrum points to the line given by (16):

$$\begin{aligned} C_{\Omega\theta} &\triangleq \sum_{m,n \in \mathbb{Z}} (m - \tan(\theta)n)^2 \cos^2(\theta) \frac{|c_{mn}|^2}{A^2} \\ &= \sin^2(\theta) \sum_{m,n \in \mathbb{Z}} n^2 \frac{|c_{mn}|^2}{A^2} + \cos^2(\theta) \sum_{m,n \in \mathbb{Z}} m^2 \frac{|c_{mn}|^2}{A^2} \\ &\quad - \sin(2\theta) \sum_{m,n \in \mathbb{Z}} mn \frac{|c_{mn}|^2}{A^2} \end{aligned} \quad (17)$$

We want to find a θ which minimizes $C_{\Omega\theta}$. This is the least square estimation of θ and it is straightforward to find θ :

$$\frac{dC_{\Omega\theta}}{d\theta} = (n_d^2 - m_d^2) \sin(2\theta) - 2p \cos(2\theta) \quad (18)$$

where

$$p = \sum_{m,n \in \mathbb{Z}} mn \frac{|c_{mn}|^2}{A^2}$$

If $n_d^2 - m_d^2 \neq 0$ or $p \neq 0$ then choose the minimizing θ as:

$$\theta_d = \frac{1}{2} \tan^{-1}(n_d^2 - m_d^2, 2p). \quad (19)$$

The degradation or uncertainty measure is given by substituting (19) in (18) and using the trigonometric half angle formulas:

$$C_{\Omega\theta} = \frac{1}{2}(n_d^2 + m_d^2 - \sqrt{(n_d^2 - m_d^2)^2 + 4p^2}). \quad (20)$$

The angle given by (19) gives the axis around which the moment of inertia is minimum and the moment of inertia is given by (20) if $\frac{|c_{mn}|^2}{A^2}$ is seen as a point mass, [5]. The omitted case when both $p = 0$ and $n_d^2 - m_d^2 = 0$ corresponds to local neighbourhoods with no specific orientation. Because $\frac{dC_{\Omega\theta}}{d\theta}$ vanishes according to (16), any θ would work as minimizing argument to (17). This case implies that

$$\begin{aligned}\sum_{m,n \in Z} m^2 \frac{|c_{mn}|^2}{A^2} &= \sum_{m,n \in Z} n^2 \frac{|c_{mn}|^2}{A^2} \\ \sum_{m,n \in Z} mn \frac{|c_{mn}|^2}{A^2} &= 0\end{aligned}$$

The class of functions having this property in their spectra is the class of functions with coinciding principal axes in the coefficient domain. Neighbourhoods of Ψ_{00} , $\Psi_{mn} + \Psi_{-m-n} + \Psi_{m-n} + \Psi_{-mn}$ are examples of such functions. We observe that $m_d^2 = 0$ and $n_d^2 = 0$ implies that the neighbourhood is a constant function and consequently has no orientation. Thus to keep the consistency of the meaning of $C_{\Omega\theta}$ in the case when $n_d^2 - m_d^2 = p = 0$, we should define $C_{\Omega\theta} = \infty$ and leave θ undefined. p which is needed to calculate θ_d and $C_{\Omega\theta}$ according to (19) and (20), can be easily found in the spatial domain to be:

$$p = \sum_{m,n \in Z} mn \frac{|c_{mn}|^2}{A^2} = \frac{1}{A^2 \omega} \left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \varphi} \right)$$

Implementation of the scalar products given above for every neighbourhood of a digitized image is straightforward after using the chain rule:

$$\begin{aligned}\frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x} \cos(\varphi) + \frac{\partial f}{\partial y} \sin(\varphi) \\ \frac{\partial f}{\partial \varphi} &= \frac{\partial f}{\partial y} r \cos(\varphi) - \frac{\partial f}{\partial x} r \sin(\varphi).\end{aligned}$$

This means that we can transfer the scalar products to be valid for functions defined in Cartesian coordinates. At this point we can use either the bandlimited signal theory or some quadrature rule to evaluate the resulting integrals, given that we know $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at a rectangular net of points. It can be shown that the scalar product evaluations at every point are obtained by convolutions with FIR-filters.

Conclusion

Both the point concentration parameters m_d , and n_d and the line concentration parameter θ_d are best fits of a point or a line respectively through the origin of coordinates of the coefficient domain. The best fit is in the sense that the two variance measures given, which are adopted as uncertainty measures, are minimized. It is interesting to note that the approach lends itself to linear symmetry parameter extraction as well, with a minor change. By linear symmetry we mean the neighbourhoods with iso-gray values being

straight lines in Cartesian coordinates. Parallel lines belong to such neighbourhoods. Hence it is possible to find the dominating frequency and the dominating orientation of a neighbourhood, [6], with the least error variance in the Fourier domain in a similar manner. The only difference is the scalar product and the shape of the neighbourhood. The scalar product of the linear symmetry case becomes the usual $\mathcal{L}^2(\Omega)$ scalar product with Ω being a rectangle. The complete ON-basis set is of course $\{e^{i(m\omega_x x + n\omega_y y)}\}_{m,n \in \mathbb{Z}}$.

References

- [1] Max Knoll, Johann Kugler, Joseph Eichmeier and Oskar Höfer: "Note on the Spectroscopy of Subjective Light Patterns" *Journal of Analytical psychology*, No 7, 1962,
- [2] Per Lövsund: "Biological Effects of Alternating Magnetic Fields with Special Reference to the Visual System" Dissertation No 47, 1980, Linköping Studies in Science and Technology, Linköping University, Sweden.
- [3] W. Rudin: "Real and Complex Analysis" McGraw Hill, 1969.
- [4] G.H. Granlund: "Hierarchical Image Processing" Proceedings of SPIE Technical conference, Geneva, April 18-27, 1983.
- [5] I.L. Meriam: "Statics" John Wiley & Sons, New York, 1980.
- [6] H. Knutsson: "Filtering and reconstruction in image processing" Dissertation No. 88, 1982, Linköpings Studies in Science and Technology, Linköping University, Sweden.
- [7] Robert A. Hummel: "Feature Detection Using Basis Functions" *Computer Graphics and Image Processing* 9, 1979.
- [8] P.E. Danielsson: "Rotation invariant linear operators with directional response" Proceedings of 5'th international conference on pattern recognition, December 1980.

Chapter III

OPTIMAL ORIENTATION DETECTION OF LINEAR SYMMETRY

What has been said about circles also applies to lines as well as to colour...

Plato

OPTIMAL ORIENTATION DETECTION OF LINEAR SYMMETRY*

Josef Bigün, Gösta H. Granlund
Linköping University
Department of Electrical Engineering
Computer Vision Laboratory
S-581 83 Linköping Sweden

Abstract

The problem of optimal detection of orientation in arbitrary neighbourhoods is solved in the least squares sense. It is shown that this corresponds to fitting an axis in the Fourier domain of the n -dimensional neighbourhood, the solution of which is a well known solution of a matrix eigenvalue problem. The eigenvalues are the variance or inertia with respect to the axes given by their respective eigenvectors. The orientation is taken as the axis given by the least eigenvalue. Moreover it is shown that the necessary computations can be performed in the spatial domain without doing a Fourier transformation. An implementation for 2-D is presented. Two certainty measures are given, corresponding to the orientation estimate. These are the relative or the absolute distances between the two eigenvalues, stating whether the fitted axis is much better than an axis orthogonal to it. The result of the implementation is verified by experiments which confirm an accurate orientation estimation and reliable certainty measure in the presence of additive noise at high as well as low level.

1 Introduction

The problem of orientation detection of lines and edges arises in many applications in image processing. One of the earliest approaches was to model the direction of a neighbourhood in terms of the direction of the gradient of the image. A drawback of this method is its noise amplification since the gradient operation enhances high frequencies of the image. Another approach is to combine linearly the magnitudes of a number, 3 or 4, of quadrature, directional filters, [1]. The coefficients in this linear functional are complex valued as well as those of the filters. This results in a complex valued variable, the argument of which is an estimate of the orientation of the local neighbourhood and

*Reprinted from PROCEEDINGS OF THE IEEE FIRST INTERNATIONAL CONFERENCE ON COMPUTER VISION, London, June 8-11, 1987 pp. 433-438.

the magnitude is an estimate of the certainty of this orientation estimation. There have also been solutions to the problem of finding the local orientation by projecting the neighbourhood to a number of fixed orthogonal functions. The projection coefficients are then used to evaluate the orientation parameter of the model [2,3].

We will propose a new approach for local orientation detection which is based on the well-known solution of the principal axis problem of rigid bodies in mechanics, but applied in the Fourier domain [5, 6,10]. In Section 1, we will define linear symmetry and describe the method for an n -dimensional Euclidean space. In Section 2 we will apply the results for 2-dimensional images and in Section 3 the experiments and results for the 2-D case will be presented. We predict that the experimental results of this approach for the 3-D case should be similar to those for 2-D. Since lines and edges are linear symmetric structures, this method can be used for detection of these structures by means of the certainty parameters introduced in Section 2, as well as orientation estimation in applications.

2 Orientation detection in n -dimensional euclidean space.

Let E_n be the Euclidean space with dimension n .

Definition 2 We will call a non-negative and bounded function f with real or complex values defined on E_n an image, and the values of f the gray values of the image. Further we will call an image linearly symmetric if the isogray values constitute parallel hyperplanes of dimension $n - 1$. That is if the image f can be expressed by a function g defined on E_1 for some vector $\bar{k} \in E_n$ as $f(\bar{r}) = g(\bar{k} \cdot \bar{r})$ for all $\bar{r} \in E_n$.

Theorem 1 A linear symmetric image has a Fourier transform concentrated to a line through the origin:

$$\mathcal{F}_n(f(\bar{k}_0^t \bar{r}))(\bar{s}) = \mathcal{F}_1(f)(\bar{s}^t \bar{k}_0) \delta(\bar{s}^t \bar{u}_1) \delta(\bar{s}^t \bar{u}_2) \cdots \delta(\bar{s}^t \bar{u}_{n-1})$$

$\bar{k}_0, \bar{u}_1, \dots, \bar{u}_{n-1}$ are orthonormal, and δ is the dirac distribution.

Proof: Decompose E_n in E_1 and E_{n-1}

$$\bar{r} = t\bar{k}_0 + u_2\bar{u}_1 + u_2\bar{u}_2 \dots u_{n-1}\bar{u}_{n-1}$$

for all $\bar{r} \in E_n$, so that $\bar{k}_0, \bar{u}_1, \bar{u}_2 \dots \bar{u}_{n-1}$ are orthonormal.

$$\begin{aligned} \mathcal{F}_n(f(\bar{k}^t \bar{r})) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t) \exp(-j2\pi t \bar{s}^t \bar{k}_0) \times \\ &\exp(-j2\pi(\bar{s}^t \bar{u}_1, u_1 + \bar{s}^t \bar{u}_2 u_2 \dots \bar{s}^t \bar{u}_{n-1} u_{n-1})) dt du_1 \cdots du_{n-1} \end{aligned}$$

then the desired result follows immediately.

To detect linearly symmetric objects is consequently the same as to check the existence of energy concentration to a line in the Fourier domain. This theorem further

states that the function $f(\bar{k}_0^t \bar{r})$, which is in general a "spread" function, is compressed to a line. This is a property which supports the idea of checking the linear symmetry in the Fourier domain rather than in the spatial domain. But the Fourier transformation of every local neighbourhood is very cumbersome. We will show that the fitting of a straight line through the origin in the Fourier domain of an image with the least square error is possible to accomplish in the spatial domain. If we define the infinitesimal energy of the Fourier transform, $|\hat{f}(\bar{r})|^2 dE_n$, as the mass distribution, then we have the variance (or the inertia) with respect to the axis $t\bar{k}_0$

$$V_f^2 = \int_{\bar{r} \in E_n} d^2(\bar{r}, \bar{k}_0) dm(\bar{r}) \quad (1)$$

for the Fourier transform of the image, f . Here $d(\bar{r}, \bar{k}_0)$ is a real valued function which gives the Euclidean distance between the point \bar{r} and a candidate axis, defined by the axis $t\bar{k}_0$, where $\|\bar{k}_0\| = 1$. The problem is to find such an axis minimizing V_f^2 :

$$\min V_f^2 = \min \int_{E_n} d^2(\bar{r}, \bar{k}_0) |\hat{f}(\bar{r})|^2 dE_n \quad (2)$$

where dE_n is $dx_1 dx_2 \dots dx_n$ when $\bar{r} = x_1 \hat{x}_1 + \hat{x}_2 \dots x_n \hat{x}_n$ for all $\bar{r} \in E_n$. The distance function is given by the usual Euclidean distance:

$$\begin{aligned} d^2(\bar{r}, \bar{k}_0) &= (\bar{r} - (\bar{r}^t \bar{k}_0) \bar{k}_0)^t (\bar{r} - (\bar{r}^t \bar{k}_0) \bar{k}_0) \\ &= \bar{k}_0^t (I \bar{r}^t \bar{r} - \bar{r} \bar{r}^t) \bar{k}_0 \end{aligned}$$

where $\|\bar{k}_0\|^2 = \bar{k}_0^t \bar{k}_0 = 1$ is assumed. In combination with (1)

$$V_f^2 = \bar{k}_0^t \mathbf{J} \bar{k}_0 \quad (3)$$

is obtained with:

$$\mathbf{J} = \begin{pmatrix} J_{11} & -J_{12} & \dots & -J_{1n} \\ -J_{21} & J_{22} & \dots & -J_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -J_{n1} & -J_{n2} & \dots & J_{nn} \end{pmatrix}$$

where

$$J_{ii} = \int_{E_n} \sum_{j \neq i} x_j^2 |\hat{f}(\bar{r})|^2 dE_n \quad (4)$$

for diagonal elements and

$$J_{ij} = \int_{E_n} x_i x_j |\hat{f}(\bar{r})|^2 dE_n \quad (5)$$

for off-diagonal elements except for a sign change. The minimization problem formulated in (2) is solved by \bar{k}_0 corresponding to the least eigenvalue of the inertia matrix \mathbf{J} of the Fourier domain, [4]. All eigenvalues are real and the smallest eigenvalue corresponds to this minimum. This matrix contains sufficient information to allow computation of the optimal \bar{k}_0 in the sense given by (2)

Lemma 1 *The inertia matrix of the energy of the Fourier domain is possible to compute in the spatial domain by the following relation:*

$$J_{ii} = \frac{1}{4\pi^2} \sum_{j \neq i} \int_{E_n} \left(\frac{\partial f}{\partial x'_j} \right)^2 dE_n \quad (6)$$

for the diagonal elements and

$$- J_{ij} = - \frac{1}{4\pi^2} \int_{E_n} \frac{\partial f}{\partial x'_i} \frac{\partial f}{\partial x'_j} dE_n \quad (7)$$

for off-diagonal elements. Here x'_j is the spatial domain coordinate corresponding to the Fourier domain coordinate x_j and $dE_n = dx'_1 dx'_2 \dots dx'_n$

Proof of the lemma is immediate by applying the Parseval relation and the fact that a differentiation in the spatial domain corresponds to multiplication by the respective coordinate in the Fourier domain, (4) and (5).

Using the previous lemma we have the tools to find an optimal orientation of any image f . The obtained orientation will be unique if all the eigenvalues differ from the least eigenvalue. Moreover the variance given by (1) would be exactly zero if and only if f is a linear symmetric image. When the multiplicity of the least eigenvalue is larger than 1, there is no unique axis $t\bar{k}_0$, by which the image can be described as $g(\bar{k}_0^t \bar{r})$ for some one-dimensional function g . Instead, the energy in the Fourier domain is distributed in such a way that there are plenty of such axes which give the least square error. More exactly these axes are any axes given by a linear combination of the eigenvector space belonging to the least eigenvalue. Here it should be observed that the dimension of this space is equal to the multiplicity of the eigenvalue it corresponds to (the least one). This is due to the fact that \mathbf{J} is positive semi-definite and symmetric by definition, (1), (3), and (5). In other words, there is no unique and optimal axis passing through the origin but an optimal and unique hyperplane passing through it in the Fourier domain when the multiplicity of the least eigenvalue is greater than one. How shall we interpret the case when the least two eigenvalues of the inertia matrix \mathbf{J} are equal? If they are equal and vanish?

Since the matrix \mathbf{J} is symmetric and positive semi-definite, it can be diagonalized by a similarity transformation to \mathbf{J}'

$$\mathbf{J}' = \mathbf{P}^t \mathbf{J} \mathbf{P}$$

This corresponds to a rotation of the coordinate axes (both in the Fourier and the spatial domain). Let the coordinate axes of the spatial domain after rotation be $\hat{u}_1, \hat{u}_2 \dots \hat{u}_n$. Then by using lemma 1) and $J'_{00} = J'_{11} = \lambda_0$ we obtain

$$\int_{E_n} \left(\frac{\partial f}{\partial u_0} \right)^2 dE_n = \int_{E_n} \left(\frac{\partial f}{\partial u_1} \right)^2 dE_n \quad (8)$$

We will call such images perfectly balanced images with respect to \hat{u}_0 and \hat{u}_1 coordinate axes due to a similar definition in mechanics when $\lambda_0 = \lambda_1$. When $J'_{00} = J'_{11} = 0$ we

obtain through lemma 1)

$$\sum_{i \neq 0} \int_{E_n} \left(\frac{\partial f}{\partial u_i} \right)^2 dE_n = 0$$

$$\sum_{i \neq 1} \int_{E_n} \left(\frac{\partial f}{\partial u_i} \right)^2 dE_n = 0$$

Since all elements in the sums above are positive we have

$$\int_{E_n} \left(\frac{\partial f}{\partial u_i} \right)^2 dE_n = 0$$

for all allowed i , which in turn leads to the fact that

$$\frac{\partial f}{\partial u_i} = 0$$

for all allowed i . But this is the same as saying that f is a constant image. If only one eigenvalue is zero then we have

$$\frac{\partial f}{\partial u_i} = 0$$

for all i except the one corresponding to the zero eigenvalue. That is, it is constant in the directions perpendicular to the axis belonging to the least eigenvalue. Thus when two eigenvalues of the matrix \mathbf{J} are zero, so are all the others, and by this we have established the following lemma.

Lemma 2 *If one of the eigenvalues of the inertia matrix \mathbf{J} of the Fourier transform of the image f , has value zero then this eigenvalue has multiplicity of either 1 or n . It is equal to 1 if and only if we have a linearly symmetric image, and equal to n if and only if we have a constant image.*

To illustrate the concept of perfectly balanced images we have such a Fourier domain and its corresponding spatial domain in 2-D, Figure 1 a). It can, for this image, be shown that any axis through the origin in the Fourier domain will give the same least square error:

$$\sum_{i=1}^4 m_i d^2(\bar{r}_i, \bar{k}) = 2a^2 m^2$$

This is a perfectly balanced image. Thus the 2-D inertia matrix \mathbf{J} of this case has one eigenvalue $2a^2 m^2$ of multiplicity 2. The spatial domain corresponding to this Fourier domain consists of two planar waves (sinusoids) in the directions of the coordinate axes. Interpreting Figure 1 a) in 3 dimensions would give a hyperplane as in Figure 1 b). In the 3-D spatial domain this corresponds to an image as the one in Figure 1 a), in every 2-D plane perpendicular to the axis defined by the vector, \hat{x}_3 . In Figure 1 a) we obtain both eigenvalues equal, namely $2m^2 a^2$, while in the 3-D case, Figure 1 b), we obtain eigenvalues $2m^2 a^2$, $2m^2 a^2$ and $4m^2 a^2$. Thus we can say that the image is perfectly balanced with respect to any axis perpendicular to \hat{x}_3 . One is tempted to infer that

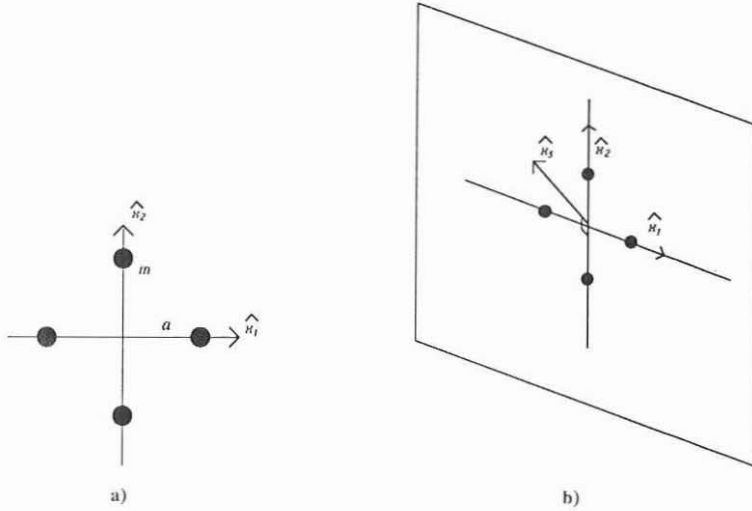


Figure 1: a) and b) illustrate perfectly balanced images in the Fourier domain for 2-D and 3-D respectively. m is energy. A spatial domain image corresponding to both a) and b) is given by: $A \sin(2\pi a u_1) + A \sin(2\pi a u_2)$, where A is a constant and u_1 and u_2 are spatial domain coordinates corresponding to \hat{x}_1 and \hat{x}_2 . It should be observed that the spatial domain image is constant for all u_3 .

the image should have its iso-gray values as parallel lines, when the least eigenvalue has multiplicity 2 for the 3-D case. However, this is not always true. A counter-example is when we have equal masses (energies) at the Fourier sites: $(\pm 1, \pm 1, 0)$, $(0, 0, \pm \frac{1}{2})$. The eigenvalues are proportional to $\frac{5}{2}$, $\frac{5}{2}$, 4, but this corresponds for instance to three sinusoids in the spatial domain:

$$A \sin(u_1) + A \sin(u_2) + A \sin\left(\frac{1}{2}u_3\right)$$

where u_1 , u_2 , u_3 are coordinates in the three orthogonal spatial domains corresponding to the Fourier coordinates, and A is some constant.

3 2-D implementation of finding the minimum variance axis

To test the theory above we have implemented two algorithms evaluating local orientation of the 2-D images. Both of the algorithms rely on finding the eigenvalues and eigenvectors of the inertia matrix of the Fourier domain. The direction measurements for both of them are the same and based on the eigenvector(s) of the least eigenvalue. They differ on the certainty of the direction estimation. What has been referred to as image in the theory above, becomes a local neighbourhood of the total image in the

following discussion. To represent this local image at the point \bar{r}_j we multiply the larger image $f(\bar{r})$, by a window function $w(\bar{r}, \bar{r}_j)$:

$$h_j(\bar{r}) = f(\bar{r})w(\bar{r}, \bar{r}_j). \quad (9)$$

For simplicity we choose w a gaussian:

$$w(\bar{r}, \bar{r}_j) = \exp\left(-\frac{4}{d_w^2}\|\bar{r} - \bar{r}_j\|^2\right) \quad (10)$$

with $\|\cdot\|$ being the Euclidean norm: $\|\bar{r}\|^2 = \bar{r}^t\bar{r}$ and d_w being a constant controlling the 'diameter' of the local neighbourhood. The algorithm fits a least square error axis in the Fourier domain of the local image h_j corresponding to the least eigenvalues of the inertia matrix of the Fourier domain, \mathbf{J} . The computation is pursued in the spatial domain by means of equations (4) and (5). The axis found, $t\bar{k}_0$ is possible to represent by \bar{k}_0 . Since this is an axis, $-\bar{k}_0$ is an equivalent representation as well. For this reason the orientation of an axis $t\bar{k}_0$ can be defined as $2\phi_0$, if ϕ_0 is the direction angle of $\bar{k}_0 = k_x\hat{x}_1 + k_y\hat{x}_2$:

$$\phi_0 = \tan^{-1}(k_x, k_y) \quad (11)$$

Consequently both $-\bar{k}$ and \bar{k} will be mapped to the same angle, through $2\phi_0$. It can easily be shown that the eigenvalues of \mathbf{J} for the 2-D case, corresponding to h_j are:

$$\lambda_{0,1} = \frac{1}{2}(J_{22} + J_{11} \pm \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}) \quad (12)$$

The eigenvector corresponding to the least eigenvalue can be found to be $\bar{k}_0 = k_x\hat{x}_1 + k_y\hat{x}_2$ satisfying:

$$(J_{11} - \lambda_0)k_x - J_{12}k_y = 0 \quad (13)$$

with the Euclidean norm of unity. Thus the orientation becomes:

$$2\phi_0 = 2 \tan^{-1}(k_x, k_y) \quad (14)$$

By using¹ (12), (13), and (14)

$$2\phi_0 = \tan^{-1}(J_{22} - J_{11}, 2J_{12})$$

is obtained. Define the complex variable z as:

$$z = J_{22} - J_{11} + j2J_{12} \quad (15)$$

thus

$$2\phi_0 = \arg z$$

The certainty in this approximation of the local orientation depends on the behaviour of the eigenvalues λ_0 and λ_1 , according to the discussion in the previous section. A linear symmetry of the neighbourhood is probable if λ_0 is small relative to λ_1 . An 'ideal' linear

¹see Appendix B at the end of this chapter

symmetry occurs when $\lambda_0 = 0$ and $\lambda_1 \gg 0$. A certainty measure, C_{f_1} , incorporating these properties is

$$C_{f_1} = \left(\frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_0} \right)^c = \left[\frac{\sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}}{J_{22} + J_{11}} \right]^c = \left(\frac{|z|}{J_{22} + J_{11}} \right)^c \quad (16)$$

Here c is a positive constant, the purpose of which is to control the dynamic range of the certainty. C_{f_1} is defined to be 0 when $J_{22} + J_{11} = 0$. According to the previous section we then have a constant image and there is not a unique orientation for such images. It attains the maximum value 1 if and only if $\lambda_0 = 0$, because both λ_0 and λ_1 are non-negative. C_{f_1} decreases when the difference between the eigenvalues decreases. This property effectively tests whether the multiplicity of λ_0 is 2, in which case there is not a unique orientation minimizing the variance.

An alternative certainty measure is:

$$C_{f_2} = \lambda_1 - \lambda_0 = |z| \quad (17)$$

In this measure we do not get a unique certainty value when $\lambda_0 = 0$. Rather the confidence varies due to the largest eigenvalue when this happens. A consequence of this is that when the neighbourhood is linear symmetric with small energy, it is considered as less reliable even though the orientation measurement is correct. This is justified when it is desired that the certainty decreases continuously as the image becomes constant. Both of these certainty measures are considered.

Thus the task is reduced to express either of the certainty parameters C_{f_1} or C_{f_2} and the orientation estimation $2\phi_0$. The evaluations of these are governed by the equations (17), (16) and (15), which in turn rely on the efficient computation of the elements of the inertia matrix. Computation of these parameters for every neighbourhood in a discrete image is accomplished as follows.

Consider a discrete representation of the 2-D, bandlimited image $f(\bar{r})$. The continuous image can be reconstructed from its discrete samples, f_i by

$$f(\bar{r}) = \sum_i f_i \mu(\bar{r} - \bar{r}_i) \quad (18)$$

where $\mu(\bar{r})$ is an analytic function governing the behaviour of the continuous function between the discrete values. We will call it the interpixel function. μ can be assumed to be known since it is theoretically the inverse Fourier transform of a function which is 1 at the passband of the considered image and 0 at the outside. For its concentration in both Fourier and spatial domains we choose μ as a gaussian as well:

$$\mu(\bar{r}) = \exp\left(-\frac{4}{d_p^2} \|\bar{r}\|^2\right) \quad (19)$$

even though it is not an ideal interpixel function, since it is not strictly bandlimited. Under these conditions an approximation to J_{11} yields:

$$J_{11}^j = \frac{1}{4\pi^2} \int_{E_2} \left(\frac{\partial h_j}{\partial x_2} \right)^2 dE_2 = \frac{1}{4\pi^2} \sum_i \left(\frac{\partial f(\bar{r}_i)}{\partial x_2} \right)^2 \times \int_{E_2} \exp\left(-\frac{4}{d_p^2} \|\bar{r} + \bar{r}_j - \bar{r}_i\|^2 - \frac{4}{d_w^2} \|\bar{r}\|^2\right) dE_n \quad (20)$$

Here $(\frac{\partial f}{\partial x_2})^2$ is reconstructed from its samples

$$\left(\frac{\partial f(\bar{r})}{\partial x_2}\right)^2 = \sum_i \left(\frac{\partial f(\bar{r}_i)}{\partial x_2}\right)^2 \mu(\bar{r} - \bar{r}_i) \quad (21)$$

since it is bandlimited as well when f is. This requires that we have a version of $\frac{\partial f}{\partial x_2}$ oversampled by at least a factor 2 in every dimension. This is due to the fact that squaring a bandlimited function doubles its passband in every dimension. This is an effect which can be removed easily by resampling, if necessary. J_{11}^j corresponding to a neighbourhood, characterized by the coordinate vector \bar{r}_j and a window function of "diameter" d_w , is then computed as

$$J_{11}^j = \frac{1}{4\pi^2} \sum_i \left(\frac{\partial f(\bar{r}_i)}{\partial x_2}\right)^2 m_i^j \quad (22)$$

where m_i^j is given by (20)

$$\begin{aligned} m_i^j &= \int_{E_2} \exp\left(-\frac{4}{d_p^2} \|\bar{r} + \bar{r}_j - \bar{r}_i\|^2 - \frac{4}{d_w^2} \|\bar{r}\|^2\right) dE_2 \\ &= \frac{\pi}{4} \frac{d_p^2 d_w^2}{d_p^2 + d_w^2} \exp\left(-\frac{4}{d_p^2 + d_w^2} \|\bar{r}_j - \bar{r}_i\|^2\right) \end{aligned} \quad (23)$$

(22) is nothing but a convolution of the discrete version of the $(\frac{\partial f}{\partial x_2})^2$ by a gaussian. Since the gaussian decreases rapidly outside of a circle with the radius $\sqrt{d_p^2 + d_w^2}$, we can truncate it when it is sufficiently small. In our experiments this is done when it has decreased to about 1% of its maximum. Similarly J_{11}^j , J_{22}^j and J_{12}^j can be approximated by averaging the discrete images:

$$\begin{aligned} J_{22}^j &= \frac{1}{4\pi^2} \sum_i \left(\frac{\partial f(\bar{r}_i)}{\partial x_1}\right)^2 m_i^j \\ J_{12}^j &= \frac{1}{4\pi^2} \sum_i \frac{\partial f(\bar{r}_i)}{\partial x_1} \frac{\partial f(\bar{r}_i)}{\partial x_2} m_i^j \end{aligned} \quad (24)$$

Thus z for the point \bar{r}_j becomes:

$$\begin{aligned} z^j &= \frac{1}{4\pi^2} \sum_i \left[\left(\frac{\partial f(\bar{r}_i)}{\partial x_1}\right)^2 - \left(\frac{\partial f(\bar{r}_i)}{\partial x_2}\right)^2 + 2j \frac{\partial f(\bar{r}_i)}{\partial x_1} \frac{\partial f(\bar{r}_i)}{\partial x_2} \right] m_i^j \\ &= \frac{1}{4\pi^2} \sum_i \left(\frac{\partial f(\bar{r}_i)}{\partial x_1} + j \frac{\partial f(\bar{r}_i)}{\partial x_2} \right)^2 m_i^j = \frac{1}{4\pi^2} \sum_i u_i m_i^j \end{aligned} \quad (25)$$

where u_i is the complex valued image obtained by taking the square of the gradient of the image, interpreted as a complex number instead of a real vector:

$$u_i = \left(\frac{\partial f(\bar{r}_i)}{\partial x_1} + j \frac{\partial f(\bar{r}_i)}{\partial x_2} \right)^2. \quad (26)$$

Similarly

$$J_{11}^j + J_{22}^j = \frac{1}{4\pi^2} \sum_i |u_i| m_i^j \quad (27)$$

is obtained. Thus calling the discrete image defined by (26) as u and the filter defined by (23) as m we have:

$$\begin{aligned} 2\phi_0 &= \arg(u * m) \\ C_{f1} &= \frac{|u * m|^c}{(|u| * m)^c} \\ C_{f2} &= \frac{1}{4\pi^2} |u * m| \end{aligned} \quad (28)$$

where the symbol $*$ represents the usual convolution operation. Here $\arg(\cdot)$, $|\cdot|$ and $(\cdot)^c$ operations are assumed to be applied pointwise to their arguments and thus $2\phi_0$, C_{f1} and C_{f2} become images representing local orientation and certainties. The first algorithm is to evaluate $2\phi_0$ and C_{f1} and the second algorithm to evaluate $2\phi_0$ and C_{f2} .

The discrete partial derivatives necessary for the evaluation of u can be produced by convolution with various filters. For the sake of completeness we just mention the technique used before: Expansion of the image in its interpixel functions and application of the derivative operation.

$$\left(\frac{\partial f(\bar{r}_j)}{\partial x_1}\right) = \sum_i f_i \frac{\partial}{\partial x_i} \mu(\bar{r}_j - \bar{r}_i) \quad (29)$$

As before this gives us a filter which decreases rapidly outside of a small region close to the examined point \bar{r}_j . The evaluation of (29) and (28) is easily computed on hardware with support for convolution. In the experiments below a GOP-300 computer has been used.

4 Experimental results

In the experiments, the implementation proposed in the previous section has been tested for detection of the linear symmetric neighbourhoods. This is carried out in two steps:

1) Evaluate a partial derivative picture. In reality the complex variable u_j of the neighbourhood at the point \bar{r}_j ,

$$\frac{\partial f(\bar{r}_j)}{\partial x_1} + j \frac{\partial f(\bar{r}_j)}{\partial x_2}$$

is computed by using (29) and then this complex number is squared.

2) Estimate local orientation, $2\phi_0$, and the certainty of this estimation, either of C_{f1} or C_{f2} according to (28) based on the image obtained in step 1).

Figure 2 shows an image containing all possible directions for sine waves with exponentially increasing frequency in the radial direction of the circles. Gaussian uncorrelated white noise is added to the right half of the image with the proportion 1:3 that



Figure 2: The test image used in the experiments. The straight line is the line along which the profiles of results are presented at the proceeding figures.

is $0.25f_i + 0.75Y_i$ where f_i is the image intensity and Y_i is the stochastic variable with the distribution of $N(0,32)$. In the experiments this proportion is varied and the local orientation is examined for different sizes of filters in the two steps mentioned before. In general it could be observed that the filter size of the first step affects the accuracy of the orientation detection more than the size of the neighbourhood given by the gaussian, (23), in the second step. This is not surprising, because the probability that the energy at a high frequency is erroneous is higher than the same probability for a low frequency. The reason is that most of the natural errors are composed of high frequencies like aliasing error, discretization error, measurement error, etc. A squaring of the gradient image causes these errors to propagate to lower parts of the local frequency spectrum. This makes it difficult to remove the noise by increasing the filter size, i.e. low pass filtering. Figure 3 shows the orientation estimation of the profile of the test image cut along the line passing through the origin of the circles shown in Figure 2. The profiles at the left part of the test circles demonstrate these phenomena. Both profiles should have constant levels, since the orientation of the profile is constant. The shown two profiles are due to two different filter configurations of two estimations. Profile 1 is due to a 9×9 filter in the first and a 15×15 gaussian averaging filter in the second step. The second profile shows the result of a 5×5 gaussian derivative filter at the first step and 21×21 at the second step. It can be observed that in the latter filter configuration the effect of the noise is removed at low and medium frequencies while at high frequencies the orientation estimate is not as good as at the lower frequencies.

Control of the certainty in the presence of noise without disturbing the certainty in

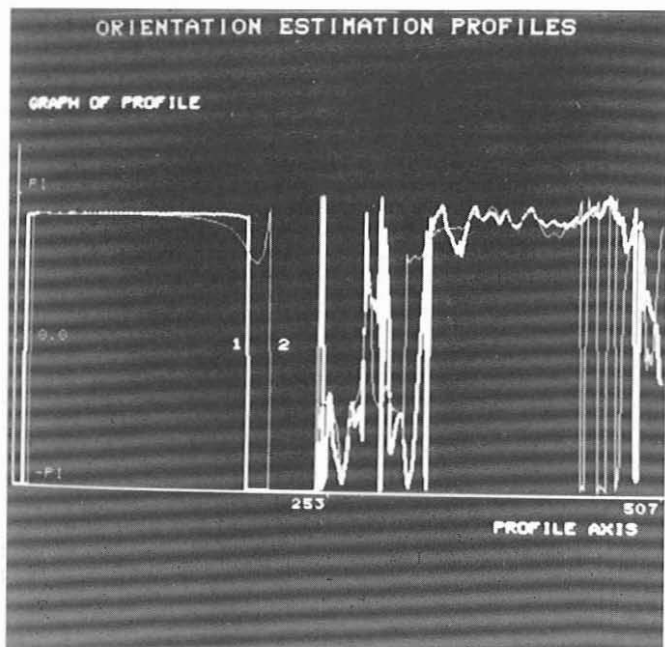


Figure 3: Orientation estimation with two different filter configurations. Graph 1 illustrates 9×9 and 15×15 configuration at the two steps of the algorithm, while 2 is due to 5×5 and 21×21 configuration.

the less noisy parts is a desirable feature for many applications. Profiles in Figure 4 show the certainty measures C_{f1} and C_{f2} of the estimation given by profile 1 in Figure 3. Profiles 1 and 3 correspond to C_{f1} with $c=6$ and $c=1$ respectively. C_{f2} is given by profile 1. Since this measure is not a relative measure like C_{f1} , it has a high degree of frequency dependence. It is considerably more 'suspect' outside of the pass band, compared to the one given by profiles 1 and 3. A natural consequence of this is that the fluctuations in the noisy part of the image are small with a low level certainty. The frequency sensitivity band of the certainty parameter C_{f2} is due to the derivation in the first step, and averaging in the second step. In the Fourier domain this corresponds to a multiplication of an increasing and a decreasing function at low frequencies. The center frequency can thus be varied by varying the scale of the filters at the two steps, Figure 5. This certainty measure can be used when it is desirable to control the orientation measurements for the neighbourhoods with a priori known frequencies.

5 Conclusion

It is experimentally verified that the problem of orientation detection within a local neighbourhood is possible to solve in the Fourier domain for the 2-D case. The problem for the 3-D case is possible to solve in a similar way. The first section is sufficiently general to handle the 3-D problem, while Section 2 uses some fundamental properties of

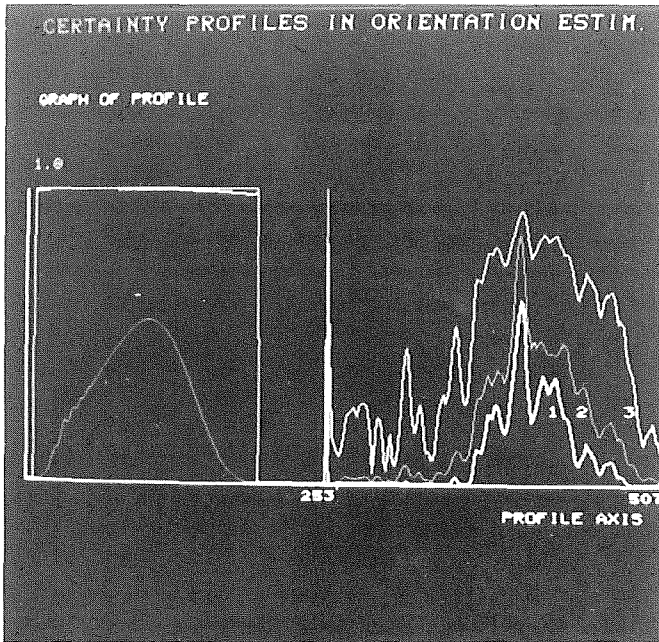


Figure 4: Certainty measures for the profile 1 in Figure 3. Graphs 1 and 3 are due to C_{f1} with $c = 6$ and $c = 1$ respectively, while 2 corresponds to C_{f2} .

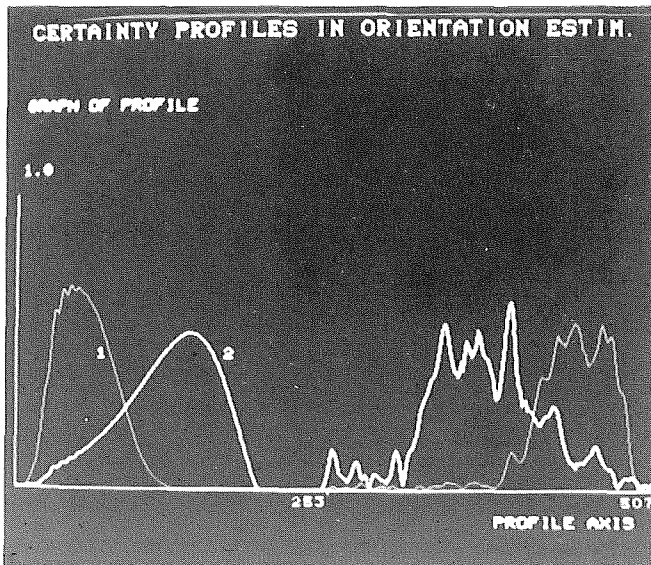


Figure 5: Frequency dependence of C_{f2} . Graph 1 corresponds to 15×15 and 21×21 configuration at the two steps of the algorithm. 2 corresponds to 5×5 and 19×19 configuration.

the 2-D case and its relation to the complex z -plane to accomplish the representation of orientation and the certainty of the estimation. The proposed certainties C_{f1} and C_{f2} and the orientation $2\phi_0$ are shown to be computable by averaging the complex valued image u which is the square of the gradient represented in the complex form. For the first algorithm producing C_{f1} and $2\phi_0$, 5 real convolutions are required, while in the second algorithm producing C_{f2} and $2\phi_0$ 4 real convolutions are required. In both cases the algorithms can, under the condition that f is bandlimited, be summarized in the compact forms: $\frac{m^*(\nabla * f)^2}{m^*|\nabla * f|^2}$ and $m * (\nabla * f)^2$ with ∇ being the complex gradient filter given by $D_{x_1} + jD_{x_2}$, and m the averaging filter. The resulting complex image can be seen and interpreted directly on a colour TV monitor, if the magnitude of the image corresponding to the certainty controls the intensity, and the argument corresponding to the orientation controls the colour [7].

One of the reasons for first evaluating a 2-D implementation, apart from it having less data compared to 3-D, is the difficulty of displaying the orientation of 3-D images together with the certainties of the given orientation estimation. The experimental work indicates clearly that, for higher dimensions, it is possible to evaluate the local orientation accurately together with its certainty with controllable behaviour when noise is present. The first section shows that even though there is no obvious linear symmetry in the neighbourhood considered, the estimation found is optimal in the least squares sense.

Acknowledgements

The authors would like to express their gratitude to the other members of the Computer Vision Laboratory group for valuable discussions and to the STU for the financial support.

References

- [1] H. Knutsson: "Filtering and reconstruction in image processing." Dissertation No. 88, 1982, Linköpings Studies in Science and Technology, Linköping University, Sweden.
- [2] R. Lenz: "Optimal filtering." Linköping university, Internal Report of the Department of Electrical Engineering, 1985.
- [3] Robert A. Hummel: "Feature Detection Using Basis Functions." Computer Graphics and Image Processing 9, 1979.
- [4] A. Wouk: "A course of applied functional analysis." Wiley, New York, 1979.
- [5] J. Bigün, G.H. Granlund: "Central symmetry modelling" proc. EUSIPCO-86 part 2 pp 883-886
- [6] J. Bigün: "Circular symmetry models in image processing" Licentiate report no: 85 1986 Linköping studies in science and technology
- [7] G.H. Granlund: "In Search of a General Picture Processing Operator." Computer Graphics and Image Processing 8, 155-173 (1978).
- [8] G.H. Granlund: "Hierarchical Image Processing." Proceedings of SPIE Technical conference, Geneva, April 18-27, 1983.
- [9] P.E. Danielsson: "Rotation invariant linear operators with directional response." Proceedings of 5'th international conference on pattern recognition, December 1980.
- [10] I.L. Meriam: "Statics." John Wiley & Sons, New York, 1980.

A Overall results

Here we will present the pictures underlying the experimental results of the 2-D implementation in this chapter. The results presented earlier correspond to certainty and orientation estimations along a line and consequently of a one-dimensional nature. The purpose of this appendix is to present the overall results. Figure 6 is included for reference. It represents the true orientation of Figure 2. It is coded in the same same way as the figures below. In colour images every pixel is a complex valued number. The magnitude of the complex number is coded as brightness of the pixel and the argument of the number is coded as its hue. The certainty estimates of Figures 8, 9 and 10 are given a close look along a line by Figure 4. It is clearly seen that the filter responses illustrated in Figure 4 are independent of the inclination angle of the chosen profile line. The orientation estimates of Figures 9 and 12 correspond to Figure 3. The accuracy of the orientation estimate at higher frequencies is thereby demonstrated to be dependent on the filter choice. Some care, in the form of resampling or interpolation, should be taken when performing partial derivations. Similarly Figure 14 and Figure 15 correspond to Figure 5 and illustrate that it is possible to tune the algorithm to certain frequency bands. Finally the results of the algorithms on a realistic image is given.

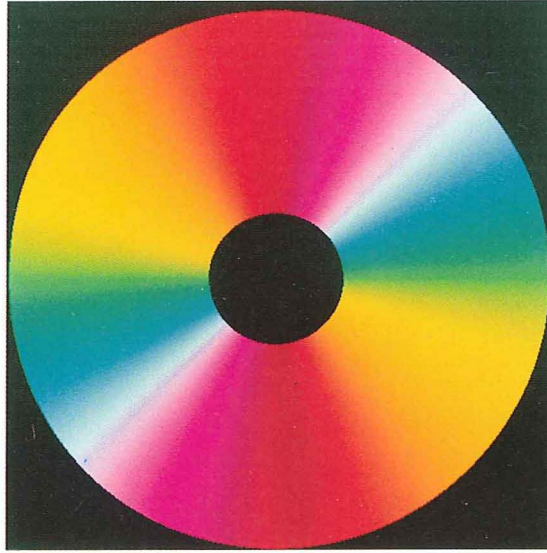


Figure 6: Local orientation of the original image, which serves as a reference image for the following colour pictures. The certainty, which is coded as brightness as before, is set to maximum value in the bright area.

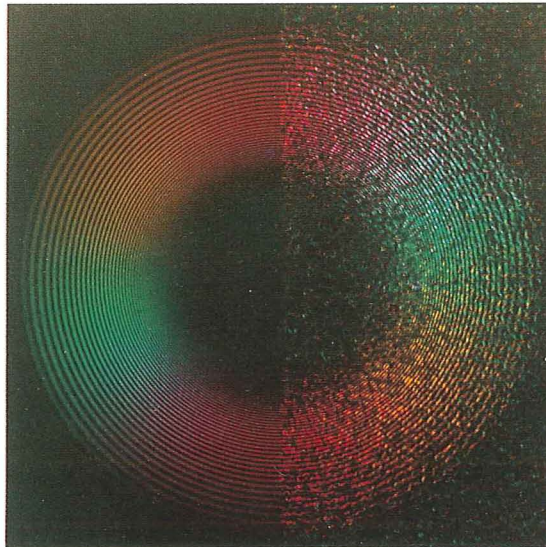


Figure 7: This image illustrates $u = z^2 = \left(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y}\right)^2$. The brightness represents $|z^2|$ and the hue represents $\arg(u)$. It is obtained by a pair of 9x9 filters based on Figure 2.

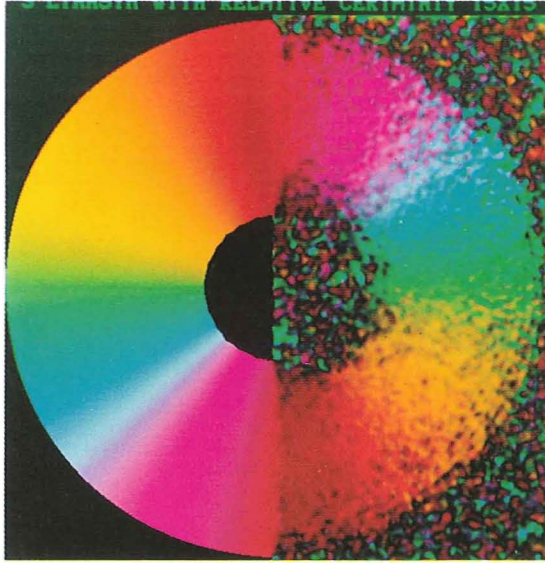


Figure 8: Local orientation estimation. The brightness represents $\frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_0}$ and the hue represents the orientation of the estimated linear symmetry, which is given by twice the inclination angle of a line or an edge. It is obtained by filtering Figure 7 with a 15x15 filter.



Figure 9: Same as Figure 8 but $(\frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_0})^6$. This suppresses noise without disturbing the pure linear symmetric neighbourhoods.

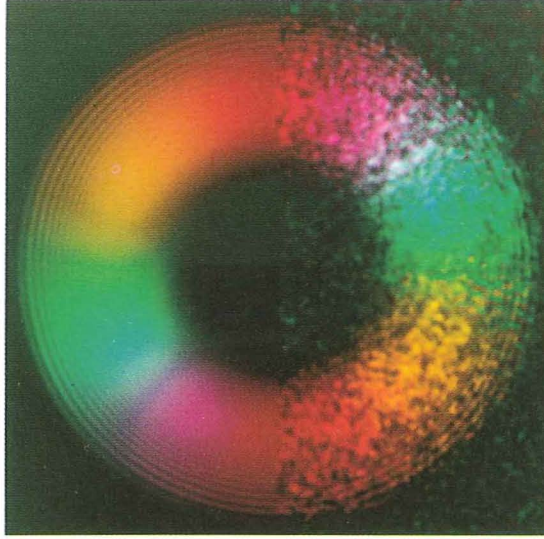


Figure 10: The image illustrates the energy dependent existence of linear symmetry, that is, the brightness represents $\lambda_1 - \lambda_0$ and the hue represents the estimated orientation of linear symmetry.

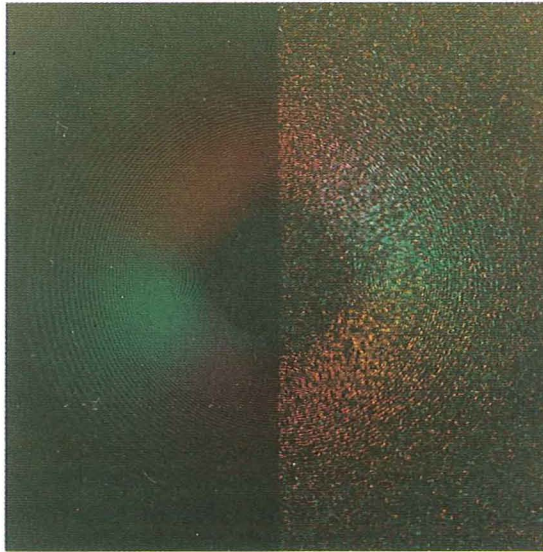


Figure 11: The image illustrates $(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y})^2$. It is obtained using a pair of 5x5 filters.



Figure 12: Indication of linear symmetry, with brightness representing $\frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_0}$. It is obtained by filtering Figure 11 with a 21x21 filter.

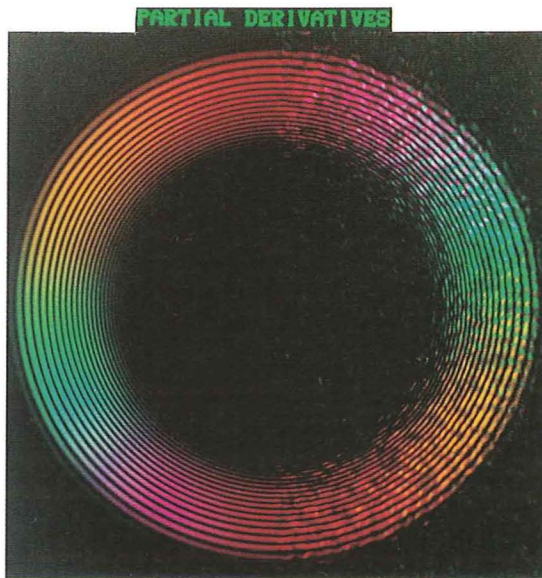


Figure 13: The image illustrates $(\frac{\partial f}{\partial x} + j \frac{\partial f}{\partial y})^2$. It is obtained using a pair of 15x15 filters.



Figure 14: Indication of linear symmetry, with the brightness $\lambda_1 - \lambda_0$ and the hue indicating orientation. It is obtained by filtering Figure 13 with a 21x21 filter.

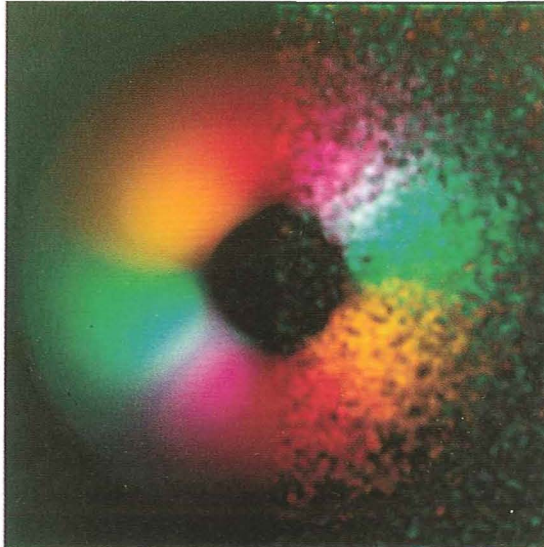


Figure 15: Indication of linear symmetry, with the brightness $\lambda_1 - \lambda_0$ and the hue indicating orientation. It is obtained by filtering Figure 11 with a 19x19 filter. Observe the change of frequency band sensitivity by comparing this figure with Figure 14. A comparison along a line is given in Figure 5.

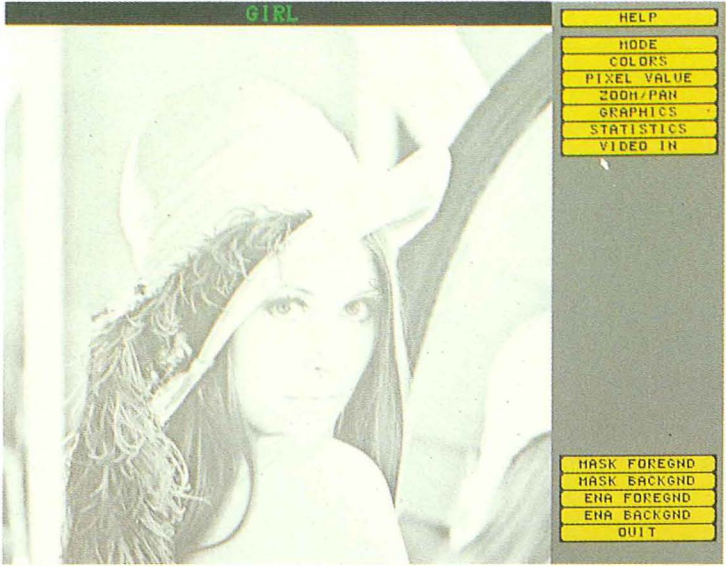


Figure 16: Test image

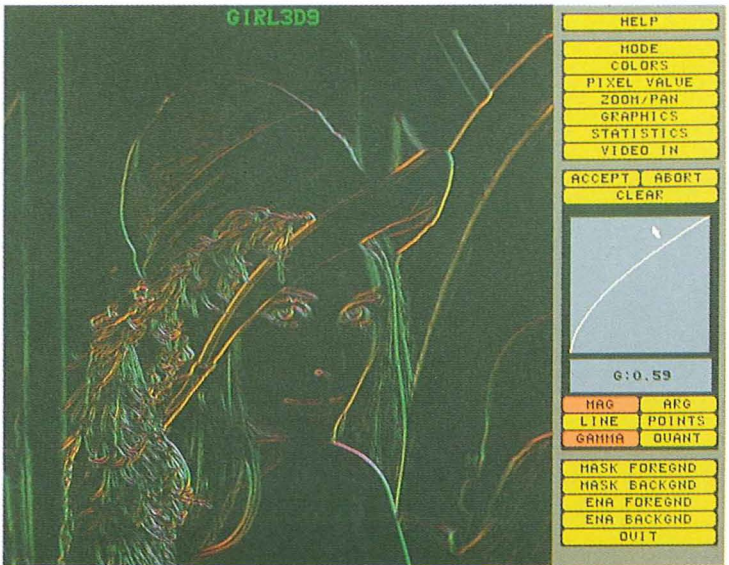


Figure 17: Result of orientation detection. The image points represent the orientation vector z_1 . Sizes of used filters were 3x3 and 9x9.

B The relation between the eigenvectors and the complex z plane.

Here we will show that the orientation angle of a line represented by an eigenvector is reduced to the simple form given by:

$$2\phi_0 = \tan^{-1}(J_{22} - J_{11}, 2J_{12}). \quad (30)$$

This relation shows that the eigenvectors of the inertia matrix in the 2-D case are related to the average of the complex number $(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})^2$.

Putting

$$k_x = J_{12}$$

and

$$\begin{aligned} k_y &= J_{11} - \lambda_0 \\ &= J_{11} - \frac{1}{2}(J_{22} + J_{11} - \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}) \\ &= \frac{1}{2}(J_{22} - J_{11} + \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}) \geq 0 \end{aligned} \quad (31)$$

provides us the eigenvector corresponding to the least eigenvalue. Using double angle formulas for sine and cosine functions, equation (14) implies:

$$2\phi_0 = 2 \tan^{-1}(k_x, k_y) = \tan^{-1}(k_x^2 - k_y^2, 2k_x k_y). \quad (32)$$

Utilizing (31)

$$\begin{aligned} k_x^2 - k_y^2 &= J_{12}^2 - (J_{11} - \lambda_0)^2 \\ &= J_{12}^2 - \frac{1}{4}[(J_{11} - J_{22})^2 + (J_{22} - J_{11})^2 + 4J_{12}^2 + \\ &\quad + 2(J_{11} - J_{22})\sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}] \\ &= -\frac{1}{2}(J_{11} - J_{22})(J_{22} - J_{11} + \sqrt{(J_{22} - J_{11})^2 + 4J_{12}^2}) = (J_{22} - J_{11})k_y \end{aligned}$$

is obtained. Thus

$$2\phi_0 = \tan^{-1}((J_{22} - J_{11})k_y, 2k_x k_y) = \tan^{-1}((J_{22} - J_{11}), 2J_{12})$$

is obtained since $k_y \geq 0$.

Chapter IV

PATTERN RECOGNITION BY
DETECTION OF LOCAL
SYMMETRIES

The philosopher who frequently meets with the divine and harmony will, as far as it is possible for man, himself become divine and harmonious. But, like everybody else, he will be exposed to slander.

Plato

PATTERN RECOGNITION BY DETECTION OF LOCAL SYMMETRIES[†]

Josef Bigün
Linköping University
Department of Electrical Engineering
Computer Vision Laboratory
S-581 83 Linköping Sweden

Abstract

The symmetries in a neighbourhood of a gray value image are modelled by conjugate harmonic function pairs. These are shown to be a suitable curve linear coordinate pair, in which the model represents a neighbourhood. In this representation the image parts, which are symmetric with respect to the chosen function pair, have iso-gray value curves which are simple lines or parallel line patterns. The detection is modelled in the special Fourier domain corresponding to the new variables by minimizing an error function. It is shown that the minimization process or detection of these patterns can be carried out for the whole image entirely in the spatial domain by convolutions. What will be defined as the partial derivative image is the image which takes part in the convolution. The convolution kernel is complex valued, as are the partial derivative image and the result. The magnitudes of the result are shown to correspond to a well defined certainty measure, while the orientation is the least square estimate of an orientation in the Fourier transform corresponding to the harmonic coordinates. Applications to four symmetries are given. These are circular, linear, hyperbolic and parabolic symmetries. Experimental results are presented.

1 Introduction

Describing events in neighbourhoods of a gray value image is an increasing need in Computer Vision. The most extensively studied event is the existence of lines and edges. Also circular patterns have been subject to investigation [9], [5], [2], [3]. The generalized Hough transform, [6], is general and accurate enough to find arbitrary curves with the drawback of being computationally demanding. In the following we will give a method for detection of a large class of symmetries in a gray value neighbourhood which is a generalization of the work done in [3], [2] for circular symmetry.

[†]PROCEEDINGS OF PATTERN RECOGNITION IN PRACTICE III, Amsterdam, May 18-20, 1988; Elsevier Science Publishers.

By a neighbourhood of a point will be understood the image multiplied by a window function placed at the point. For its behaviour to be satisfactory in practical situations we will assume this window function to be a gaussian with a sufficiently large standard deviation instead of a sharp window. But the method is not restricted to this choice. In Section 2, a definition of symmetry will be given by means of a pair of conjugate harmonic functions in which the neighbourhood coordinates are represented. After this representation, the neighbourhoods with iso-gray value curves associated with a linear combination of these chosen coordinates constitute the family of neighbourhoods for which the detection is developed.

The detection is based on minimization of an error function in the Fourier domain, but computed entirely in the spatial domain. In Section 3, this minimization process is shown to be a convolution of the complex valued partial derivative image with a complex valued filter. The result delivers an angle corresponding to a subclass of neighbourhoods within the family of the neighbourhoods the a priori chosen function pair can handle. By changing this angle, all patterns in the symmetry model can be reached which is in analogy of a class of lines and edges with the same orientation. All lines and edges are covered by changing the orientation of the latter class. In fact, in Section 4 it is shown that lines and edges can be modelled in this general framework just as any other symmetry.

Besides the orientation of the found symmetry, the minimization process delivers a certainty defined by the minimum and maximum error. The higher the significance of the found symmetry orientation for the neighbourhood, the higher this certainty becomes. Also a non-energy dependent certainty measure is derived, which is useful for pictures with different light conditions in different parts of the picture. In Section 4 applications and experimental results are given.

Four symmetry models are covered by using the general methodology given in Sections 3 and 4. These are circular, linear, hyperbolic and parabolic symmetries. The conjugate harmonic function pairs are easily established by observing that all analytic functions' real and imaginary parts are such pairs. The analytic functions connected with the symmetries mentioned are the elementary functions: $\log z$, z , z^2 and \sqrt{z} .

2 Modeling the local neighbourhoods by harmonic functions

Let $u(x, y)$ be a *harmonic function* defined in the neighbourhood, that is, it is continuous together with its partial derivatives of the first two orders and satisfies the *Laplace's equation*:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (1)$$

Due to the linear character of Laplace's equation, the linear combinations of the harmonic functions are also harmonic. If two harmonic functions u and v satisfy the

Cauchy-Riemann equations:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad (2)$$

then v is said to be the *conjugate harmonic function* of u . The imaginary part of any analytic function is the conjugate harmonic function of the real part. In general v does not need to be a single valued function even if u is. Here we will assume both u and v to be single valued. By definition, (2), a curve pair defined by:

$$\xi = u(x, y) \quad (3)$$

$$\eta = v(x, y) \quad (4)$$

are orthogonal to each other at their intersection points for any constants ξ and η . For non trivial $u(x, y)$ and $v(x, y)$, (3-4) define a coordinate transformation which is reversible almost everywhere.

Consider a neighbourhood around a point in a gray value image. Let this neighbourhood be represented by the real function $f_1(x, y)$ which attains positive real values. For simplicity we will assume that the origin of the Cartesian coordinate system in which the neighbourhood is represented as being the considered point. Let $\bar{r} = (\xi, \eta)^t$ be defined through (3-4). The representation of the neighbourhood is then possible in these coordinates and yields $f_1(x, y) = f_2(\xi, \eta)$.

Definition 1 *The local neighbourhood $f(x, y)$ represented in its local Cartesian coordinates, is said to be symmetric with respect to the coordinates $(\xi, \eta)^t$ if there exists a one dimensional function g so that $f(x, y) = g(a\xi + b\eta)$ for some real constants, a and b . Here $(\xi, \eta)^t = (u(x, y), v(x, y))^t$ and v is the harmonic conjugate function of u .*

This definition suggests that the iso-gray value curves of a neighbourhood, which is symmetric with respect to a coordinate pair $(\xi, \eta)^t$, are *parallel lines* in this coordinate system. In the following we will develop a method to select all neighbourhoods of an image, which are symmetric with respect to an a priori chosen local coordinate transformation given by the harmonic pair $(\xi, \eta)^t = (u(x, y), v(x, y))^t$. It may be thought that every neighbourhood is first represented in its local Cartesian coordinates and then represented in the transformed coordinate system which in turn tested whether its iso-gray value curves are parallel lines. Of course we will not suggest doing all these time-consuming transformations in practice but rather their equivalents which are computationally easy to evaluate. In general a neighbourhood is not symmetric with respect to an a priori coordinate transformation. But when it is symmetric, we will observe that the energy of the neighbourhood is concentrated to a line through the origin in its Fourier transform domain corresponding to the new coordinates. Our approach will be to check whether there exists a line concentration in the local Fourier domain obtained by means of the new coordinates. But before doing this, we should specify what we mean by this special Fourier transformation and a line concentration in it.

Let the local image be represented by $f(\xi, \eta)$, then the corresponding Fourier transformation is defined as:

$$F(\omega, \zeta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-i(\omega\xi + \zeta\eta)] f(\xi, \eta) d\xi d\eta \quad (5)$$

Here we assumed that ξ and η have the range $[-\infty, \infty]$. For cases where either of the variables have a finite range, f is put to zero outside its value set. Then the corresponding transform variable (ω or ζ) constitutes a numerable periodic discrete set (Fourier series expansion in that variable). In the following, all integrations in the spatial domain should be interpreted in this sense. The inverse transform is obtained through:

$$f(\xi, \eta) = (2\pi)^{-2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[i(\omega\xi + \zeta\eta)] F(\omega, \zeta) d\omega d\zeta. \quad (6)$$

where one of the integrations is changed to a summation over a discrete periodic set if ξ or η has finite range. In the following, all integrations in the Fourier domain should be interpreted in this sense.

Theorem 2 *A symmetric neighbourhood with respect to the coordinates $(\xi, \eta)^t$, that is $f(a\xi + b\eta)$, has a Fourier transform, in these coordinates, which is concentrated to a line through the origin:*

$$(\mathcal{F}f)(\omega, \zeta) = \frac{1}{2\pi} \delta(b\omega - a\zeta) \int_{-\infty}^{\infty} f(t) \exp(-it(a\omega + b\zeta)) dt \quad (7)$$

where δ is the dirac distribution, and the symmetry direction vector, $(a, b)^t$, has the modulus of unity:

$$\sqrt{a^2 + b^2} = 1. \quad (8)$$

The proof of the theorem is straight forward and will be omitted here, [1]. As already mentioned, the initial position for determining the symmetry is the corresponding Fourier domain. Any neighbourhood in the image will then have a Fourier transform which is not necessarily concentrated to a line through the origin. We will fit the best line to the corresponding Fourier domain in the least square sense. If there exists a symmetry according to an a priori model then the error will be low. The greater the size of the minimum error, the more degradation from this symmetry will be observed. To any neighbourhood, a line through the origin of its Fourier transform, or rather its orientation, will be given together with the minimum error expressing the significance of the found symmetry. The degradation of $F(\omega, \zeta)$ from a line $t\bar{k}_\theta = t \cdot (\cos \theta, \sin \theta)^t$ with $t \in R$, or the error in the least square estimation, will be given by:

$$E(\theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d^2(\bar{k}, \bar{k}_\theta) |F(\bar{k})|^2 d\omega d\zeta. \quad (9)$$

Here $d^2(\bar{k}, \bar{k}_\theta)$ is the squared Euclidean distance between the vector $\bar{k} = (\omega, \zeta)^t$ and $\bar{k}_\theta = (\cos \theta, \sin \theta)^t$, of which the former symbolizes the transposed coordinate vector in the special Fourier domain, Figure 1. Without proof, [2], we mention the following theorem.

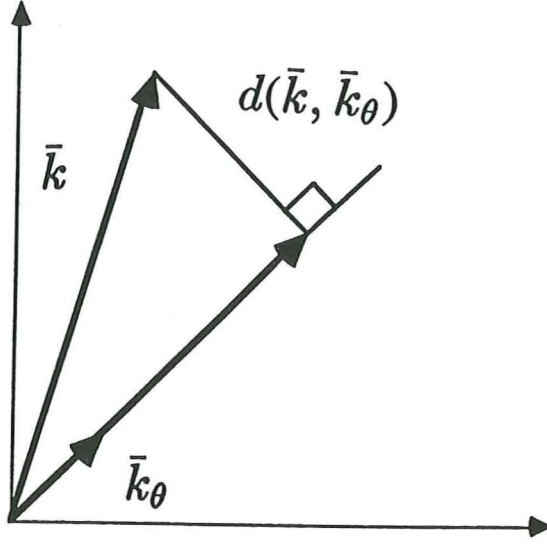


Figure 1: The figure illustrates the Euclidean distance in the special Fourier domain, $d(\bar{k}, \bar{k}_\theta)$, between the coordinate point \bar{k} and the axis denoted by \bar{k}_θ .

Theorem 3 The double angle minimizing $E(\theta)$, e.g. $2\theta_{\min}$ is given by the formula:

$$2\theta_{\min} = \tan^{-1}(\omega_d^2 - \zeta_d^2, 2p_d) \quad (10)$$

where

$$\begin{aligned} \omega_d^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega^2 |F(\omega, \zeta)|^2 d\omega d\zeta \\ \zeta_d^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \zeta^2 |F(\omega, \zeta)|^2 d\omega d\zeta \\ p_d &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \zeta \omega |F(\omega, \zeta)|^2 d\omega d\zeta. \end{aligned}$$

The error corresponding to the line given by θ_{\min} is obtained through

$$E(\theta_{\min}) = \frac{1}{2}[\omega_d^2 + \zeta_d^2 - \sqrt{(\omega_d^2 - \zeta_d^2)^2 + 4p_d^2}]. \quad (11)$$

Moreover, the line corresponding to the angle maximizing the error $E(\theta)$, is orthogonal to the line minimizing the error. The maximum error is given by:

$$E(\theta_{\max}) = \frac{1}{2}[\omega_d^2 + \zeta_d^2 + \sqrt{(\omega_d^2 - \zeta_d^2)^2 + 4p_d^2}]. \quad (12)$$

Since we know that the least square error line passes through the origin, it is possible to specify it by a single real variable. We will choose the label 2θ for a line given by

$t \cdot (\cos \theta, \sin \theta)^t$. Labelling a line by the angle of 2θ instead of θ is more suitable since 2θ maps θ and $\theta + \pi$ to the same angle. This is a desirable property, since we do not want to claim that we have found two different lines when we obtain two angles differing by π .

The least square error $E(\theta_{min})$ is a positive number which we really cannot interpret if it is not related to another quantity, since we should know what is a large error and what is a small one. One way to do that, is to consider the quantity:

$$C_{f1} = E(\theta_{max}) - E(\theta_{min}) = \sqrt{(\omega_d^2 - \zeta_d^2)^2 + 4p_d^2} \quad (13)$$

which can be used as a certainty measure for the found symmetry. This is simply a quantity which measures how much better a best fitted line is compared to the worst line. It is non-negative and is large for neighbourhoods possessing the symmetry under consideration and it decreases smoothly for the neighbourhoods degrading from this type of symmetry. This definition allows us to consolidate the obtained orientation and the corresponding symmetry to a complex number z_1 , [8]:

$$z_1 = \omega_d^2 - n_d^2 + i2p_d = C_{f1} \exp(i2\theta_{min}). \quad (14)$$

Another way to interpret $E(\theta_{min})$ is to use

$$C_{f2} = \frac{E(\theta_{max}) - E(\theta_{min})}{E(\theta_{max}) + E(\theta_{min})} \quad (15)$$

as a certainty measure. Unlike C_{f1} , it is not energy dependent and $C_{f2} = 1$ if and only if $E(\theta_{min}) = 0$ which is the condition for a truly symmetric neighbourhood. Moreover it varies between 0 and 1. The closer this measure is to 1 the more significant the found symmetry is. If desired, C_{f2} and $2\theta_{min}$ can be consolidated to a complex number $z_2 = C_{f2} \exp i2\theta$ as before.

If the two symmetry models specified by the local coordinate transformations

$$(\xi_1, \eta_1) = (u_1(x, y), v_1(x, y)) \quad (\xi_2, \eta_2) = (u_2(x, y), v_2(x, y)) \quad (16)$$

are given, then can we relate the two certainty measures obtained for a neighbourhood? That is can we say that one symmetry model describes the neighbourhood better than the other when the certainty for one of them is higher? Since equality occurs for $C_{f2} \leq 1$ only when we have a truly symmetric neighbourhood and the right hand side is independent of the chosen transformations describing the symmetries, the two certainties of type C_{f2} can be related to each other. The following theorem will help us to answer the question when the certainties are of C_{f1} type.

Theorem 4 (Energy Conservation) *The sum of the maximum and the minimum error is independent of the coordinate system chosen for symmetry investigation of a given neighbourhood f :*

$$E(\theta_{max}) + E(\theta_{min}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\omega^2 + \zeta^2) |F(\omega, \zeta)|^2 d\omega d\zeta \quad (17)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial f}{\partial \xi} \right)^2 + \left(\frac{\partial f}{\partial \eta} \right)^2 d\xi d\eta \quad (18)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 dx dy. \quad (19)$$

The Fourier transform of f in the new coordinates is F as before.

Proof:The first equality is obtained by theorem 3 the second is provided by the Parseval relation. To prove the third relation we utilize the chain rule:

$$\frac{\partial f}{\partial \xi} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \xi}$$

and then utilize the fact that the Jacobian

$$\frac{\partial(\xi, \eta)}{\partial x \partial y} \triangleq \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{pmatrix}$$

is given by

$$\frac{\partial(\xi, \eta)}{\partial x \partial y} = \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ -\frac{\partial \xi}{\partial y} & \frac{\partial \xi}{\partial x} \end{pmatrix}$$

since $\xi = u(x, y)$ and $\eta = v(x, y)$ constitute a harmonic pair fulfilling the Cauchy-Riemann equations (2), almost everywhere. Remembering the relation

$$\frac{\partial(x, y)}{\partial \xi \partial \eta} = \left(\frac{\partial(\xi, \eta)}{\partial x \partial y} \right)^{-1} = \frac{1}{\left(\frac{\partial \xi}{\partial x} \right)^2 + \left(\frac{\partial \xi}{\partial y} \right)^2} \begin{pmatrix} \frac{\partial \xi}{\partial x} & -\frac{\partial \xi}{\partial y} \\ \frac{\partial \xi}{\partial y} & \frac{\partial \xi}{\partial x} \end{pmatrix}$$

we obtain

$$\frac{\partial f}{\partial \xi} = \frac{1}{\left(\frac{\partial \xi}{\partial x} \right)^2 + \left(\frac{\partial \xi}{\partial y} \right)^2} \left(\frac{\partial f}{\partial x} \frac{\partial \xi}{\partial x} - \frac{\partial f}{\partial y} \frac{\partial \xi}{\partial y} \right).$$

Similarly

$$\frac{\partial f}{\partial \eta} = \frac{1}{\left(\frac{\partial \xi}{\partial x} \right)^2 + \left(\frac{\partial \xi}{\partial y} \right)^2} \left(\frac{\partial f}{\partial x} \frac{\partial \xi}{\partial y} + \frac{\partial f}{\partial y} \frac{\partial \xi}{\partial x} \right)$$

is obtained. Thus

$$\left(\frac{\partial f}{\partial \xi} \right)^2 + \left(\frac{\partial f}{\partial \eta} \right)^2 = \frac{1}{\left(\frac{\partial \xi}{\partial x} \right)^2 + \left(\frac{\partial \xi}{\partial y} \right)^2} \left(\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right)$$

together with the variable substitution according to (3) and (4) concludes the proof.

It follows from the theorem even the certainties of type C_{f1} , can be related to each other since

$$0 \leq C_{f1} \leq E(\theta_{max}) + E(\theta_{min}) \quad (20)$$

and the right hand of the inequality is independent of the choice of the transformation for description of symmetry. Equality occurs only when the neighbourhood is truly symmetric with respect to the chosen coordinates. The closer C_{f1} is to this common upper bound, the better the chosen transformation describes the underlying symmetry.

The certainties and the orientation above are evaluated in the Fourier domain corresponding to the a priori chosen coordinate transformation. By using the Parseval

relation it is possible to transfer these evaluations to the spatial domain:

$$\omega_d^2 = (2\pi)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial f}{\partial \xi}\right)^2 d\xi d\eta \quad (21)$$

$$n_d^2 = (2\pi)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial f}{\partial \eta}\right)^2 d\xi d\eta \quad (22)$$

$$p_d = (2\pi)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial f}{\partial \eta} \frac{\partial f}{\partial \xi} d\xi d\eta. \quad (23)$$

3 Detection of local symmetries

In the previous section we have presented a model by which we could test the symmetry with respect to a fixed coordinate system in the entire image. Although (21–23) presents a way to evaluate the quantities necessary for the symmetry test of a neighbourhood in the spatial domain, it is still cumbersome to evaluate these integrals for every neighbourhood. The chain rule applied to these formulas along with the Cauchy-Riemann equations, (2), establishes:

$$z_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right)^2 \exp[-i2 \arg\left(\frac{\partial \xi}{\partial x} + i\frac{\partial \xi}{\partial y}\right)] dx dy. \quad (24)$$

By this result, z_1 the unit incorporating the certainty and the found symmetry orientation, is ready for approximation by discrete image data. Assume that f is a bandlimited function, that is, its Fourier transform in the Cartesian coordinates is such that it vanishes outside a bounded region. Then it is possible to reconstruct f from its discrete samples. The same is true for f 's partial derivatives with respect to x and y , the functions obtained by f 's products and sums with other bandlimited functions, etc.

Assume that the local neighbourhood is such a bandlimited function. Then for sufficient dense discretization

$$z_1 = \sum_j (f_{xj} + if_{yj})^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_j(\bar{r}) \exp[-i2 \arg\left(\frac{\partial \xi}{\partial x} + i\frac{\partial \xi}{\partial y}\right)] dx dy \quad (25)$$

is evaluated from the samples f_{xj} and f_{yj} . The latter are the values of the functions $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at the discrete image position \bar{r}_j . The image given by $(f_{xj} + if_{yj})^2$ will be referred to as the *partial derivative image*. The analytic function μ_j will be called the *interpixel function* since it is the function which makes it possible to evaluate values between the image pixels. It is the function obtained by inverse Fourier transforming a region function which is 1 inside of a region and 0 outside of it. The region itself is the region in which the Fourier transform of the bandlimited function $(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})^2$ does not vanish. Even if a gaussian is not an ideal inter pixel function in the strict sense it is proposed to utilize as μ_j to obtain reasonable filter sizes.

Consider (25), which is a convolution with a filter with the coefficients:

$$w^j = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_j(\bar{r}) \exp[-i2 \arg\left(\frac{\partial \xi}{\partial x} + i\frac{\partial \xi}{\partial y}\right)] dx dy. \quad (26)$$

The filter coefficients decrease rapidly as $\|\bar{r}_j\|$ become large when μ_j is chosen as:

$$\mu_j(\bar{r}) = \exp(-\beta\|\bar{r} - \bar{r}_j\|^2) \exp(-\alpha\|\bar{r}\|^2). \quad (27)$$

Here α controls the size of the neighbourhood and β controls the low pass character of the filter. The image should be multiplied by the gaussian term containing α to define a neighbourhood, but for convenience it is incorporated to the interpixel function, resulting in the same effect.

Since $z_2 = z_1/[E(\theta_{max}) + E(\theta_{min})]$ and the denominator is independent of the chosen coordinate system it is easy to obtain z_2 from z_1 . We will not consider the derivation of the error sum in detail. We will only mention that it is a gaussian filtering of $\|\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\|^2$ and refer to [2] for further details.

4 Applications and experiments

To apply the method derived above, a model for the symmetry in terms of harmonic functions and a partial derivative image,

$$(f_{xj} + if_{yj})^2 \quad (28)$$

are required. A special technique is not required to obtain f_{xj} and f_{yj} . Many methods exist in the literature for this purpose. However an efficient and easily implemented one is to convolve the image by the partial derivatives of a gaussian even though this has some drawbacks. In the experiments we have made this choice to obtain the partial derivative image.

For test purposes we define a one dimensional function g as

$$g(x) = (1 + \cos x)/2 \quad (29)$$

which is positive. The parameter x of this one dimensional function is replaced by $a\xi + b\eta$ to illustrate the different symmetries given below.

1. The function $\log z$, except for the origin, is analytic and single valued if one defines the *principal branch* as the value set. Since the imaginary part of an analytic function is the conjugate harmonic function of its real part, then

$$\log z = \ln r + i\varphi \quad (30)$$

where $r = |z|$, $\varphi = \arg(z)$ and z is the complex variable $x + iy$. Thus

$$\begin{aligned} \xi &= \ln r \\ \eta &= \varphi \end{aligned}$$

is obtained. The only singularity is at the origin and does not cause any problem since (21)-(23) are not affected by the values of the integrands at enumerable points (Lebesgue integrals, [12]).

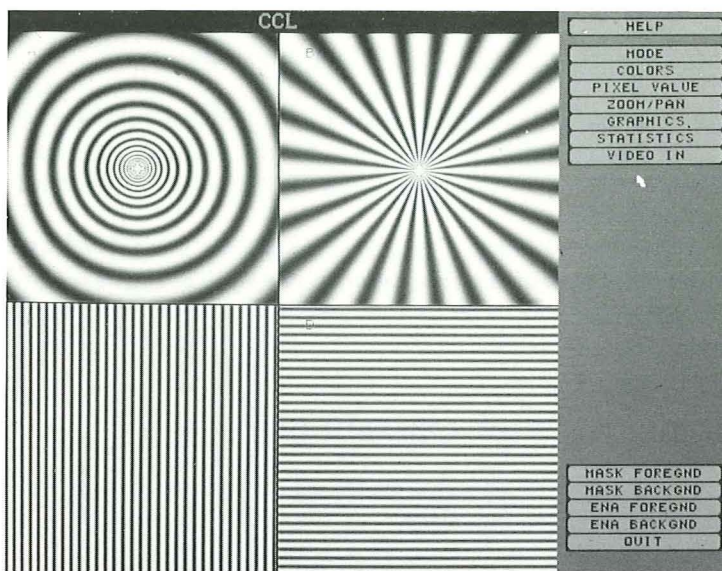


Figure 2: The figure illustrates the iso-gray value curves of a)-b) the circular symmetric harmonic function pair c)-d) linear symmetric harmonic function pair.

Figure 2 a) illustrates a neighbourhood described by $g(a\xi)$ and Figure 2 b) illustrates $g(b\eta)$. According to the previous sections all neighbourhoods, with iso-gray values being $a\xi + b\eta$ with any real constants a and b , are included in the symmetry model associated with the coordinates $\xi = \ln r$ and $\eta = \varphi$. We will call this type of neighbourhood *circularly symmetric*. Some of the neighbourhoods in this model are given in Figure 3. Figure 7 illustrates the result of the convolution proposed by (25) with the filter coefficients v^j which are obtained by (26),

$$v^j = \int_0^\infty \int_0^{2\pi} \mu_j(\bar{r}) \exp(-i2\varphi) r dr d\varphi. \quad (31)$$

A plot of these complex valued coefficients is given in [2]. The intensity of the picture in Figure 7 is the certainty in the symmetry, while the colour indicates its orientation, that is, the double angle representation of the inclination angle of the line $a\xi + b\eta$.

The certainty at the borders of the test blocks is approximately half the amount of the block centers. Green colour represents circles, which is a representation of the angle 0 and is mapped to the integer 0, while red represents star-shaped neighbourhoods which is a representation of the angle π and is mapped to the integer 127. (See the scale at right in Figure 8 for angle representation.) These two neighbourhoods are represented by two different complex variables: the first one has the argument zero, the latter the argument π . The fan-shaped neighbourhoods

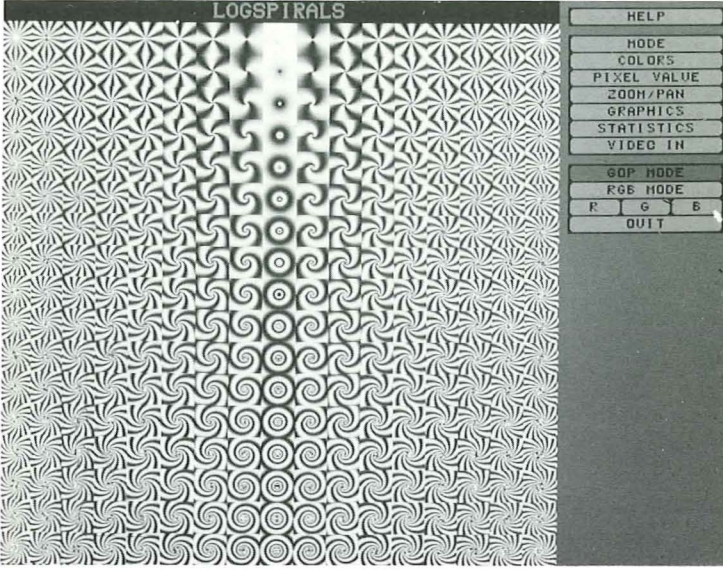


Figure 3: The figure illustrates some neighbourhoods included in the circular symmetry model. The orientation of the model corresponds to the “twistedness” of a neighbourhood and is well-defined.

are mapped to z_1 with $\text{im } z_1 > 0$ or $\text{im } z_1 < 0$ in accordance with the direction of rotation. The amount of “twistedness” that is double the inclination angle of $a\xi + b\eta$, is continuously mapped to the argument of z_1 , which is colour. The dark areas do have an orientation estimate but since the certainty level is too low in these areas, the visibility is automatically suppressed due to the low intensities. An experiment with natural images has been made. The circular symmetry detection is utilized in the classification tasks. Figure 5 illustrates a sea bottom image with sea anemones and their identification only by the circular symmetry image and the original image. Box classification is used.

2. Another pair of harmonic functions, the simplest one, is obtained by the analytic function z :

$$z = x + iy = \xi + i\eta \quad (32)$$

This is simply a model of neighbourhoods having edge or line forms. It delivers the orientation of these lines together with a certainty, [1]. The iso-gray value curves generating these symmetries, (*linear symmetries*), are given in Figure 2 c) and d), that is $g(a\xi)$ and $g(b\eta)$. The necessary filter turns out to be real and a gaussian, [1].

3. Choose the analytic function z^2 to generate the harmonic pair

$$z^2 = x^2 - y^2 + i2xy = \xi + i\eta \quad (33)$$

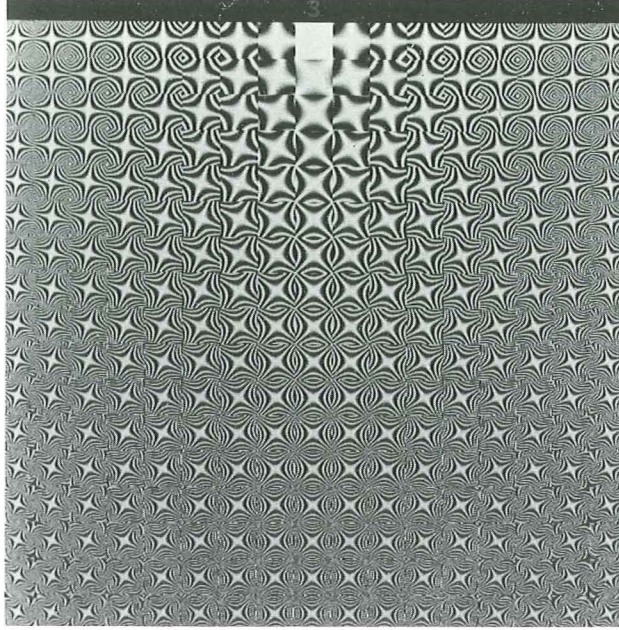


Figure 4: The figure illustrates some neighbourhoods included in the hyperbolic symmetry model. The orientation of the model corresponds to the orientation of the orthogonal asymptotes and is well-defined.

generating a symmetry which is given in Figure 5 a)-b). We will call this type of symmetry *hyperbolic symmetry*. The corresponding filter coefficients are given by:

$$v^j = \int_0^\infty \int_0^{2\pi} \mu_j(\bar{r}) \exp(i2\varphi) r dr d\varphi. \quad (34)$$

which is the complex conjugate of the filter obtained for circular symmetry, (31). A number of the neighbourhoods which are included in this model are given in Figure 4. The function generating these is $g(a\xi + b\eta)$ where $\xi = x^2 - y^2$ and $\eta = 2xy$. The neighbourhoods are generated by changing a and b as before. The orientation obtained in this symmetry obviously corresponds to the angle of rotation of the asymptotes. The two asymptotes are orthogonal and make this operator useful for detection of crosses in the natural images. The result of the symmetry detection is given in Figure 8.

4. Yet another symmetry will be generated by the real and the imaginary part of the analytic function \sqrt{z} , (the principal branch of the log is utilized):

$$\sqrt{z} = \sqrt{r} \exp(i\frac{\varphi}{2}) = \sqrt{r} \cos(\frac{\varphi}{2}) + i\sqrt{r} \sin(\frac{\varphi}{2}) = \xi + i\eta. \quad (35)$$

The symmetric neighbourhood pair generating this symmetry model is given in Figure 5 a)-b). The corresponding filter coefficients are given by:

$$v^j = \int_0^\infty \int_0^{2\pi} \mu_j(\bar{r}) \exp(-i\varphi) r dr d\varphi. \quad (36)$$

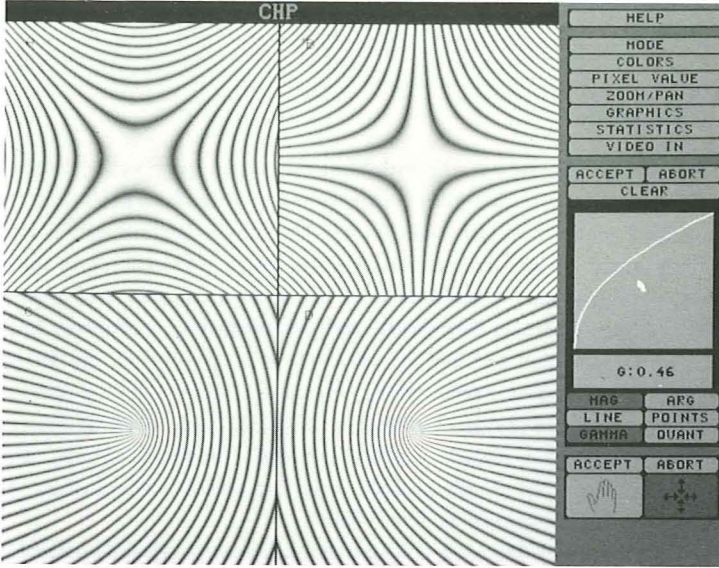


Figure 5: The figure illustrates the iso-gray value curves of a)-b) the hyperbolic symmetric harmonic function pair c)-d) parabolic symmetric harmonic function pair.

This operator is observed to be useful in finger print images to detect patterns having *parabolic symmetries*. The linear combinations of these coordinates, $a\xi + b\eta$, result in rotated versions of the parabolas shown in Figure 5 a)-b).

The list of symmetric patterns detectable by the formula (25) can be made long. It is sufficient to know one of the coordinate curves or its gradient, to be able to construct a symmetry model for the family of curves associated with this curve. In fact, since the the coordinates are assumed to be harmonic and one is the conjugate of the other makes it possible to use the theory developed for harmonic and analytic functions. For example if a coordinate curve function $\xi = u(x, y)$ is known on a circle with radius R and is harmonic within the circle then the other points on the disc are possible to obtain by *Poisson's formula*, [7]:

$$u(z_0) = \frac{1}{2\pi} \int_{|z|=R} \frac{R^2 - |z_0|^2}{|z - z_0|^2} u(z) d\varphi. \quad (37)$$

We have already used the fact that the real and imaginary parts of an analytic function are harmonic and the imaginary part is the harmonic conjugate of the real part.

The results of experiments with energy independent certainty, C_{f_2} , are deliberately left out since these are very similar to the results in [1], [2]. However we can comment on the implementation of equation (15) should be made. A threshold value should be specified for the entire image, telling the level below which numerator is considered zero. For values below this threshold the obtained certainties are set to zero. This effectively

eliminates the division by zero problem as well since the denominator is greater than the numerator.

5 Conclusion

A method to model symmetries of the neighbourhoods in gray value images is derived. It is based on the form of the iso-gray value curves. It is shown that it is possible to check in a special Fourier domain whether all iso-curves in a neighbourhood can be described by an a priori chosen harmonic function pair. However, it is also shown that the equivalent procedure can be performed in the spatial domain without performing a local Fourier transform, which is computationally demanding. For every neighbourhood a complex number is obtained through a convolution of a complex valued image with a complex valued filter. The magnitude of the complex number is the degree of symmetry with respect to the a priori chosen harmonic function pair. The degree of symmetry has a clear definition which is based on the error in the Fourier domain. The argument of the complex number is the angle representing the relative dominance of one of the harmonic pair functions compared to the other.

Since the a priori chosen harmonic function pair is such that the curves associated with these, intersect in right angles (except for some singularities), it is possible to represent a neighbourhood in this curvilinear coordinate system. And hence the problem becomes an edge detection problem in another coordinate system. The obtained orientation corresponds then to the orientation of an edge and therefore has a geometric interpretation. An advantage of using harmonic functions in symmetry descriptions, beside their ability to lead to easily implementable methods, is that one can use many powerful results obtained for analytic functions. The fact that the sum of the maximum and minimum errors for a given neighbourhood is invariant under different tests of symmetry models has potential advantages which can be used to prefer one symmetry model to another in updating the description of the neighbourhoods.

Acknowledgements

I should like to thank the Computer Vision Laboratory staff at Linköping University, especially Prof. G. H. Granlund and Dr. H. Knutsson for inspiring discussions. Also the Swedish National Board for Technical Development is acknowledged gratefully for their financial support during the course of this project.

References

- [1] J. Bigün, G.H. Granlund "Optimal orientation detection of linear symmetry", First International Conf. on Computer Vision, London, June 1987. pp. 433-438
- [2] J. Bigün, "Circular symmetry models in image processing", Linköping Studies in Science and Technology thesis No:85, 1986.
- [3] J. Bigün, G.H. Granlund "Central symmetry modeling", Third European Signal Processing Conference, The Hague, sep 3-5 1986 pp. 883-886
- [4] J. Bigün "Optimal orientatation detection of circular symmetry parameters" submitted for publication. Internal Report LiTH-ISY october 1987.
- [5] R. O. Duda, P. E. Hart " Use of the Hough transformation to detect lines and curves in pictures" Comm. ACM 15, 1, January 1972, 11-15.
- [6] B. H. Ballard "Generalizing the Hough transform to detect arbitrary shapes" Pattern Recognition 13,2, 1981, 111-112.
- [7] L. V. Ahlfors, "Complex Analysis" McGraw-Hill, New york, 1966.
- [8] G.H. Granlund: "In Search of a General Picture Processing Operator", Computer Graphics and Image Processing 8, 155-173 (1978).
- [9] W.C. Hoffman "The lie algebra of visual perception" J.math. Psychol. 1966 3 p65-98
- [10] W.C. Hoffman "Higher visual perception as prolantation of the basic Lie transformation group" Mathematical Biosciences No 6 (1970), pp. 437-471
- [11] P. C. Dodwell "The transformation group model of visual perception" Perception & Psychophysics 34 (1) 1983, pp. 1-16
- [12] R. L. Wheeden, A Zygmund, "Measure And Integral" Marcel Dekker, Inc., Basel, 1977.

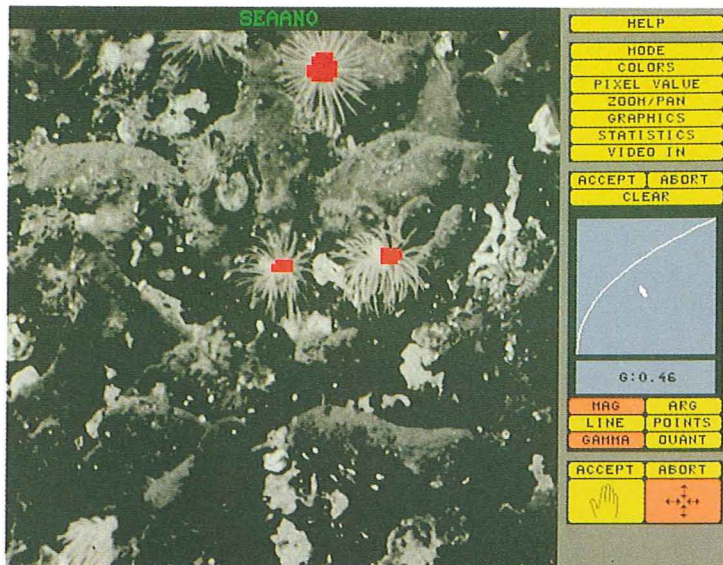


Figure 6: The image is a sea bottom photograph. The objective is to identify the sea anemones. The labels are obtained as a result of box classification.

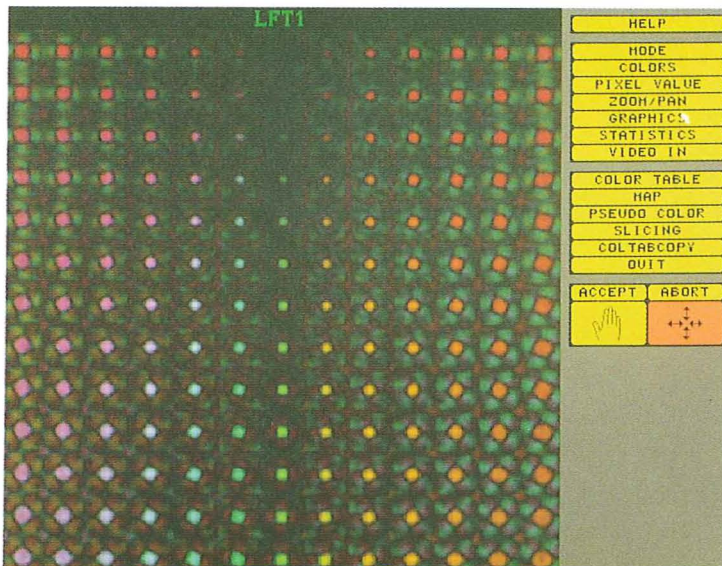


Figure 7: Result of circular symmetry estimation on the circularly symmetric neighbourhoods given in figure 3. The intensity indicates certainty while the colour indicates orientation.

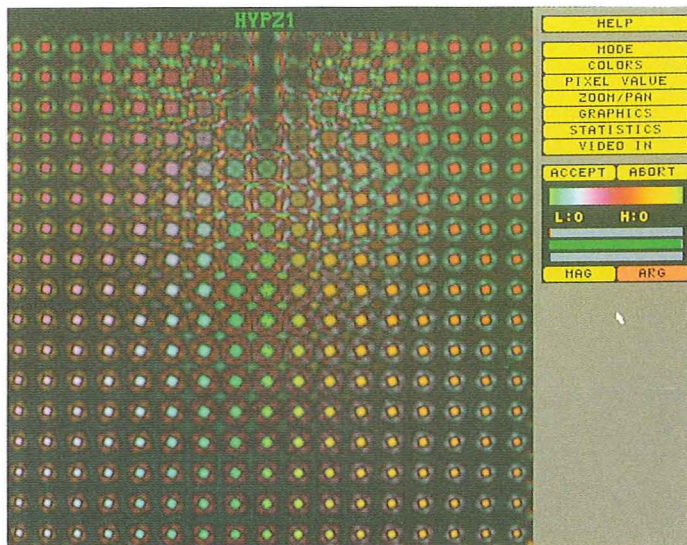


Figure 8: Result of a hyperbolic symmetry operation on the corresponding neighbourhoods given in Figure 4. On the right the colour scale used to map $[0, 2\pi]$ to the integers $\{0..255\}$. The mapping is modulo 2π .

Chapter V

OPTICAL FLOW BASED ON THE INERTIA MATRIX OF THE FREQUENCY DOMAIN

*So what is time? If nobody asks me I know.
If I try to explain it to someone who asks
me I do not know... ...Do you command
me to agree if someone says that time is a
motion of a body? Do not command! For
I hear that no body moves except in time:
this you tell me.*

St. Augustine

OPTICAL FLOW BASED ON THE INERTIA MATRIX OF THE FREQUENCY DOMAIN*

Josef Bigün, Gösta Granlund
Department of Electrical Engineering
Computer Vision Laboratory
Linköping University
S-581 83 Linköping Sweden

Abstract

An algorithm which utilizes the behaviour of the frequency domain to obtain the optical flow is presented. The solution is based on the inertia matrix of the Fourier domain corresponding to small regions of the observed visual field. It gives the optimal velocity vector for the regions together with a confidence measure for the estimates. Although the method is Fourier domain based, all computations are carried out in the spatio-temporal domain as filterings with separable filters. The frequency sensitivity associated with both spatial and temporal domain is possible to control by utilizing different filters, which is a property of human perception of motion according to current understanding of neurophysiology and psychophysics of motion.

1 Introduction

In recent years an increasing amount of the computer vision research effort has been devoted to motion. This is partly because of the powerful computers that are available. It is also believed that many important clues about the environment can be obtained by using motion analysis of each region of the visual field. One of the tools developed is called optical flow and refers to the two-dimensional velocity field of the visual field. There are gradient based approaches as well as other spatio temporal filter response approaches to solve the optical flow problem, [5], [4], [6]. An extensive list of references relating to the motion analysis of images can be found in [7].

In the following we will report about a gradient based method which is actually an application of the principles derived in [1]. The method can be thought of as Fourier domain based, even though all computations are finally carried out in the spatio-temporal domain. The obtained velocities are determined by a closed solution of an eigenvalue

*PROCEEDINGS OF SSAB SYMPOSIUM ON IMAGE PROCESING, Lund, March 10-11, 1988.

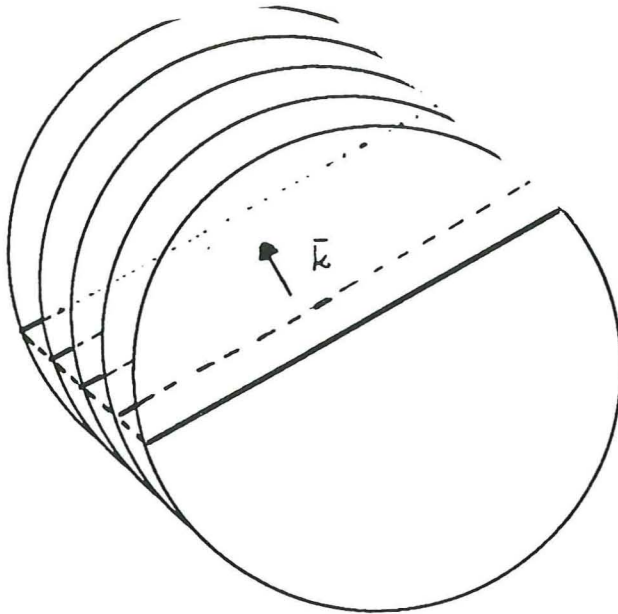


Figure 1: The figure illustrates a small region of the visual field in which a translation of a line occurs.

problem associated with the scatter matrix of the frequency domain. This guarantees the optimality of the velocities found. The differences in the eigenvalues are used as statements of certainty. The next section discusses the formulation of the optical flow problem and its relation to the frequency domain. Moreover the solution of the problem in terms of the inertia matrix of the frequency domain is proposed. The section on *Experimental results* reports about an implementation of the algorithm and the experimental results of an image of a rotating fan with different rotation frequencies. The algorithm turns out to be implementable with a relatively small effort, using the separability of the filters and the closed solution of the eigenvalue problem.

2 Formulation of the problem

Let $f(x, y, t)$ represent the local intensity function registered by a visual system. The parameters x and y represent the spatial coordinates of the visual system while t represents the time variable of the system. By the local intensity function or a neighbourhood we mean the product of a smooth window function and the intensity function. Assume that the local intensity function is generated by a line translated in a constant direction and velocity according to Figure 1.

The local intensity function, f , then describes a plane. The velocity component of translation parallel to the line is not possible to obtain. This natural phenomenon is often referred to as the aperture problem. The velocity component of translation

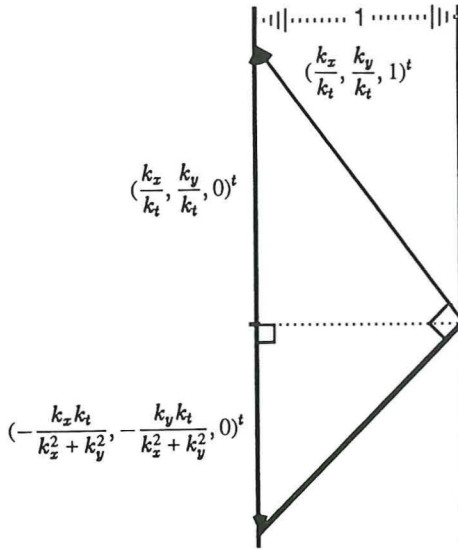


Figure 2: The figure illustrates the geometry used to derive the 2D velocity vector from the 3D normal vector associated with the translation.

perpendicular to the line is implicitly given by the normal vector of the plane. If this vector is given by $\bar{k} = (k_x, k_y, k_t)^t$ then using geometry as indicated in figure 2, the velocity, $\bar{v} = (v_x, v_y)^t$, of translation perpendicular to the line is obtained by the formulas:

$$v_x = -\frac{k_x k_t}{k_x^2 + k_y^2} \quad (1)$$

$$v_y = -\frac{k_y k_t}{k_x^2 + k_y^2}. \quad (2)$$

As expected, the velocity vector \bar{v} does not change when \bar{k} is replaced by $-\bar{k}$.

When the local intensity function f is generated by translation of a line the formulas (1) and (2) can be used to obtain the local velocity field. By the local velocity field is understood a “representative” velocity of a neighbourhood. However, \bar{k} is not known in general, and an immediate way to find an estimate of it is to estimate \bar{k} with the gradient of the scalar field f . But this intuitive estimation is not always meaningful for a general local intensity function, since we can not expect it to be a perfect plane. Moreover, neither is it clear what a representative velocity of such a neighbourhood is. If the local image degrades severely from a perfect plane, say that it is like a cloud, then motion no longer consists of translation of a single line. Obviously, when the local optical flow consists of a perfect plane, the estimation $\bar{k} = \nabla f$ works. But what do (1) and (2) estimate in terms of physical observations when f is no longer a plane? The question of when to trust such an estimation of the velocity, and when not to, has to

be raised. This implies that an estimation of the velocity of a neighbourhood needs a reliable certainty measure informing about the credibility of the estimate.

In the following we will model a general local intensity image using sine and cosine functions. This is equivalent to the Fourier transform of the local image. Without proving it, see [1], we will state that if the graph of the iso-values of the intensity function f consists of parallel planes, then the energy is always concentrated to an axis through the origin in the Fourier transform domain. Any local intensity function corresponding to the visual world has a representation in the frequency domain. Thus one can reformulate the problem of finding a representative velocity for the local image as the problem of fitting an axis through the origin of the Fourier representation of the local image. The reason for reformulation of the optical flow problem is of course to concretize the rather vague concept of a "representative" velocity of a local image. We will fit an axis to the Fourier domain in the least square error sense, which corresponds to solving an eigenvalue problem in the spatial domain. The 3X3 matrix involved in this eigenvalue problem is the inertia matrix of the Fourier transform of the local intensity function, and is found to be

$$J = I\text{trace}(A) - A$$

with

$$A = \begin{pmatrix} \int (\frac{\partial f}{\partial x})^2 & \int \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} & \int \frac{\partial f}{\partial x} \frac{\partial f}{\partial t} \\ \int \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} & \int (\frac{\partial f}{\partial y})^2 & \int \frac{\partial f}{\partial y} \frac{\partial f}{\partial t} \\ \int \frac{\partial f}{\partial x} \frac{\partial f}{\partial t} & \int \frac{\partial f}{\partial y} \frac{\partial f}{\partial t} & \int (\frac{\partial f}{\partial t})^2 \end{pmatrix} = \int (\nabla f)(\nabla f)^t. \quad (3)$$

The least square error axis in the Fourier domain is given by the eigenvector corresponding to the least eigenvalue of J . In fact, given J for a neighbourhood f , the error or the inertia with respect to any axis $\bar{k} \in R^3$ is given by $E(\bar{k}) = \bar{k}^t J \bar{k}$. There is no risk of $E(\bar{k})$ being negative, since J is positive semi-definite. Thus the eigenvalues of J are the errors obtained when the corresponding eigenvector is fitted to the Fourier domain. If the least eigenvalue is zero with multiplicity one, then one can be sure that the local intensity function corresponds to a translation of an edge. Another way to interpret the eigenvalues is to consider the difference between the second least and the least eigenvalue. If this difference is small, then one can say that the error obtained by an alternative axis is not much worse than the optimal one. Consequently, when this difference is large, one cannot only say that the axis found is optimal, but also that the energy in the Fourier domain is concentrated on the optimal axis.

Label the eigenvalues of A in ascending order with λ_j as:

$$\lambda_2 \geq \lambda_1 \geq \lambda_0 \geq 0$$

and the corresponding eigenvectors by $\bar{u}_2, \bar{u}_1, \bar{u}_0$. Since $A\bar{u}_j = \lambda_j\bar{u}_j$

$$J\bar{u}_j = (I\text{trace}(A) - A)\bar{u}_j = \lambda'_j\bar{u}_j \quad (4)$$

is obtained, where λ'_j is given by

$$\lambda'_j = (\lambda_2 + \lambda_1 + \lambda_0) - \lambda_j. \quad (5)$$

Here the relation

$$\text{trace}(A) = A_{11} + A_{22} + A_{33} = \lambda_2 + \lambda_1 + \lambda_0$$

is utilized. Thus the matrices J and A share the same eigenvectors with the corresponding eigenvalues being the primed and unprimed lambdas satisfying equation (6). Finding the eigenvector corresponding to the least eigenvalue of J is the same as finding the eigenvector corresponding to the largest eigenvalue of A . The certainty measure representing the reliability of the obtained motion direction is proposed to be

$$C_{f1} = \lambda'_1 - \lambda'_0 = \lambda_2 - \lambda_1. \quad (7)$$

The images are represented in computers by a discrete sample of data. Assuming that the sampling is dense enough, one can reconstruct the local image as well as its partial derivatives from the discrete data. The reconstructibility assumption is that the images as well as the products of their partial derivatives are possible to reconstruct from their discrete samples by means of an interpolation function. For example the matrix element a_{33} is possible to compute by means of the reconstructibility assumption according to:

$$\begin{aligned} a_{33} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial f}{\partial t}\right)^2 dx dy dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_j f_{tj}^2 \mu_j(x, y, t) dx dy dt \\ &= \sum_j f_{tj}^2 w_j. \end{aligned}$$

Here f_{tj} is the value of $\frac{\partial f}{\partial t}$ at the discrete point \bar{r}_j , and μ_j is the interpixel or the interpolation function associated with the reconstructibility assumption. The other elements of A can be evaluated in a similar way. Moreover all matrices A_n associated with the neighbourhoods around the discrete image points \bar{r}_n can be evaluated through simple convolutions, which turns out to be simple averaging, [1]. For compactness in the frequency and the spatial domain, the μ_j 's are chosen as gaussians resulting in smooth filter coefficients. To be able to obtain the elements of the scatter matrix, A , the gradient image

$$\nabla f(x_j, y_j, t_j) = (f_{xj}, f_{yj}, f_{tj})^t \quad (8)$$

is necessary. If the gradient image is known, then the scatter matrix associated with a local image is the local average of $\nabla f(\nabla f)^t$:

$$A = \langle [\nabla f(x_j, y_j, t_j)][\nabla f(x_j, y_j, t_j)]^t \rangle. \quad (9)$$

The gradient image ∇f_j is of course also possible to estimate through convolutions with partial derivative filters obtained through the same technique, resulting in discrete values of the partial derivatives of three-dimensional gaussians. In the experiments below this technique is utilized. It should be mentioned that ∇f_j and A_j can be obtained using recursive filters as well as FIR filters. However, we have only considered FIR filters in our preliminary implementation.

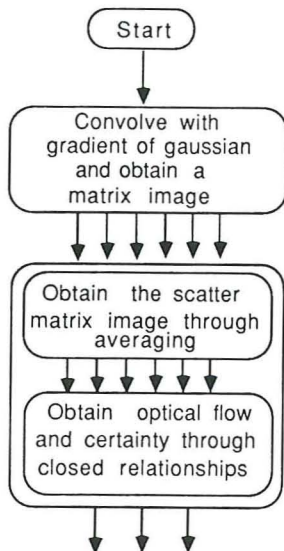


Figure 3: The flow chart illustrates the implemented algorithm.

3 Experimental results

To test the proposed method we have implemented the scheme given in Figure 3 on a GOP-300.

1. The first step of this implementation starts with convolving the discrete intensity function $f(x_j, y_j, t_j)$ with three partial derivative filters. The filter coefficients in the gradient filter are chosen as $\nabla g(\bar{r}_j)$ where

$$g(\bar{r}_j) = \exp(-\beta(x_j^2 + y_j^2 + t_j^2)). \quad (10)$$

Here the spatio-temporal variables x_j, y_j, t_j are assumed to be scaled appropriately to accomplish a desired relation between the spatial and temporal variables. Then $[\nabla f_j][\nabla f_j]^t$ is obtained, where ∇f_j is the three-dimensional vector of the three partial derivative convolutions. Since the result is a symmetric matrix of dimension three, only six of the matrix elements need to be stored. The spatio-temporal frequencies in which the algorithm produces reliable results are controlled by the size of the gaussian in equation (10) through β . In the experiments the convolution with the gradient filter is implemented as a FIR filter. The three partial derivative filters are the truncated versions of the gradient of a gaussian. The truncation is such that at the borders of the filters the magnitude of the filter coefficients is less than 1 % of the maximum magnitude of the filter coefficients. In the following results the used filter was 5x5x5. The fact that the gaussian filters are separable is utilized, resulting in a speed-up of the algorithm.

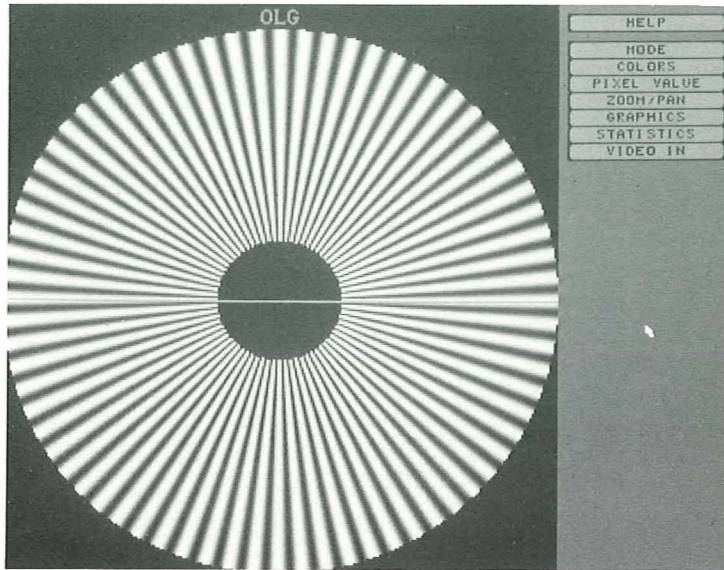


Figure 4: The figure illustrates the rotated fan. All black and white figures have an intensity dynamic range consisting of the integers in the interval $[0,255]$.

2. Convolve the matrix image obtained in the previous step by a gaussian and obtain the scatter matrix, A_j , of every point. The size of the gaussian controls the size of the neighbourhood for which a representative velocity is to be found. In the experiments the size of the gaussian was $7 \times 7 \times 7$. The separability is utilized as before. The eigenvector corresponding to the largest eigenvalue is utilized as \bar{k} in equations (1) and (2) to obtain the local velocities. The certainties in these velocity estimations are obtained by evaluating the difference of the eigenvalues of the scatter matrix according to (7). Without writing down the solution it should be mentioned that the eigenvalues and eigenvectors for 3×3 matrices are possible to obtain in closed forms. This enables the algorithm to find the optimal orientation without implicit optimization techniques as is the case when a set of responses of the Gabor filters are used, [4].

As a test image we have used the one in Figure 4 which illustrates a fan produced according to the formula:

$$[1 + \cos(72 \tan^{-1}(x_j, y_j))]/2. \quad (11)$$

It is rotated and 64 frames of the motion are observed under slow-down of the rotation, to study the response in different spatio-temporal frequencies. If the rotation velocity is kept constant, the intensity at any particular spatial position on the fan frames describes a sinusoid. The slow-down of the rotation is such that the sinusoids in

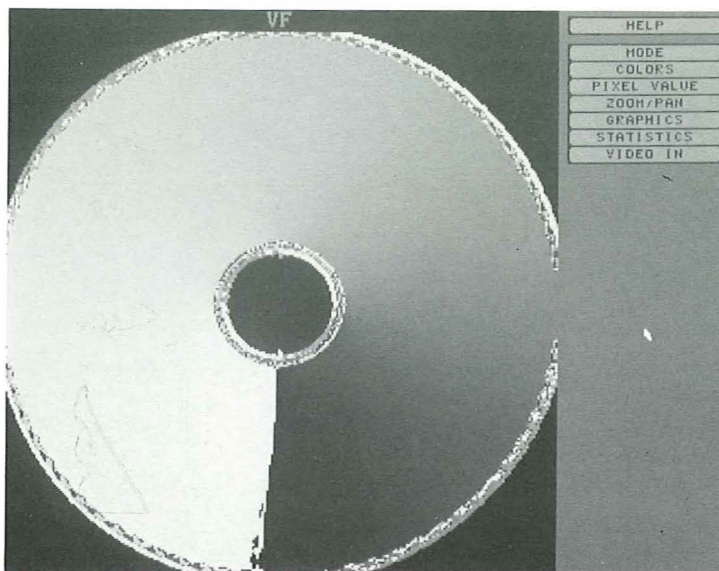


Figure 5: The figure illustrates the direction of velocity estimation.

the time direction have a frequency decreasing as $o(1/t)$, chosen to match the natural spatial frequency slow-down, $o(1/r)$.

Figure 14 illustrates the obtained velocities for one frame. The intensities are proportional to the obtained velocities. Figure 5 illustrates the direction of the velocities.

The interval $[0, 2\pi]$ is mapped to integers in the interval $[0, 255]$, which represent the available intensity values in all images. Figure 6 illustrates the obtained certainties for the same frame.

Figure 7 illustrates the cross-section of 64 observed frames along the horizontal midline of the frames. The slow-down of the motion is visible and the decrease in frequency is illustrated in Figure 10, which is the graphical illustration of the intensity variation of a point along the time axis. Figure 15 illustrates the velocities along the same cross-section. The intensities represent the magnitudes of the velocities. Figure 8 illustrates the direction of the velocity. The intensities are proportional to the angles $[0, 2\pi]$. The certainty image which is illustrated by Figure 9 predicts the confidence in the velocity estimations in Figures 15 and 5. It is clearly visible that the results are not reliable at very high spatial frequencies, which is a consequence of insufficient sampling density. This certainty image can also be used as a tool to determine the frequency sensitivity band of the filters. The iso-values of 9 describe curves like c/r . This indicates that the product of temporal and spatial frequency is an important factor which influences the frequency sensitivity of the used filters.

Figure 11 illustrates the velocity profile along a horizontal line in Figure 14 which clearly shows a linear behaviour as it should. The irregularities at the boundaries are due

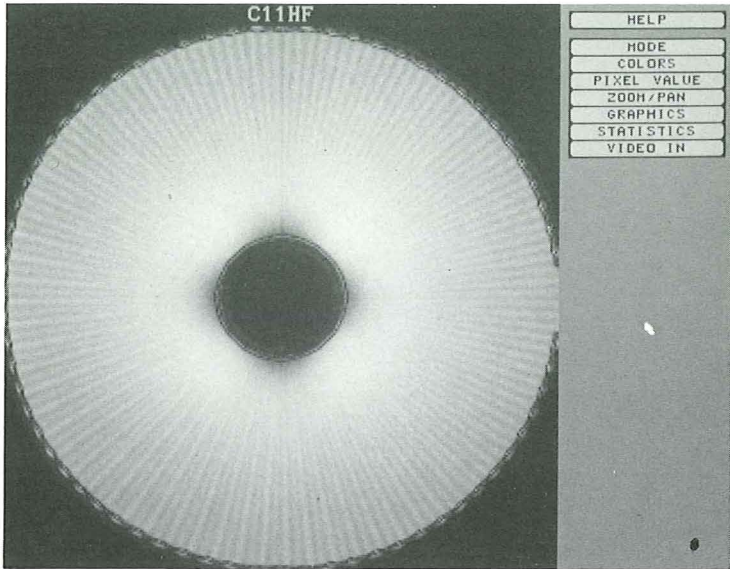


Figure 6: The figure illustrates the certainty estimation.

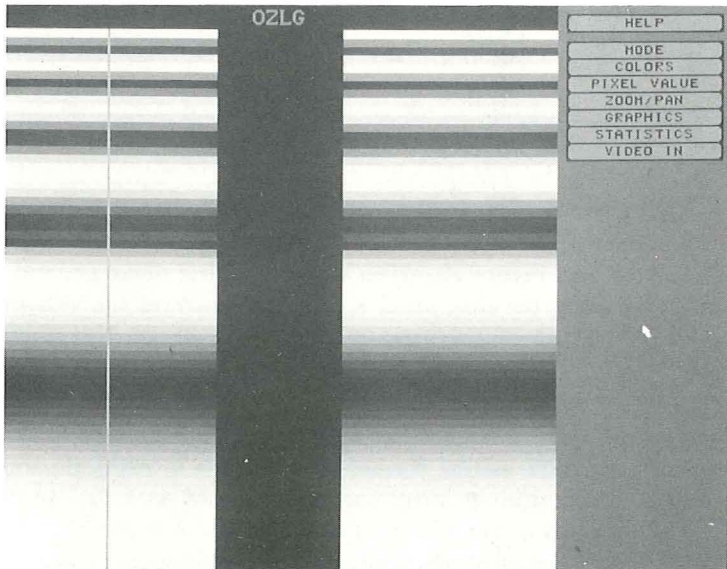


Figure 7: The figure illustrates the obtained horizontal cross-section of the fan along the time axis.

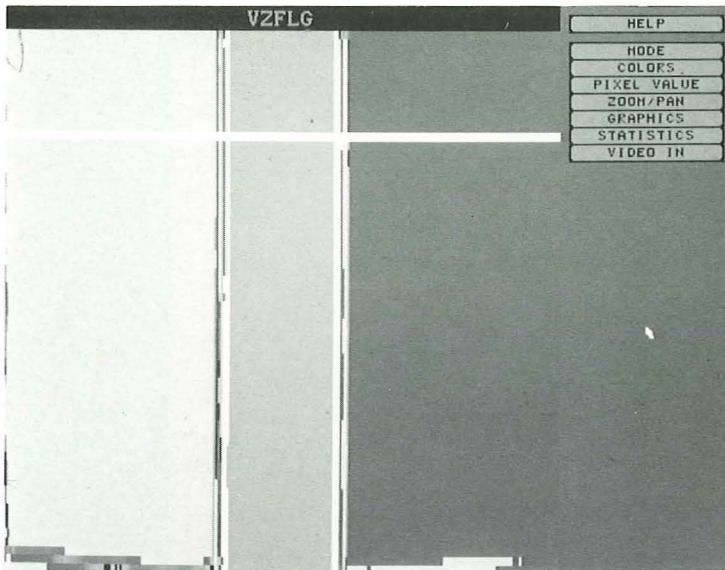


Figure 8: The figure illustrates the obtained horizontal cross-section of the rotation direction along the time axis.

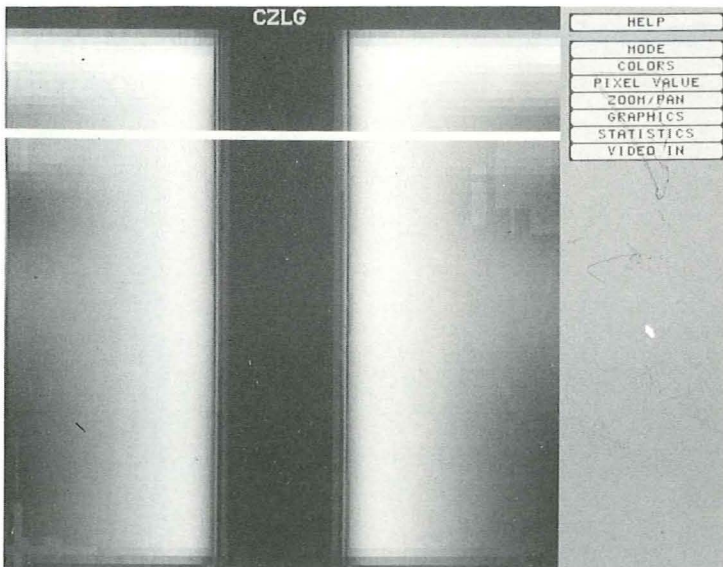


Figure 9: The figure illustrates the obtained horizontal cross-section of the certainty of velocity vector along the time axis.

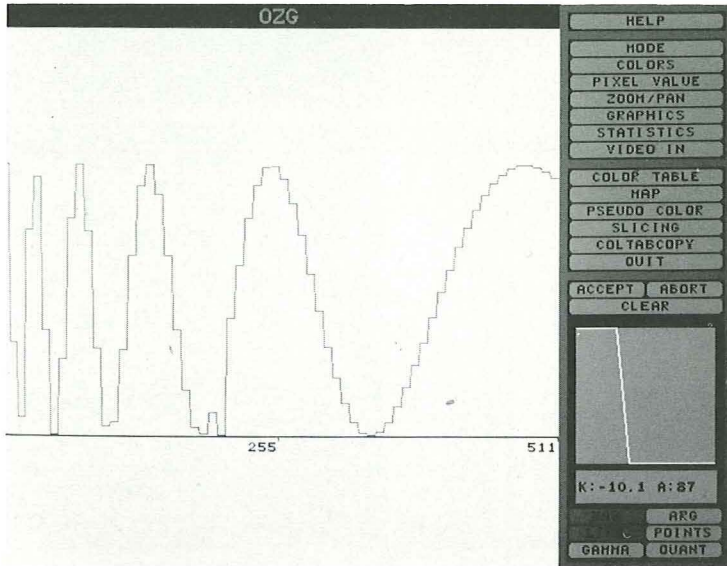


Figure 10: The figure illustrates the profile of the figure 7 along the indicated line.

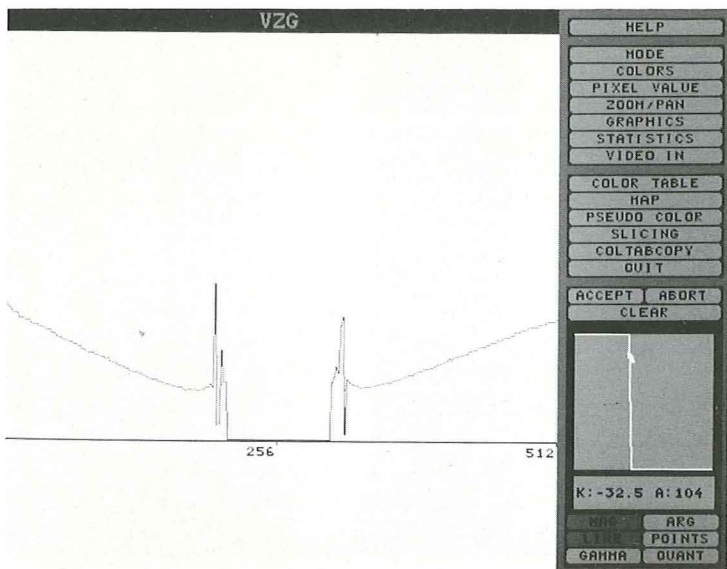


Figure 11: The figure illustrates the velocity estimation along the line in figure 15.

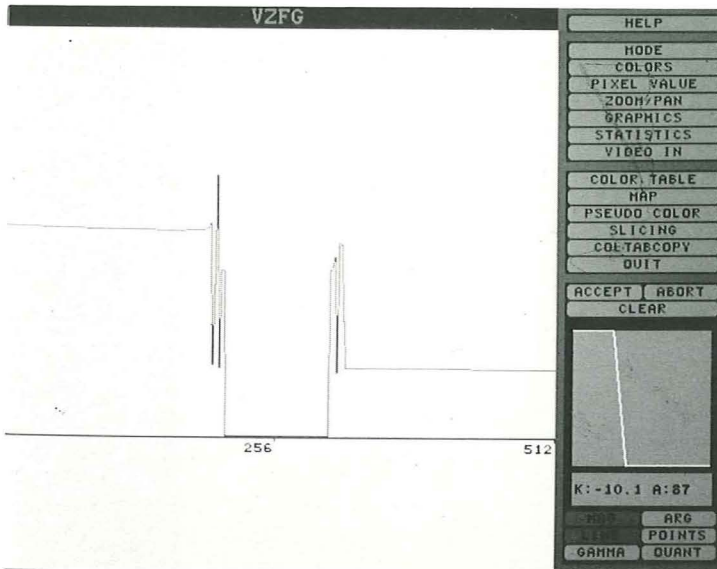


Figure 12: The figure illustrates the direction of velocity estimation along the line in figure 8.

to insufficient sampling density as well as to the discontinuities at the boundaries and are indicated to be less relevant by the certainty image. Figure 13 shows the certainty profile corresponding to Figures 14 and 12, of which the latter describes the direction of the velocity. The direction is described by the argument angle of the velocity vector and attains two constant values differing by half of the representation dynamic range, 128 which corresponds to π . The source of the error at the boundaries is the same as in Figure 11. The optimal spatial frequency response at a given temporal frequency is visible as a top in Figure 13, the position of which can be controlled by changing the size of the gradient filter.

4 Conclusion

The presented experiments show that frequency domain based minimization of the error function or the moment of inertia gives good estimations of the optical flow together with a confidence measure. The presented algorithm, which is entirely spatial domain based, implicitly performs the least square fitting through averaging and eigenvalue analysis. The averaging over a functional of the gradient is a property which is shared by symmetry descriptions other than linear symmetric, [2], [3]. This is interesting since the optical flow problem is considered as a part of symmetry description of a neighbourhood. It can be shown that the obtained estimate of the the optimal axis is also the normal of an optimal plane parallel to which a translation of the entire

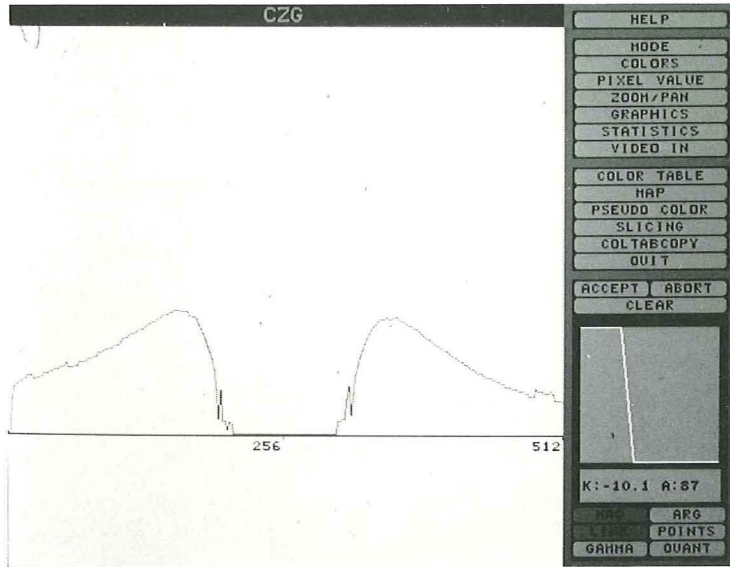


Figure 13: The figure illustrates the certainty estimation along the line in figure 9.

neighbourhood can be performed with minimal error in the L^2 sense. The closed form solutions of the eigenvalue problem and the separability of the filters contribute to a decrease in the execution time compared to the iterative methods and non-separable filters.

Acknowledgements

The financial support of the Swedish National Board for Technical Development is gratefully acknowledged.

References

- [1] J. Bigün, G.H. Granlund "Optimal orientation detection of linear symmetry", First International Conf. on Computer Vision, London, June 1987. pp. 433-438
- [2] J. Bigün "Optimal orientatation detection of circular symmetry parameters" submitted for publication. Internal Report LiTH-ISY october 1987.
- [3] J. Bigün "Pattern recognition by detection of local symmetries", Proceedings of Pattern Recognition in Practice iii, Amsterdam, May 18-20, 1988; Elsevier Science Publishers.
- [4] D. J. Heeger "Optical Flow from Spatiotemporal Filters", First International Conf. on Computer Vision, London, June 1987. pp. 181-190
- [5] Ballard, D.H., Brown, C.H. "Computer Vision", Prentice Hall, 1987, NJ
- [6] E. H. Adelson, J. R. Bergen "Spatiotemporal energy models for the perception of motion" Jour. of the Opt. Soc. of America A, 1(2) 1985 pp 284-299
- [7] A. Rosenfeld "Survey, Picture Processing: 1984" Computer Vision, Graphics, and Image Processing **30**, pp. 189-242 (1985), section (G) of the paper.
- [8] H. Knutsson "A tensor representation of 3-D structure" 5.th workshop on multidimensional signal processing. Noordwijkerhout, Netherlands september 14-16, 1987.

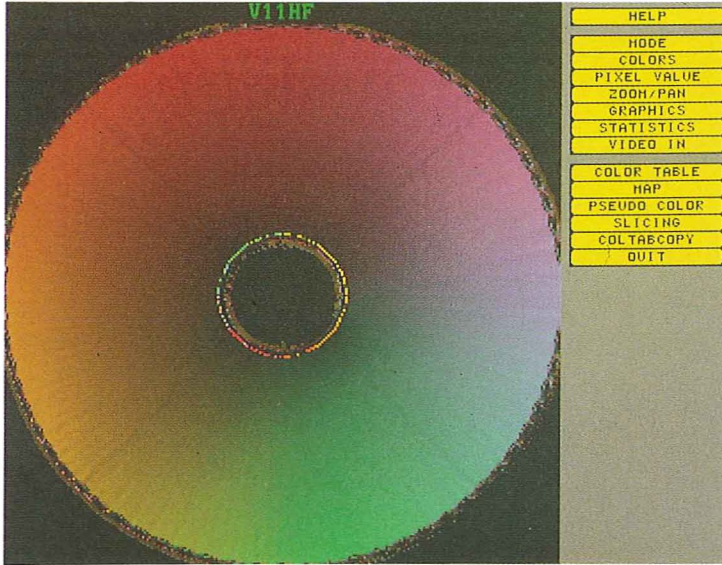


Figure 14: Optical flow estimation for one frame. Intensity represents the magnitude of velocity, hue represents the direction of velocity.

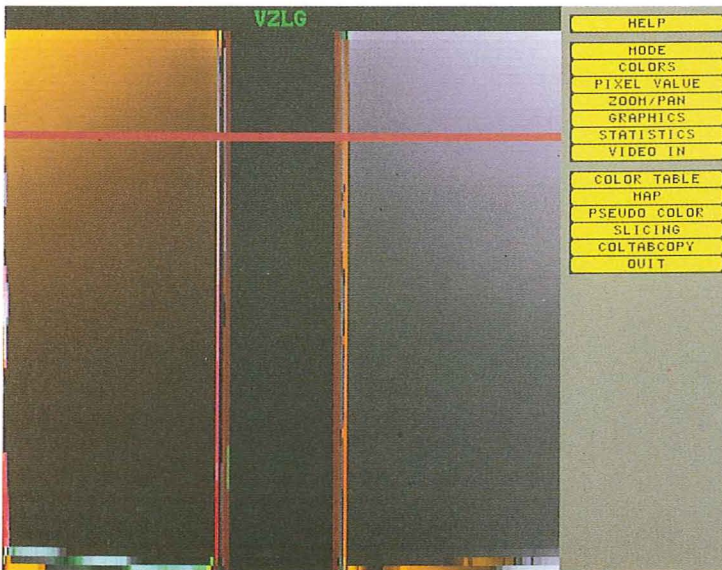


Figure 15: Horizontal cross-section of the fan velocity along the time axis.