(C) Computer Vision and Image Understanding 68(3):pp.290-307:1997

Pattern recognition in images by symmetries and coordinate transformations

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

Abbreviations

CT:	Coordinate transformation
FIR:	Finite impulse response
HFP:	Harmonic function pair
ILS:	Infinitesimal linear symmetry
GHT:	Generalized Hough transform
GLS:	Generalized linear symmetry
SNR:	Signal to noise ratio

Key words: deformation, coordinate transformation, Lie groups, Lie operators, symmetry, continuous symmetry, chirality, pattern recognition, image analysis

Abstract

A theory for detecting general curve families by means of symmetry measurements in the coordinate transformed originals is presented. Symmetries are modeled by iso-gray curves of conjugate harmonic function pairs which also define the coordinate transformations. Harmonic function pair coordinates render the target curve patterns as parallel lines, which is defined here as linear symmetry. Detecting these lines, or generalized linear symmetry fitting as it will be called, corresponds to finding invariants of Lie groups of transformations. A technique based on least square error minimization for estimating the invariance parameters is presented. It uses the Lie infinitesimal operators to construct feature extraction methods that are efficient and simple to implement. The technique, which is shown to be an extension of the generalized Hough transform, enables detection by voting and accumulating evidence for the searched pattern. In this approach complex valued votes are permitted, where the phase of the vote identifies the member of the family of patterns that are detectable. Experimental results illustrating the theory and its application to real as well as synthetic images are presented.

1 Introduction

Traditionally, edge images of the original image have been used to detect symmetries or complex structures. From a theoretical point of view, the Generalized Hough Transform (GHT), [27, 16], has been the most widely used in addressing the general curve detection problem in images. It uses edges and their orientations in order to conclude whether or not a template/curve exists in an image. Apart from being computationally expensive, its inability to cope with gray level images, implies difficulties for its use in image processing chains e.g. multiscale decompositions in texture segmentation. Nevertheless, it remains a conceptually important tool to which, the symmetry techniques developed here will be related. Other methods of capturing the shape information include moment invariants [28], Fourier descriptors [44, 20, 3] and morphological shape descriptors [37].

Encouraged by support of the relationship of Lie operators to the biological vision systems on the functional level [25, 26, 42, 36], pattern recognition and measurement tools based on coordinate transformation (CT) have been emerging. The latter is fully equivalent to the study of Lie groups. The application of CT theory to the modelling of general shapes has been studied in [10, 18, 23, 19]. Also, the work in [35, 43, 39, 40, 22] is related to CT theory although their emphasis is not on the theory. Motion parameter detection can also be formulated by the Lie groups of infinitesimal transformations, [29, 17, 6, 15].

The theory developed in Kass and Witkin, [30], is related to that of Bigün, [4, 11, 10]. These two independently developed theories share some fundamentals, but their tools and conclusions address different applications i.e. enhancement and reconstruction versus pattern recognition. Although some applications and tools of the theory of Bigün have been published earlier [8, 6, 7], a comprehensive and accesible version of the work in [4, 11, 10], has been lacking. This paper addresses that need. Another related work using local orientation measurements in shape detection is the work of Granlund and Knutsson [31]. However, they do not provide a theoretical analysis.

The objective of this paper is to provide a general technique, for detection of patterns possesing linear symmetry, with respect to a known harmonic function pair (HFP). Since very complex patterns can be described by such CTs, the technique is a general toolbox for geometric

pattern detection. We also develop a unifying concept based on [10, 8, 6, 27, 16], for geometric shape feature computations based on symmetries in the coordinate transformed images. In this theory, detection of repetitive patterns (texture) as well as non-repetitive patterns i.e. any given prototype, or analytical curve such as circle, cross junction, spiral, pencil of lines, is possible using the same tool box with different filters. We present four experiments which use real data as support for the validation of the theory and synthetic data as for determining accuracy. The data in our experiments include i) aerial images for automatic texture segmentation ii) underwater images and iii) electronic circuit design photographes for general object recognition. Our investigation shows that detection of even complicated target objects is equivalent to a problem of symmetry detection in the HFP coordinate system.

Although the proposed technique is equivalent to finding symmetry axes in the HFP coordinate system, it does not perform CTs explicitly. Instead, Lie operators are applied to Infinitesimal Linear Symmetry (ILS) of images. This extracts the generalized linear symmetry (GLS) features.

In Section 2 we develop the concept of locally orthogonal CTs and pattern families generated by them. A theory is presented in Section 3 for using CTs in order to represent and detect curves, by means of linear symmetries. In Section 4, the solution of the problem and its properties are discussed. In Section 5 we present the generalized linear symmetry detection technique for the traditional image grid, i.e. original Cartesian coordinates. Experimental results are presented in Section 6. Finally, Section 7 contains discussion and conclusion.

2 Generating symmetric patterns by conjugate harmonic functions

In this section we define families of symmetric patterns which are modelled and quantified. We illustrate symmetric patterns generated by HFP's through four examples. As will be seen, the local orthogonality of the HFP's simplifies the algorithms for estimating the symmetry parameters.

Let $\xi(x, y)$ be a *harmonic function,* that is, its partial derivatives of the first two orders are continious and it satisfies the *Laplace equation:*

$$\Delta\xi = \frac{\partial^2\xi}{\partial x^2} + \frac{\partial^2\xi}{\partial y^2} = 0. \tag{1}$$

Due to the linearity of Laplace's equation, linear combinations of harmonic functions are also harmonic. If two harmonic functions ξ and η satisfy the *Cauchy-Riemann equations:*

$$\frac{\partial\xi}{\partial x} = \frac{\partial\eta}{\partial y}, \quad \frac{\partial\xi}{\partial y} = -\frac{\partial\eta}{\partial x} \tag{2}$$

then η is said to be the *conjugate harmonic function* of ξ. Equivalently the pair (ξ, η) is said to be a HFP.

The imaginary part of any analytic function is the conjugate harmonic function of the real part. Without loss of generality we can assume both ξ and η to be single valued by imposing proper restrictions. Then by definition (2) a curve pair,

$$\xi_0 = \xi(x, y) \tag{3}$$

$$\eta_0 = \eta(x, y) \tag{4}$$

has orthogonal gradients at the same point. For non-trivial $\xi(x,y)$ and $\eta(x,y)$, (3-4) define a CT which is invertible almost everywhere.

Let an image be represented by the real function $f_1(x,y)$. Another representation of the image f_2, can be obtained, by means of a CT using the HFP, $f_1(x(\xi,\eta),y(\xi,\eta)) = f_2(\xi,\eta)$. The term image will refer to both a sub-image and an entire image, depending on the context.

Definition 1 *The image $f(\xi,\eta)$, is said to be linearly symmetric in the coordinates (ξ,η) if there exists a one-dimensional function g such that $f(\xi,\eta) = g(a\xi+b\eta)$ for some real constants, a and b. Here $\xi(x,y),\eta(x,y)$ is a HFP and the symmetry direction vector, (a,b), has its length normalized to unity, i.e. $\sqrt{a^2+b^2}=1$.*

The notion "linearly symmetric in (ξ,η)" is motivated by the fact that the iso-gray curves (or shortly iso-curves) of such images, are *parallel lines* in the ξ,η coordinate system. Also such images have a high concentration of their spectral power along a line through the origin.

Starting with the trivial, unity, transformation we give examples of CTs for pattern families which can be modeled and detected by the toolbox to be presented.

Example 1

$$w(z) = z = \xi(x,y) + i\eta(x,y) = x + iy$$

Since w is an analytic function in z, (x,y) is a HFP. For illustration, we let the one dimensional function g be

$$g(\tau) = (1+\cos\tau)/2 \tag{5}$$

while bearing in mind that g can be any 1-D function. The argument, τ, is replaced by $a\xi+b\eta$ to generate the family of iso-curves defined by this transformation. Figure 1 (top) illustrates the two basis patterns $g(\xi) = constant$, and $g(\eta) = constant$ respectively. Since g is general, not only parallel lines but also an image consisting of one edge or one line is part of this family. The lengths of the "parallel lines" in the image are formally infinite. We will see that this assumption is not a serious restriction even if the image is finite in size e.g. a neighborhood.

********** Figure 1 about here**********

Example 2

By using the same g as in Example 1, we can illustrate the transformation defined by,

$$w(z) = \log z = \xi(x,y) + i\eta(x,y) = \log\sqrt{x^2+y^2} + i\tan^{-1}(x,y)$$

which is analytic every where except at the origin. We assume the *principal branch* as the value set of w in order to avoid multiple valued functions. Figure 1 (bottom) shows the basis pair of this transformation. The local orthogonality can be seen by superimposing the two figures so that the origins coincide. The linear combinations of the basis pair, $a\xi + b\eta$, generate the family of "logarithmic spirals". Some members of this family are displayed in Figure 4. We note that the sign of $a \cdot b$ determines the chirality of the spirals, i.e. whether they are twisted to the left or to the right. Thus, by measuring the orientation angle of the vector (a,b) it is possible to tell apart left-handed and right-handed patterns, as well as whether it is circular or star shaped without actually knowing the exact form of the pattern, g, in advance.

Example 3

We use the analytic function z^2 to obtain the HFP, ξ, η

$$w(z) = z^2 = \xi(x,y) + i\eta(x,y) = x^2 - y^2 + i2xy \tag{6}$$

which are illustrated in Figure 2 (top). The generated pattern family, $a\xi + b\eta$, is given by Figure 6 and corresponds to rotated versions of a basis pattern. The asymptotes of the generated hyperbolic patterns are orthogonal and the orientation of the cross is given by the orientation of (a, b) which, as will be discussed further below, represents the orientation of a "cross".

********** Figure 2 **********

Example 4

The analytic function $\sqrt{z}$:

$$w(z) = \sqrt{z} = \xi(x,y) + i\eta(x,y) = \sqrt{r}\exp(i\frac{\varphi}{2}) = \sqrt{r}\cos(\frac{\varphi}{2}) + i\sqrt{r}\sin(\frac{\varphi}{2}). \tag{7}$$

generates the basis patterns which are illustrated in Figure 2 (bottom). The pattern family generated by this pair consists of rotated versions of one of the basis patterns, that is the rotation angle corresponds to the orientation of the vector (a, b).

3 Pattern recognition by linear symmetries

Here we briefly review the essential parts of the theory of the differential operators in order to detect symmetric pattern families. The degree of "belongingness" of a pattern to a symmetric pattern family will be measured by the amount of its invariance when exposed to differential operators characterizing the family. These operators perform CTs which are all equivalent to translations when represented in invariant coordinates.

We start by performing translations of f in ξ and η coordinates, begining with ξ. The translated coordinates yield:

$$T_1(\xi, \eta, \epsilon_1) = \begin{cases} \xi^* = \xi + \epsilon_1 \\ \eta^* = \eta \end{cases} \tag{8}$$

Two successive translations are equivalent to a single translation which can be obtained by using the parameter combination rule:

$$\phi(\epsilon_1, \delta) = \epsilon_1 + \delta \tag{9}$$

ϕ is analytic with respect to both of its arguments and fulfills the group axioms with $\epsilon = 0$ being the identity element of the group. These properties make T_1 a one parameter Lie group of transformations. With each Lie group of transformation an infinitesimal generator is associated. In this case this is $\mathcal{L}_1 = \frac{\partial}{\partial \xi}$. When applied to f the CT results in a translation of the iso-curves of f along the basis vector, $\hat{\xi}$, which is the curvilinear basis related to ξ. $\mathcal{L}_1$ applied to any function whose iso-curves consists of $\xi(x,y) = \lambda$, delivers the tangent fields of ξ. $\mathcal{L}_1$ is a basis of an abstract vector space of dimension one at a given position and can be utilized to represent a tangent vector or vice-versa interchangeably. Furthermore it can be shown that as long as

the functions representing the components of a vector field, $v^i(x^1, x^2)$, with $i = 1, 2$, in some coordinate system $\{x^i\}$, with $i = 1, 2$, have continuous partial derivatives with respect to x^i, this vector field is a tangent vector field. Arbitrary amount of translations in $\hat{\xi}$ direction can be obtained by applying the exponential form of $\mathcal{L}_1$

$$f(\xi^*, \eta^*) = (1 + \epsilon_1 \mathcal{L}_1 + \frac{\epsilon_1^2}{2!}\mathcal{L}_1^2 + \cdots) f(\xi, \eta) = \exp(\epsilon_1 \mathcal{L}_1) f(\xi, \eta) \tag{10}$$

which is a Taylor expansion of $f(\xi + \epsilon_1, \eta)$ around (ξ, η). Thus, the CT (ξ^*, η^*) is completely determined by $\mathcal{L}_1$'s actions on the ξ, η coordinates. As a special case, if the iso-curves of f are given by $\xi = constant$, i.e. $f(\xi, \eta) = g(\xi)$ for some g, then we have $\mathcal{L}_1 f = g'(\xi)$, leaving the *iso-curve families* of f invariant. $\xi(x, y) = \lambda$ are integral curves of $\mathcal{L}_1$, that is the iso-curve $\xi(x, y) = \lambda$ maps to another iso-curve $\xi(x, y) + \epsilon_1 = \lambda$ which is within the same family. A more powerful invariance is obtained, if $f(\xi, \eta) = g(\eta)$ for some g. This yields $\mathcal{L}_1 f = 0$ that is the iso-curves $\eta(x, y) = \lambda$, are invariant. That is one iso-curve maps into itself.

Similarly, $T_2(\xi, \eta, \epsilon_2)$ which translates the η coordinate, is a one parameter Lie group of transformation:

$$T_2(\xi, \eta, \epsilon_2) = \begin{cases} \xi^* = \xi \\ \eta^* = \eta + \epsilon_2 \end{cases} \tag{11}$$

with the corresponding infinitesimal generator

$$\mathcal{L}_2 = \frac{\partial}{\partial \eta} \tag{12}$$

As before, one can reconstruct T_2, by means of the operator $\exp(\epsilon_2 \mathcal{L}_2)$.

We can formally define a new infinitesimal operator

$$\mathcal{L} = \cos\theta \mathcal{L}_1 + \sin\theta \mathcal{L}_2 \tag{13}$$

which is linear in $\mathcal{L}_1$, $\mathcal{L}_2$ and expect to have a one parameter group of transformations, $T(\xi, \eta, \epsilon)$ corresponding to it. This expectation can be realized precisely because L_1 and L_2 are tangent vector fields of a coordinate basis, by construction. (For general vector fields represented by $\frac{\partial}{\partial \lambda}$ and $\frac{\partial}{\partial \mu}$, in order that a coordinate basis generating them via tangents exist, the vector fields must commute and be independent) Consequently, the commutator $[\mathcal{L}_1, \mathcal{L}_2]$ vanishes, so that a finite motion along the integral curves of $\mathcal{L}$ can be obtained by successive exponentiations (Taylor series):

$$\exp(\epsilon \mathcal{L}) = \exp(\epsilon \cos\theta \mathcal{L}_1) \exp(\epsilon \sin\theta \mathcal{L}_2) = \exp(\epsilon \sin\theta \mathcal{L}_2) \exp(\epsilon \cos\theta \mathcal{L}_1) \tag{14}$$

This reveals that $\mathcal{L}_1$ and $\mathcal{L}_2$ together act as an operator basis pair for *any* translation, [12], which in turn make $\mathcal{L}$ a classical *directional gradient* but in the curvilinear coordinates (ξ, η). The corresponding one parameter Lie group of transformation is given by

$$T(\xi, \eta, \epsilon_1) = \begin{cases} \xi^* = \xi + \epsilon \cos\theta \\ \eta^* = \eta + \epsilon \sin\theta \end{cases} \tag{15}$$

with θ being a "directional" constant, characterizing a unique family of the invariant curves of the operator $\mathcal{L}$, (13).

$$-\sin\theta \xi + \cos\theta \eta = constant \tag{16}$$

The converse is also true, i.e. (16) uniquely represents $\mathcal{L}$. In our approach ξ and η are known a priori, but not θ. We try to fit a family of curves as defined by (16) to an image by minimizing the error energy

$$E(\theta) = \int |\mathcal{L}f(\xi,\eta)|^2 d\xi d\eta = \int |(\cos\theta\mathcal{L}_1 + \sin\theta\mathcal{L}_2)f(\xi,\eta)|^2 d\xi d\eta \tag{17}$$

with respect to θ, since $\mathcal{L}f = 0$, almost everywhere, whenever $E(\theta) = 0$. Throughout this paper we assume that the integrals are of Lebesgue type. By using Equation (10), Equation (17), can be seen as the total error in a small translation.

For illustration, we go back to Example 2 and note that the operators $\frac{\partial}{\partial\xi}$ and $\frac{\partial}{\partial\eta}$ have the curves $\eta = \tan^{-1}(x,y) = constant$ and $\xi = \log(\sqrt{x^2+y^2}) = constant$ as invariants. Since $\mathcal{L}_1 = \frac{\partial}{\partial\xi}$ acts as a zoomer and $\mathcal{L}_2 = \frac{\partial}{\partial\eta}$ as a rotator, the direction given by θ_{min} represents the amount of scaling versus rotation leaving f unchanged (in practice least changed). The $E(\theta_{min})$ contains information as to whether θ_{min} which represents symmetry is or is not significant.

The basis of recognition is that $E(\theta)$, is small for some θ. We define

$$k(\theta) \triangleq (\cos\theta, \sin\theta)' \tag{18}$$

where "$\prime$" represents transposition and we rewrite the error function in the well-known quadratic form

$$E(\theta) = \int k(\theta)' \begin{pmatrix} \frac{\partial f}{\partial\xi} \\ \frac{\partial f}{\partial\eta} \end{pmatrix} \left(\frac{\partial f}{\partial\xi}, \frac{\partial f}{\partial\eta}\right) k(\theta) d\xi d\eta = k(\theta)' A k(\theta) \tag{19}$$

where we have defined A as

$$A \triangleq \begin{pmatrix} \int \frac{\partial f}{\partial\xi}\frac{\partial f}{\partial\xi} d\xi d\eta & \int \frac{\partial f}{\partial\xi}\frac{\partial f}{\partial\eta} d\xi d\eta \\ \int \frac{\partial f}{\partial\xi}\frac{\partial f}{\partial\eta} d\xi d\eta & \int \frac{\partial f}{\partial\eta}\frac{\partial f}{\partial\eta} d\xi d\eta \end{pmatrix} \tag{20}$$

Since A is a symmetric semi positive definite matrix (because $E(\theta) \geq 0$), the minimum of $E(\theta)$ is the smallest eigenvalue of A, $\lambda_{min} \geq 0$. This is attained when the corresponding eigenvector is utilized as k. When we know A we can, not only minimize $E(\theta)$ but also maximize it. That is for a certain θ_{max}, $E(\theta_{max}) = \lambda_{max}$. Furthermore, the maximization will prove to be useful when determining whether the error obtained is large or small, since we do not have a-priori knowledge about which ranges of the error can be considered as large or small. Because of its information redundancy, since it is always orthogonal to k_{min}, k_{max} is of no practical interest. Remembering that $k_{min} = (\cos\theta_{min}, \sin\theta_{min})'$ can be determined by one parameter, it is clear that the information content in A is equivalent to that of the tuple

$$(\lambda_{min}, \lambda_{max}, \theta) = (E(\theta_{min}), E(\theta_{max}), \theta)'. \tag{21}$$

When A is known, finding the tuple (21) is very simple, since the solution can be expressed in closed form following elementary algebra schemes. However, representation using complex numbers leads to efficient data structures in implementation and also easier visual interpretation.

Applying Parseval-Plancherel identity in elements of matrix A (20) gives

$$A = \begin{pmatrix} \mu_{20}(|F|^2) & \mu_{11}(|F|^2) \\ \mu_{11}(|F|^2) & \mu_{02}(|F|^2) \end{pmatrix} \tag{22}$$

Hence, elements of A correspond to second order moments of the power spectrum $|F|^2$, which is obtained by Fourier transforming the image f in (ξ, η) coordinates. The power spectrum always has its mass center at the origin since f is real valued. For Cartesian coordinates it has been shown that fitting a line to the power spectrum in the LSE sense is equivalent to minimizing (17), e.g. [6]. The line fitting interpretation is also valid for harmonic coordinates. The second order moments can also be equivalently, represented by the second order complex moments, I_{pq} with $p+q=2$:

$$I_{pq} = \int (\omega + i\zeta)^p (\omega - i\zeta)^q \, |F(\omega, \zeta)|^2 \, d\omega d\zeta \tag{23}$$

We note that $I_{pq} = I_{qp}^*$, and hence given I_{pq}, I_{qp}^* is redundant. Therefore, we have only the second order complex moments I_{20} (complex), and I_{11} (real) which are informative. By using the next theorem, our minimization can be carried out utilizing the complex moments formalism.

Theorem 1 *For the function $|F|^2$, the minimum error line (through the origin) as well as the minimum error are determined by the second order complex moments of the power spectrum, in ξ and η coordinates.*

$$|I_{20}| = E(\theta_{max}) - E(\theta_{min}), \qquad \arg I_{20} = 2\theta_{min} \tag{24}$$

$$I_{11} = E(\theta_{max}) + E(\theta_{min}) \tag{25}$$

$$\tag{26}$$

The maximum error line is orthogonal to the minimum error line if $I_{20} \neq 0$. In the degenerate case when $I_{20} = 0$ no single orientation performs better than any other, i.e. $E(\theta) = constant$.

We omit the proof since it is similar to the proofs of the same theorem in Cartesian or harmonic polar coordinates, [4, 5, 8]. As a direct consequence of the theorem we have the following result, which will be used in goodness of fit assessments.

Corrolary 1 *The inequality*

$$|I_{20}| \leq I_{11} \tag{27}$$

holds with equality if and only if f is linearly symmetric with respect to ξ, and η, i.e. when $E(\theta_{min}) = 0$.

By applying the definition of the complex moments, (23), and using the conservation of the scalar products, the second order complex moments of the power spectrum can be written as

$$I_{20} = \int (\omega + i\zeta)^2 \, |F(\omega, \zeta)|^2 \, d\omega d\zeta = \mu_{20} - \mu_{02} + i2\mu_{11} = \int (\mathcal{L}_1 f(\xi, \eta) + i\mathcal{L}_2 f(\xi, \eta))^2 d\xi d\eta \tag{28}$$

and

$$I_{11} = \int (\omega + i\zeta)(\omega - i\zeta) \, |F(\omega, \zeta)|^2 \, d\omega d\zeta = \mu_{20} + \mu_{02} = \int |\mathcal{L}_1 f(\xi, \eta) + i\mathcal{L}_2 f(\xi, \eta)|^2 d\xi d\eta \tag{29}$$

We define the *infinitesimal linear symmetry operator* , ILS, as

$$S(f)(\xi, \eta) = (\mathcal{L}_1 f(\xi, \eta) + i\mathcal{L}_2 f(\xi, \eta))^2, \tag{30}$$

which is non-linear in f and complex valued. $S(f)$ is all that is needed in order to explicitly perform the minimization of $E(\theta)$ since the quantities (I_{20}, I_{11}) are the mean values of $S(f)$, and $|S(f)|$, respectively.

$$I_{20} = \int S(f)(\xi,\eta)d\xi d\eta \tag{31}$$

$$I_{11} = \int |S(f)(\xi,\eta)|d\xi d\eta \tag{32}$$

The parameters represented by I_{20} and I_{11} will be called the GLS parameters.

Representing the eigenvalues and eigenvectors of A by means of the scalar pair (I_{20}, I_{11}) is appropriate since it provides all information about the line fitting process directly. From a mathematical view point, the tuple is of-course equivalent to the elements of A, or to the tuple given by (21). However, since the notion image may correspond to a local image and therefore the minimization process would need to be carried out many times, the data structure and the computation of eigenvalue and eigenvectors of even a simple matrix must be designed carefully. Next we will comment on the magnitude and the argument of I_{20}.

First the ILS as operator is applied to the image, then the averages of $S(f)(\xi,\eta)$, and $|S(f)(\xi,\eta)|$ are taken (in analogy with tensors in physics) so that the GLS parameters are obtained. The magnitudes of the components have direct interpretations which are related to the errors. For example, $|I_{20}| \approx 0$, predicts a very poor fit when $|I_{20}| << I_{11}$, or conversely $0 << |I_{20}|$, represents a very good line fit when $|I_{20}| \approx I_{11}$. Provided that f is non-zero, the corresponding $|I_{20}|$ can be made arbitrarily large or small by multiplying f with a scalar. Since $|I_{20}| \leq I_{11}$, in which the inequality holds with equality exclusively when $E(\theta_{min}) = 0$, measuring I_{11} establishes the level of $|I_{20}|$ to expect when there is a good fit of line independent of the contrast in the image.

The argument of I_{20} is 2θ which maps θ and $\theta + \pi$ to the same angle making the numerical representation of $\mathcal{L}$ unambiguous and continuous. This is important because GLS parameters can then be smoothed further, for example for sub sampling purposes. The smoothing of k_θ vector is not appropriate as the average of the two orientations k_θ and $-k_\theta + \pi$ which represent identical patterns, and hence can very likely occur in the same neighborhood, is 0, instead of k_θ or $-k_\theta$.

4 Estimation of symmetry parameters

Equations (31-32) formally describe how to obtain the GLS parameters. However in that form these equations are not practical to implement since the representation of the linear symmetry operator as well as the averaging, are in ξ, η coordinates. The canonical representation of the transformation group, with the properties

$$\begin{aligned} \mathcal{L}_1\xi &= 1 \quad \mathcal{L}_1\eta = 0 \\ \mathcal{L}_2\xi &= 0 \quad \mathcal{L}_2\eta = 1, \end{aligned} \tag{33}$$

was motivated by the powerful interpretation (uncoupled translations) it offered during the analytic treatment.

In different coordinate systems the representation of the Lie operators and the linear symmetry operator may look quite different despite the identical physical effect when applied to a function. The representation of the Lie operators in the Cartesian coordinates can be obtained by using the chain rule:

$$\mathcal{L}_1 = x_\xi \frac{\partial}{\partial x} + y_\xi \frac{\partial}{\partial y} = \frac{\xi_x}{\xi_x^2+\xi_y^2}\frac{\partial}{\partial x} + \frac{\xi_y}{\xi_x^2+\xi_y^2}\frac{\partial}{\partial y} = \frac{\nabla'\xi}{\|\nabla\xi\|^2}\nabla \tag{34}$$

$$\mathcal{L}_2 = x_\eta \frac{\partial}{\partial x} + y_\eta \frac{\partial}{\partial y} = -\frac{\xi_y}{\xi_x^2+\xi_y^2}\frac{\partial}{\partial x} + \frac{\xi_x}{\xi_x^2+\xi_y^2}\frac{\partial}{\partial y} = \frac{\nabla'_\perp\xi}{\|\nabla\xi\|^2}\nabla \tag{35}$$

where the definitions $\xi_x = \frac{\partial\xi(x,y)}{\partial x}$ and $\nabla'_\perp\xi = (-\xi_y, \xi_x)$ are used for simplicity. Moreover, the partial derivatives of ξ, and η are obtained by inverting the Jacobian of T, and using the Cauchy-Riemann equations, (2):

$$\frac{\partial(x,y)}{\partial\xi\partial\eta} = \begin{pmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{pmatrix} = \left(\frac{\partial(\xi,\eta)}{\partial x\partial y}\right)^{-1} = \begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix}^{-1} = \frac{1}{\xi_x^2+\xi_y^2}\begin{pmatrix} \xi_x & -\xi_y \\ \xi_y & \xi_x \end{pmatrix} \tag{36}$$

We note that $\mathcal{L}_2$ does not explicitly depend on partial derivatives of η with respect to x or y which of-course is the consequence of the HFP assumption which binds the gradients of ξ and η together.

For illustration we go back to Example 2, and see that the corresponding infinitesimal operators are easily found by using (34-35) :

$$\mathcal{L}_1 = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$$

$$\mathcal{L}_2 = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial x}$$

These are, from the differential geometry, well known scaling and rotation operators.

When the Lie operators are represented in Cartesian coordinates, we can write the ILS operator as

$$S(f)(\xi,\eta) = (\mathcal{L}_1 f + i\mathcal{L}_2 f)^2 = [(\nabla\xi + i\nabla_\perp\xi)'\nabla f]^2/\|\nabla\xi\|^4. \tag{37}$$

However, since

$$\nabla\xi + i\nabla_\perp\xi = \begin{pmatrix} \xi_x - i\xi_y \\ i(\xi_x - i\xi_y) \end{pmatrix} = (\xi_x - i\xi_y)\begin{pmatrix} 1 \\ i \end{pmatrix} \tag{38}$$

where $\xi_x - i\xi_y$ is complex valued (scalar function), we have the result

$$\begin{aligned} S(f)(\xi,\eta) &= \frac{(\xi_x - i\xi_y)^2}{|\xi_x - i\xi_y|^4}[(1,i)\begin{pmatrix} f_x \\ f_y \end{pmatrix}]^2 = \frac{(\xi_x - i\xi_y)^2}{|\xi_x - i\xi_y|^4}(f_x + if_y)^2 \\ &= \frac{S^*(\xi)(x,y)}{|S^*(\xi)(x,y)|^2}S(f)(x,y). \end{aligned} \tag{39}$$

which we restate in the following theorem.

Theorem 2 *Under a harmonic conjugate basis change given by ξ,η, the ILS operator changes basis according to:*

$$S(f)(\xi,\eta) = \frac{S^*(\xi)(x,y)}{|S^*(\xi)(x,y)|^2}S(f)(x,y) \tag{40}$$

Consequently I_{20} is given by

$$\begin{aligned} I_{20} &= \int S(f)(\xi,\eta)d\xi d\eta = \int \frac{S^*(\xi)(x,y)}{|S^*(\xi)(x,y)|^2} S(f)(x,y)\det(\frac{\partial\xi(x,y)}{\partial x\partial y})dxdy \\ &= \int \frac{S^*(\xi)(x,y)}{|S^*(\xi)(x,y)|} S(f)(x,y)dxdy = \int w^{20^*}(x,y)S(f)(x,y)dxdy \end{aligned} \tag{41}$$

where the Cauchy-Riemann equations applied to the functional determinant

$$\det(\frac{\partial\xi(x,y)}{\partial x\partial y}) = \begin{vmatrix} \xi_x & \xi_y \\ -\xi_y & \xi_x \end{vmatrix} = \xi_x^2 + \xi_y^2 = |S^*(\xi)(x,y)| \tag{42}$$

has been used. In equation (41), we note that I_{20} is obtained by projecting the ILS of the image in Cartesian coordinates, which does not depend on the pattern family ξ, on the kernel w^{20} which depends on ξ. Consequently, once the ILS of the image is available it can be tested for containing a multitude of pattern families by changing the corresponding kernel w^{20} without recalculating the ILS of the image. This is not surprising as the ILS of the image represents "universal" information for pattern matching, namely the local edges and the double of their direction. The double angle representation of local edges which preserves the numerical continuity of the angle information under the smoothing operation, has been proposed by Granlund, [21], and Knutsson [32].

Similarly, I_{11} which is the average ILS magnitude can be obtained yielding the following interesting property.

Theorem 3 (Energy Conservation) *The sum of the maximum and the minimum error is independent of the coordinate system chosen for symmetry investigation of the image f:*

$$\begin{aligned} I_{11} &= E(\theta_{max}) + E(\theta_{min}) = \int |S(f)(\xi,\eta)|d\xi d\eta \\ &= \int |S(f)(x,y)|dxdy = \int (\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 dxdy. \end{aligned} \tag{43}$$

The theorem suggests that

$$0 \le |I_{20}(|F(\omega,\zeta)|^2)| \le E(\theta_{max}) + E(\theta_{min}) = I_{11}(|F'(\omega',\zeta')|^2) \tag{44}$$

where the $I_{..}(F)$ explicitly displays the function which is used to obtain the respective complex moments. $F'(\omega',\zeta')$ represents the Fourier transform of $f'(\xi',\eta')$ which is the original image f represented in another HFP coordinates (ξ',η'). Since the upper bound is independent of the chosen CT (ξ,η), I_{11} can be computed in Cartesian coordinates. Equality occurs only when the neighborhood is truly symmetric with respect to the chosen coordinates. The closer $|I_{20}|$ is to this common upper bound, the better the chosen transformation describes the underlying symmetry. Equation (42) is the projection of the magnitude of the ILS of the image to the magnitude of w^{20} which is the constant function 1. From now on the term ILS of f, where f can be the original image or the prototype of a target pattern, will be assumed to refer to the ILS of f in Cartesian coordinates, that is $S(f)(x,y) = (f_x + if_y)^2$.

5 Discrete approximation

5.1 Estimation of the ILS image

For computation of GLS parameters, the ILS of the image f,

$$S(f)(x,y) = (f_x + if_y)^2, \tag{45}$$

must be computed on line and discretely. For simplicity, we will assume that f is band-limited although analogues of the argumentation below will also apply to other classes of functions, e.g. piecewise polynomial functions. A band-limited function $\alpha(x,y)$ can be represented by means of its samples α_j through interpolation functions ψ_j

$$\alpha(x,y) = \sum_j \alpha_j \psi_j(x,y) \tag{46}$$

provided that the discretization frequency, DF, is higher than twice the Nyquist frequency, NF. The square of a band-limited signal is also band limited, since squaring corresponds to auto convolution in the frequency domain. Then for $4NF < DF$ in each dimension we may write

$$\alpha^2(x,y) = \sum_j \alpha_j^2 \psi_j(x,y) \tag{47}$$

over-sampling at least twice is not a severe restriction as images are either i) already over-sampled that much, or ii) they can be easily over-sampled by using the discrete data. Therefore, without loss of generality we assume that (47) is valid.

For band-limited functions, the interpolation function ψ_j can, theoretically, be obtained by inverse Fourier transforming the frequency support function of the assumed band-limited signal family. But it is also known that interpolation functions obtained by this formal procedure have very large spatial supports with frequent sign changes, e.g. the Sinc function, and cause numerical errors since numerous additions and subtractions must be carried out with limited word lengths of computers. The family of Gaussians

$$\psi(x,y) = G_\sigma(x,y) = C \exp\,(- \frac{(x)^2 + (y)^2}{2\sigma^2}) \tag{48}$$

where C is a constant normalizing the integral of ψ to 1, is convenient for approximating all images of practical interest and is often used as interpolation function in image analysis. This is motivated by its simplicity, positivity, and its joint concentration in the Fourier and spatial domain, and as well as for its other known advantages, [33]. We adopt it as well despite the fact that its support is not finite in the frequency domain. The constant σ, is a parameter which is to be determined experimentally since in practice the image is already discretized and one can only speculate over as to how much it is over-sampled or to which function family it belongs to.

Thus, provided that the gradient image is densely sampled the ILS image can be written as

$$S(f)(x_j,y_j) = (f_x + if_y)^2|_{(x,y)=(x_j,y_j)} = [(f_{x_j} + if_{y_j})]^2 \tag{49}$$

where f_{x_j} is the value of $\frac{\partial f}{\partial x}$ at the grid point $(x,y) = (x_j,y_j)$. The sampled gradient image represented as a complex scalar, is obtained by applying partial differentiation to (46):

$$f_{x_j} + if_{y_j} = \sum_l f_l(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})\psi(x - x_l, y - y_l)|_{(x,y)=(x_j,y_j)} = [f_k * (\psi_{x_l} + i\psi_{y_l})](x_j,y_j) \tag{50}$$

where * represents discrete convolution. Technically, ILS image is nothing but a non-thresholded and squared edge image when the edge image is interpreted as a complex image rather than a vector image. The estimation of GLS parameters consists of projecting the ILS image and its magnitude to the discrete kernels w_j^{20} w_j^{11}, respectively, as will be seen next.

5.2 Estimation of I_{20}: case 1

Here we assume that $\nabla\xi$, and therefore $S(\xi)(x,y)$ is known for continuous x, y. In section 5.3 we will study the case when this continuous information is not available. An approximation of I_{20} can be obtained by substituting the reconstructed (from its samples) $S(f)$,

$$S(f)(x,y) = \sum_j \psi(x - x_j, y - y_j) S(f)(x_j, y_j) \tag{51}$$

into equation (41)

$$I_{20} = \sum_j S(f)(x_j, y_j) \int \psi(x - x_j, y - y_j) [\frac{S(\xi)(x,y)}{|S(\xi)(x,y)|}]^* dxdy = \sum_j S(f)(x_j, y_j)(w_j^{20})^* \tag{52}$$

where w_j^{20} is given by

$$w_j^{20} = \int \psi(x - x_j, y - y_j) \exp(i2 \tan^{-1}(\xi_x, \xi_y)) dxdy \tag{53}$$

Here, we observe that the discrete kernel of I_{20}, that is w_j^{20}, is obtained by projecting the continuous kernel defined through equation (41) as,

$$w^{20}(x,y) = \frac{S(\xi)(x,y)}{|S(\xi)(x,y)|} = \exp(i2 \tan^{-1}(\xi_x, \xi_y)) \tag{54}$$

onto the space of band-limited signals. We note that the continuous kernel has modulus 1 except at $\nabla\xi(x,y) = 0$, where $S(\xi)$ is undefined. At these points, w^{20} can safely be assumed to be 0, as the values of the integrand on a set of points with zero measure does not affect I_{20}. Technically (53) is a low-pass filtering followed by discretization, which is also known as perfect sampling. Thus equation (52) is essentially a matching of the orientation of the basis tangent vector field with the tangent vector field of the image. This observation will prove to be useful, in Section (5.3). Naturally, the closed form of the integral in (53) is not possible to obtain for most ξ's. However, w_j^{20} can be computed numerically and off line, e.g. [8], for pattern recognition purposes. In one important case, when $\xi = x$, though, (53) can be derived analytically and reduces to $w_j^{20} = 1$

Is it really worth computing w_j^{20} exactly through (53) in order to obtain a useful approximation of I_{20}? The answer to this question depends on the application at hand. The computation of w_j^{20} through equation (53) and then substituting it in (52) yields robust approximations of I_{20} since the weight zero is automatically given to the appropriate points of the kernel at the same time as all kernel coefficients vary smoothly. If the digitized image f_j, represents a small neighborhood it might be worth to compute w_j^{20} in the afore-mentioned "orthodox" fashion (i.e. by projection on the band-limited functions), as this yields less biased estimates. However, this may not be appropriate, if the number of singularity points is negligible compared to the total

number of the image points, since the bias of the singularity points will be negligible. In this case one can use the approximation,

$$w_j = \begin{cases} \exp(i2\tan^{-1}(\xi_x, \xi_y))|_{x=x_j y=y_j}, & \text{if } \nabla\xi(x_j, y_j) \neq 0; \\ 0, & \text{if } \nabla\xi(x_j, y_j) = 0. \end{cases} \tag{55}$$

5.3 Estimation of I_{20}: Case 2

Often a digital image of a prototype pattern, $\tilde{\xi}_j$ is all that is known and one would like to know whether $\tilde{\xi}_j$ occurs in a discrete image f_j or not. One can not assume that iso-curves of the prototype are sampled iso-curves of harmonic functions, i.e. the iso-curves fulfill the Laplacian equation, as this is i) difficult to verify for an arbitrary digital prototype ii) it is not very helpful when $\tilde{\xi}$ is not strictly harmonic. The tilde in $\tilde{\xi}_j$ is used in order to note that we do not know whether $\tilde{\xi}$ is harmonic or not. We can exploit the fact that the computation of I_{20} is essentially a matching between on the ILS image and the sampled version of the normalized ILS of the prototype. As an algorithm implementing the discrete ILS operator according to equation (49) is assumed to exist, we can apply such an algorithm to $\tilde{\xi}_j$ to obtain,

$$S(\tilde{\xi})(x_j, y_j) = [(\tilde{\xi}_{x_j} + i\tilde{\xi}_{y_j})]^2, \tag{56}$$

We can then proceed as if $\tilde{\xi}_j$ is a sampled harmonic function and find the kernel w_j^{20} as

$$w_j^{20} = \begin{cases} \exp(i2\tan^{-1}(\tilde{\xi}_{x_j}, \tilde{\xi}_{y_j})), & \text{if } \nabla\xi(x_j, y_j) \neq 0; \\ 0, & \text{if } \nabla\xi(x_j, y_j) = 0. \end{cases} \tag{57}$$

Then, in analogy with (52), I_{20} can be computed as

$$I_{20} = \sum_j S(f)(x_j, y_j)(w_j^{20})^* \tag{58}$$

The classical alternative to the case dealt within this subsection is to match (correlate) directly the two digital images, f_j, and $\tilde{\xi}_j$, without filtering them through the ILS operator. However, matching the ILS image with an appropriate kernel has certain advantages. In the ILS approach it is the edges of f_j and not the gray values which are aligned in case of match. As a consequence we can expect a high localization. However, this is only a by product, the main advantage is the complex voting process and its rich interpretability as will be discussed in Section 6.

5.4 Estimation of I_{11}

According to equation (43), I_{11} is obtained as

$$I_{11} = \int |S(f(x, y))| \cdot 1 dxdy = \int |S(f)(x, y)||w^{20}(x, y)| dxdy \tag{59}$$

Consequently, the continuous kernel of I_{11} is

$$w^{11} = |w^{20}| \tag{60}$$

By using the reconstructed $|S(f)|$, and in analogy with (52),

$$\bar{I}_{11} = \sum_j |S(f)(x_j, y_j)| \bar{w}_j^{11} \tag{61}$$

where the discrete kernel

$$\bar{w}_j^{11} = \int \psi(x - x_j, y - y_j) |w^{20}(x, y)| dxdy \tag{62}$$

is the projection of the continuous kernel $w^{11}(x, y)$ onto the space of band-limited signals. Using the kernel $\bar{w}_j^{11}$ in equation (61) yields $\bar{I}_{11}$– the biased estimate of I_{11}– as is shown below. The discrete kernel $\bar{w}_j^{11}$ is the perfect sampling of $w^{11}(x, y)$. The magnitudes of the two kernels fulfill the inequality $|w_j^{20}| \leq \bar{w}_j^{11}$ which is a weaker relationship than the equality relationship in the continuous case, (60). Consequently, when $|w_j^{20}| < \bar{w}_j^{11}$ even for one kernel coefficient j, we get

$$|I_{20}| = |\sum_j S(f)(x_j, y_j) w_j^{20^*}| \leq \sum_j |S(f)(x_j, y_j)||w_j^{20}| < \sum_j |S(f)(x_j, y_j)| \bar{w}_j^{11} = \bar{I}_{11} \tag{63}$$

This implies that, with this inequality $|I_{20}|$ will not attain the value $\bar{I}_{11}$ even if the image is linearly symmetric. By using the triangle inequality, it can be shown that $|I_{20}|$ will attain the upper bound only when the discrete ILS of the image and the kernel coefficients are co-linear. This behavior is similar to the continuous case, as can be seen by applying the triangle inequality to (41) and comparing the result to (43). In order to avoid the bias introduced by the discretization process, we will use the discrete kernel

$$w_j^{11} = |w_j^{20}| \tag{64}$$

instead of Equation (62), in order to compute

$$I_{11} = \sum_j |S(f)(x_j, y_j)||w_j^{20}| \tag{65}$$

where w^{20} is assumed to be available through the processes described in Section 5.2 and Section 5.3.

6 Symmetry detection is GHT with complex votes

The GLS parameter estimation is more than a correlation of edge magnitudes. It carefully takes into account the directions of the edges, too. Here we show that the GLS detection extends GHT, to a complex GHT in which the votes are allowed to assume complex values.

GHT has been popular for its robustness even when the prototype $\tilde{\xi}_j$ is partly missing in the image f_j. It consists of a table look-up and a voting procedure; see also [1]. The table entries are discrete edge directions of the edges of the prototype. To each entry corresponds a list of edge positions, expressed relative to a unique reference point (origin) of the prototype. Given the table one can easily reconstruct the *unitary gradient of the prototype* , or inversely construct the table, given the unitary gradient of the prototype. By unitary gradient we mean the gradient vectors with unity (edge) or 0 (non-edge) magnitudes. This is because edges

participating in the GHT poll are binary, i.e. they either exist or do not. Consequently the table is essentially another representation of the unitary gradient of the prototype. It can be shown that the GHT procedure is identical to i) sliding the unitary gradient of the prototype over the unitary gradient of the image and counting the number of unitary gradient agreements (edge and direction matches) per sliding position ii) the process of assigning this count to the current sliding position, which qualifies as a reference point if enough votes are cast for it. The determination of the threshold for qualification as a referenc point is a step that can assume varying degrees of complexity, but it is not of interest here. The votes cast for a point to be a reference point is then,

$$A(x_j, y_j) = \sum_l \delta(\varphi^r(x_j^r + x_l^p, y_j^r + y_l^p) - \varphi^p(x_l^p, y_l^p)) \tag{66}$$

where $\varphi^p(x_l^p, y_l^p)$ is the direction of the prototype edge pixel at the position (x_j^p, y_j^p) in the unitary gradient of the prototype. $\varphi^r(x_j^r + x_l^p, y_j^r + y_l^p)$ corresponds to the direction of the edge at the position $(x_j^r + x_l^p, y_j^r + y_l^p)$ in the unitary gradient of the image, whenever it is an edge pixel, or $i\infty$. That φ^r is defined to be $i\infty$ is purely symbolic and serves to make φ^r different from φ^p, resulting in the vanishing of the δ function. $A(x_j, y_j)$ represents the vote accumulator of a pixel position and can be compared to,

$$\begin{aligned} I_{20}(x_j, y_j)) &= \sum_l S(f)(x_j^r + x_l^p, y_j^r + y_l^p) w^{20^*}(x_l^p, y_l^p) \\ &= \sum_l \exp[i2\varphi^r(x_j^r + x_l^p, y_j^r + y_l^p) - i2\varphi^p(x_l^p, y_l^p)] \end{aligned} \tag{67}$$

The latter assumes that the same unitary gradient of the image, and the unitary gradient of the prototype as those used in the computation of A are employed, in order to obtain $S(f)(x_j, y_j)$ and $w^{20}(x_j, y_j)$, via (57), and (58). Consistent with equation (66), $\varphi^r = i\infty$ whenever it is not defined, i.e. when a pixel position does not represent an edge pixel position, the contribution of that point to I_{20} is reduced to 0. While $A(x_j, y_j)$ is the count of the positive matches only, $|I_{20}|$ is the positive matching score *adjusted downwards* with the amount of negative matches. At the reference point, when the image is the same as the prototype itself, we obtain the maximum match with both of the techniques

$$A = L \qquad I_{20} = L \tag{68}$$

where L represents the number of edge pixels in the prototype. When 100% of the edge directions mismatch maximally i.e. when all prototype directions are orthogonal to the image directions we obtain,

$$A = 0 \qquad I_{20} = -L \tag{69}$$

while if 50% of the edge directions match perfectly and 50% mismatch maximally we obtain,

$$A = \frac{L}{2} \qquad I_{20} = 0 \tag{70}$$

Now, it is clear that in computing $I_{20}(x_j, y_j)$ the scores of the positions with contradictory unitary gradient matches will see their scores reduced compared to the GHT scores, $A(x_j, y_j)$. The GHT is without score reduction since a score of $A(x_j, y_j)$ is only allowed to increase (in

case of unitary gradient match) or it is unchanged (in case of mismatch). For GHT this is a necessity, because negative scores would not be meaningful in case they were given to empty accumulators, while in the case of I_{20} this is allowed as this simply corresponds to a vote for a pattern which is locally orthogonal to the prototype, *anti-prototype*. Clearly, the computation of I_{20} is a voting process in which not only negative but also complex votes are allowed.

However, this has to be carefully considered: Formally, non-prototype ILS-images are possible to generate by a phase shift of the prototype ILS-image:

$$\exp(i\varphi_0)\exp(i2\tan^{-1}(\tilde{\xi}_{x_j},\tilde{\xi}_{y_j})) \tag{71}$$

But to which real patterns $\varphi_0 \neq 0$ corresponds, as this new ILS-image is a purely synthetic construct, is not obvious. This is because a phase shift of the prototype ILS may result in tangent fields which are not always imaginable or intelligible by visual inspection of (71). To give every φ_0 an exact meaning, i.e. to find a ξ_j which approximates $\tilde{\xi}_j$, an estimation should be made by numerical methods. For pattern recognition purposes, however, this will not be necessary given the control possibility I_{11} offers. If, for an image for which $0 << |I_{20}| \approx I_{11}$, $\varphi_0 = \arg I_{20} \neq 0$ is obtained, the non-prototype is known in reality too (as the gradients come from a real image). Thus in practice only when $|I_{20}| << I_{11}$ may pose interpretation difficulties of $\arg I_{20}$ in which case no member of this class is a good fit to the data anyway.

7 Experimental results

********** Figure 3 about here **********

Experiment 1: Linear symmetry as texture measure

Unsupervised texture segmentation consists of the sub tasks: feature extraction, feature selection and segmentation.We will concentrate on the theory of the feature extraction part and will show that some already proposed, [4, 5], and independently tested [6, 7, 39] texture discrimination measures are covered by the pattern recognition theory developed here. Thus we do not propose a new set of texture measures, but rather show that the detection of repetitive patterns as well as the detection of non repetitive patterns, see further below, share the same theory.

In texture analysis one wishes to design features (measures) which can discriminate textures having the structural properties

1. "No" particular dominant direction (isotropic), versus "One" dominant direction (directional)

2. "Fine" (rich in high spatial frequencies) versus "Coarse" (rich in low spatial frequencies)

A band pass decomposition (non-directional) of the image can provide the necessary measures to test for Property 2, e.g. Laplacian pyramid, [13]. Independent of their orientations fine textures will appear in high frequency bands while they will be suppressed in low frequency bands. Coarse textures will have an analogous response in the decomposition. However, this decomposition can not offer information to discriminate isotropy from non-isotropy e.g. whether a texture which is fine and has a dominant direction versus a texture which is also fine but has

no direction. Here we propose to measure this property by use of the transformation $\xi = x$ which models the linear patterns with a dominant orientation. Since we dispose two goodness of fit measures we can discriminate textures with one dominant direction from those without. That is, a directional texture will be characterized by $0 << |I_{20}| \approx I_{11}$ while an isotropic texture will be characterized by $0 \approx |I_{20}|$ and $0 << I_{11}$. We observe that the symmetry model $\xi = x$ implies that a line is fit to the local power spectrum in the Cartesian coordinates. It is known that the Gabor responses are discrete approximations of the local power spectrum (windowed Fourier transform) e.g. see [2, 14]. The second order moments of the Gabor power spectrum which are equivalent to I_{20} and I_{11} when $\xi = x$ as in this experiment, are known texture measures with high discrimination power, [32, 38, 9].

As already mentioned, $w_j = 1$ for the symmetries characterized by $\xi = x$. Since GLS parameters are to be computed for every neighborhood, we approximate the ILS image of the windowed image with the windowed ILS image. Consequently, I_{20}, for every neighborhood is an image which can be computed by smoothing $S(f)(x_j, y_j)$ with the windowed kernel, (the hat underlines that it is a windowed kernel)

$$\hat{w}_j^{20} = G_\sigma(x_j, y_j) w_j = \exp(-\frac{x_j^2 + y_j^2}{2\sigma_w^2}) \cdot 1 \tag{72}$$

The original image Figure 3 (a), has been used for evaluating the performance of texture analysis methods in our laboratory. It is composed of real aerial image patches (7 in total) repeated in such a way that every combination of texture boundary crossover is present. The patches represent an urban area, natural forest, cultivated forest, and agricultural fields all imaged at the same height above the ground.

The different ILS images are obtained by using a $\sigma = 0.25$ in (48-50). By choosing 3 levels in the Laplacian pyramid we obtained 3 sub-bands which differed roughly, in octaves. Since we have 3 real images per band, we obtained 9 feature images which differed in size. In order to make use of the segmentation method, [41], we had access to, we equalized the image sizes by proper interpolation.

We asked for 7, 6 and 8 classes from our unsupervised segmentation algorithm and obtained the results in Figure 3 (b)-(d) when 9 linear symmetry features are used. The segmentation was unsupervised, that is training was not employed in class center determination. Class centers were determined by automatic clustering in the feature space so that only labels e.g. "A", "B"... appear as class names in Figure 3 (b)-(d). Nor was an automatic feature reduction mechanism enhancing the segmentation result but obscuring a direct illustration of GLS parameters, was utilized. For a complete segmentation strategy that includes these improvements, refer to [41]. The number of classes needed to be specified as the classifier uses the fuzzy c-means clustering, and a pyramidal boundary estimation technique in order to obtain the labels and the boundaries. In the case of Figure 3 (d) there are 7 classes despite asked 8 since the segmentation algorithm is designed to merge close classes. However, the boundaries are better estimated when the true number of classes are requested, (c). The boundaries of (b) are closer to the true boundaries than those of (c) with the drawback of one class is missing. In conclusion, these results indicate that the linear symmetry features have the ability to discriminate between natural textures even without refinements. For extensive comparisons with other features and the extension of the linear symmetry features using the same test images we refer to [6, 7, 9].

Experiment 2: chirality determination

In many registration or object position identification problems, one would like to identify the so called singularity points e.g. neighborhoods which may look like a "circle", or a "star". This is because these neighborhoods offer invariance to many geometric transformations such as zooming and rotation. Also, with the increased use of digital imaging in applications the automatic determination of chirality of the visual objects, i.e. left handedness versus right handedness are gaining importance. For example, in chemistry and material sciences, it has been known that the chirality of molecule chains often dramatically affects the organic properties (e.g. natural sugar and diet sugar), and image processing is considered as an analytic tool see [24].

In Figure 4, we are interested in whether I_{20} can be used to discriminate between classes *left handed, right handed, circular patterns, and star shaped patterns.* To the right half of the image, Gaussian white noise is added in order to test the orientation measurement sensitivity at low SNR's. Every pattern in the left half also exists in the right half but with noise added. In Figure 5 we are interested in detecting the sea anemones present in the under water photograph.

Letting

$$\xi = \log(r) \tag{73}$$

is appropriate for both tasks since these basically represent the same problem in our formulation. We assume that the notion image corresponds to a local image, that is we compute GLS parameters at every pixel by using a Gaussian window function as in the previous experiment. Since we know the analytic expression of $\nabla\xi$ we obtain the kernel of I_{20}, by employing (53):

$$\hat{w}_j^{20} = G_{\sigma_w}(x_j, y_j) w_j^{20} = G_{\sigma_w}(x_j, y_j) \cdot \int G_\sigma(x - x_j, y - y_j) \exp(i2 \tan^{-1}(x, y)) dxdy \tag{74}$$

The kernel w_j^{11} is given by $|\hat{w}_j^{20}|$. The exact values of $\hat{w}_j^{20}$ for Gaussian window and interpolation functions can be found in [8].

Figure 4 displays the original superimposed with lines representing the fit $\mathcal{L}$, i.e. $\arg(I_{20})$ is used for orientation. The ILS image was obtained by using $\sigma = 0.85$ and $\sigma_w = 2.8$. The lengths of the lines are proportional to $|I_{20}|$. We observe that $|I_{20}| \approx 0$ for patterns with low frequencies which fall outside of the frequency band supported by the derivation filters. Such patterns can be detected by using larger σ's. The role of σ is primarily to select the frequency band at which the detection is effective. Clearly, a classification based on the arguments, and the magnitudes of I_{20} as well as on I_{11} can identify the objects which are left or right handed ...etc. Furthermore, the $\arg I_{20}$ on the right half of the Figure 4 (Right), indicates a low sensitivity of the orientation parameter estimation to Gaussian noise. The line lengths, which are modulated by $|I_{20}|$, in the right half of the image are slightly shorter suggesting a relatively good quality estimation compared to the noiseless pattern. This is not surprising since the least-square models are optimal for Gaussian noise. Optimal estimates assuming other noise, (such as impulsive noise) are possible to obtain by applying order statistics, [37], to the ILS of the image instead of averaging. But a discussion of this type of extension is too much of a digression from the purpose of this paper.

********** Figure 4 about here**********

In Figure 5, overlayed to the original (Right), three black patches display the result of supervised box classification (component wise thresholding) on the $|I_{20}|$, $\arg I_{20}$, and I_{11} as they were obtained for the under water photograph. We used a different window size, $\sigma_w = 11.2$, since the anemones cover larger portions of the image than the spirals in the previous test

image. Despite that, the anemones differ in size and appearance 3 (4 present in the picture) were detected. The missing anemone deviates too much from the model. By including further features, e.g. the gray value itself, it is possible to detect all anemones. Here we wanted to show the properties of the GLS features.

********** Figure 5 about here **********

Experiment 3: orientation determination of crosses

To illustrate that the located $\mathcal{L}$'s (i.e. orientation in the $\xi - \eta$ space) are meaningful and accurate, we attempt to detect the curve family represented by

$$z^2 = \xi(x,y) + i\eta(x,y) = x^2 - y^2 + i2xy \tag{75}$$

The patches in Figure 6 consist of the members of this family. The procedure of detecting and identifying the members is analogous to that of the spiral family presented in Experiment 2. In fact, the only difference is in the used windowed kernel, $\hat{w}_j^{20}$, which is given by the complex conjugate of (74). Overlayed the original, the crosses in Figure 6 represent the found members at the center of the patches. They show a good degree of orientation fit to the asymptotes. The crosses are generated synthetically with orientations as found by the method which delivers 4θ with θ being the angle representing any asymptotic direction of the four possible. As already pointed out the $\arg I_{20}$ automatically solves the numerical representation ambiguity. The size of the crosses are modulated by $|I_{20}|$.

**********Figure 6 about here **********

Experiment 4: Arbitrary patterns

Figure 7 a,b display an electronic circuit picture taken by a CCD camera with 256 gray levels at the output, along with the prototype transistor which needs to be identified. The ILS of the prototype and the original were quantized to two levels (thresholded) with respect to their magnitudes which essentially capture the edge strength. The arguments of the ILS images are twice the gradient direction angles, which were also quantized and stored as an image with 256 gray levels.

Figure 7 c,d display $|I_{20}|$ and I_{11}. Both values of these measures are equal and high i.e. $0 << |I_{20}| \approx I_{11}$, only at very localized points corresponding to the true positions of the transistors i.e. the center of the prototype image (compare to $|I_{20}|$). No other points than the true transistor positions have such a high agreement of the $|I_{20}|$ and I_{11} indicating that only the prototype, $\varphi_0 = 0$, and no other patterns of the same family, $\varphi_0 \neq 0$ exist in the original. The results were practically unchanged when we varied the quantization levels of the magnitudes or the arguments or when not applying the quantization at all.

Figure 7 e, illustrates the accumulator image when the GHT transform with 256 directional entries were applied. The peaks of the transistor positions have several orders of magnitudes difference and therefore only the largest peak is clearly identifiable. This behavior is expected as the δ function in (66) rejects to cast a vote even when the directions differ only by a small amount, whereas the $\exp(i\cdot)$ function in (67) processes these differences in a continuous manner. However, because of this insensitivity to quantization the localization is somewhat worse in the GLS than in the GHT approach for the found transistors (one for GHT).

Since GLS parameters can be obtained by means of convolutions, we have chosen FFT to compute them while GHT was implemented in a straight forward manner resulting in longer computation times.

8 Discussion and conclusion

In the theory presented and in the experiments, we have dealt with GLS parameters that measure the fit of the linear symmetry to the original image. Knowing that $|I_{20}| \leq I_{11}$ holds with equality exclusively when the examined image is a member of the pattern family, i.e. symmetry exists, allowed us to interpret the scalars $|I_{20}|$, and I_{11} as quality measures. However, one can define

$$D = I_{11} - |I_{20}| \tag{76}$$

which measures the lack of symmetry. In particular when $\xi = x$, D represents "non-lines", i.e. corners, high curvature points, T-junctions, ...etc. This is illustrated by Figure 8, where the result represents lack of linear symmetry with respect to Cartesian coordinates ($\sigma = 0.5$ and $\sigma_w = 1.0$). Thus, other uses of these two scalars are possible since their significance is well understood.

We have derived a method to model symmetries of the neighborhoods in gray value images for pattern recognition. It is based on the iso-curves which are invariants of commutative Lie groups of transformations. It uses the gray image directly without intermediate thresholding. It is shown that the GLS parameter extraction can be performed in the spatial domain by means of projections and squaring, yielding two scalars, I_{20} (complex) and I_{11} (non-negative). In case the "image" represents a neighborhood, these scalars are obtained through filtering of a complex valued image (ILS image). The magnitudes of the resulting pair represent the degree of symmetry with respect to the a-priori chosen HFP. The degree of symmetry has a clear definition which is based on the maximum and minimum error and can be used in order to quantify the "symmetry orientation" estimation quality. The argument of I_{20} represents the member of the Lie group operator family, generated by the HFP, that leaves the image invariant.

We have also shown that GLS filtering is equivalent to a voting process similar to that of GHT with the extension that complex votes are allowed. The complex votes, permit the detection and identification of "prototypes" as well as "anti-prototypes" concurrently.

In the paper we have neither addressed how to optimally select the filter parameters for a pattern/curve recognition task nor have we addressed the automatic selection of models (CTs). These problems need further investigation. A pragmatic way to circumvent the problem is to try many parameters, as in Experiment 1 where the same filter parameters were applied to many resolutions, or many CTs. GLS parameter estimations are implemented via matched filters applied to the complex valued ILS image. Techiniques allowing selective application of matched filters to few locations in images rather than exhaustive application to the entire image, need to be developed. Currently, if the image represents a local neighborhood, estimation of the GLS parameters are equivalent to the steps i) Linear filtering (one complex FIR filter which is the sum of two separable FIR filters) ii) pixelwise complex squaring iii) Linear Filtering (one complex FIR filter). For general CT's, step (iii), which is the most computationally expensive, can be implemented using Fast Fourier Transform. For the CT $\xi = x$ the step (iii)

can be implemented efficiently through separable filters as this step is equivalent to ordinary smoothing.

Acknowledgments

I would like to thank to the Computer Vision Laboratory staff at Linköping University where I worked out the essential parts of the theory during my Ph.D. studies, 1983-88, especially to Prof. G. H. Granlund, my supervisor then, and Dr. H. Knutsson for enthusiastic discussions. The Swedish National Board for Technical Development (now NUTEK) is gratefully acknowledged for its financial support.

References

[1] B. H. Ballard *Generalizing the Hough transform to detect arbitrary shapes* Pattern Recognition 13,2, pp. 111-112, (1981).

[2] M. J. Bastian *A sampling theorem for the complex spectrogram, and Gabor's expansion of a signal in Gaussian elementary signals* Optical engineering, vol. 20, No. 4, pp. 594-598 (1981)

[3] O. Bertrand, R. Queval, and H. Maitre *Shape interpolation using Fourier descriptors with application to animation graphics* Signal Processing, vol. 4, pp. 53-58, (1982)

[4] J. Bigun, *Circular symmetry models in image processing*, Linköping studies in science and technology; Tekn. Lic. thesis No:85, (1986).

[5] J. Bigun, G.H. Granlund *Optimal orientation detection of linear symmetry*, First international conf. on computer vision, London, pp. 433-438, June (1987).

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

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

[8] J. Bigun *A structure feature for image processing applications based on spiral functions* Computer vision, graphics and image processing. No. 51, pp. 166-194 (1990)

[9] J. Bigun and J. M. H. du Buf *N-folded symmetries by complex moments in Gabor space.* IEEE-PAMI vol. 16, No.1, pp. 80-87 (1994)

[10] J. Bigun *Pattern recognition by detection of local symmetries.* Pattern recognition and artificial intelligence, North-Holland, Amsterdam, pp. 75-90, (1988), Editors: E.S. Gelsema and L.N. Kanal.

[11] J. Bigun *Local symmetry features in image processing.* PhD thesis No. 179, ISBN 91-7870-334-4, (1988), Dept. of Electrical Eng., Linköping University, S-581 83, Sweden.

[12] G. W. Bluman and S. Kumei *Symmetries and differential equations* Springer, Berlin (1989)

[13] P. Burt *Fast filter transforms for image processing* Computer graphics and image processing 16, pp. 20-51 (1981)

[14] I. Daubechies *The wavelet transform, time-frequency localization and signal analysis* IEEE Trans. on Inf. Theory, vol. 36, no.5, pp. 961-1005, (1990)

[15] B. Duc *Motion estimation using invariance under group transformations* In Proc. ICPR-94, Jerusalem, (1994)

[16] 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.

[17] R. Eagleson, and T. Caelli *Group theoretic analysis of local flow characteristics while visually tracking a textured surface* in proc. 5'th int. conf. Image Analysis and Processing, V. Cantoni et. al. Ed., Positano, Italy, 20-22 Sept., World scientific, pp. 443-450 (1989)

[18] M. Ferraro and T. M. Caelli *Relationship between integral transform invariances and Lie group theory* J. Opt. Soc. Am. A; Vol. 5, No.5 pp. 738–742 (1988)

[19] L.M.J. Florack, B.M. ter H. Romeny, J.J. Koenderink and M.A. Viergever *General intensity transformations and second order invariants* 7'th Scandinavian conference on image analysis, Aalborg, 13-16 august, pp. 338–345 (1991)

[20] G. H. Granlund *Fourier preprocessing for hand print character recognition* IEEE trans. Computers, Vol. 21, pp. 195-201 (1972)

[21] G.H. Granlund *In Search of a General Picture Processing Operator*, Computer Graphics and Image Processing 8, 155–173 (1978).

[22] H. Greenspan, M. Porat, and Y. Y. Zeevi *Projection based approach to image analysis: Pattern recognition and representation in the position-orientation space* PAMI vol. 14, No. 11, pp. 1105-1110 (1992)

[23] O. Hansen and J. Bigun *Local symmetry modeling in multi-dimensional images* Pattern Recognition Letters, 13, pp. 253-262 (1992)

[24] Y. Hel-Or, S. Peleg and D. avnir *Characterization of right-handed and left-handed shapes, CVGIP-Image understanding, Vol. 53, pp. 297-302 (1991)*

[25] W.C. Hoffman *The Lie algebra of visual perception* J.math. Psychol. (1966) 3 p65-98

[26] W.C. Hoffman *Higher visual perception as prolangation of the basic Lie transformation group* Mathematical Biosciences Vol. 6 (1970), pp. 437–471

[27] P. V. C. Hough *Method and means for recognizing complex patterns*, U.S. Patent 3,069,654 (1962)

[28] M. K. Hu, *Visual pattern recognition by moment invariants* IRE Tr. on information Theory, pp. 179- 187 (1962)

[29] K. Kanatani *Coordinate rotation of image characteristics for 3D shape and motion recovery* ICCV, London, june 8-11, pp. 55-64 (1987)

[30] M. Kass, and A. Witkin *Analyzing oriented patterns* CVGIP, vol. 37, pp. 362-385 (1987)

[31] H. Knutsson, M. Hedlund, G. H. Granlund *Apparatus for Determining the Degree of Variation of a Feature in a Region of an Image that is Divided into Discrete Picture Elements* US. patent, 4.747.151, Filed 1986, (1988)

[32] H. Knutsson *Filtering and reconstruction in image processing*, Linköping studies in science and technology Dissertations No:88, (1982).

[33] T. Lindeberg *Scale space theory in computer vision* Monograph Kluwer, (1994)

[34] T. Lindeberg *On scale selection for differential operators* 8'th SCIA May 25-28, Tromsoe, pp. 857-866 (1993)

[35] P. Minovic, S. Ishikawa, and K. Kato *Symmetry identification of a 3-D object represented by octree.* IEEE-PAMI vol. 15, Nr. 5, pp. 507-513, (1993)

[36] T. V. Papathomas and B. Julesz *Lie differential operators in animal and machine vision* From pixels to features, J.C. Simon (ed) Elsevier, pp. 115-126 (1989)

[37] Y. Ping, V. Anastassopoulos and A. N. Venetsanopoulos *Pattern recognition based on morphological shape analysis and neural networks* Special issue of the Journal fo Mathematics and computers in simulation (IMACS Transaction) on Neural Networks/Neural Computing pp.28, in press (1994).

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

[39] A. R. Rao and R. C. Jain *Computerized flow field analysis: oriented texture fields* PAMI vol. 14, No. 7, pp. 693-709 (1992)

[40] P. T. Sander, and S. W. Zucker *Singularities of principal direction fields from 3-D images* PAMI 309-317 (1992)

[41] P. Schroeter and J. Bigun *Hierarchical image segmentation by multi-dimensional clustering and orientation adaptive boundary refinement* Pattern Recognition, vol. 28, No. 5, pp.695-709, (1995)

[42] E. L. Schwartz *Computational Anatomy and Functional Architecture of Striate Cortex: A spatial Mapping Approach to Perceptual Coding* Visual Research, vol. 20, pp. 645–669, (1980)

[43] H. Zabrodsky, S. Peleg, and D. Avnir *A measure of symmetry based on shape similarity* CVPR-92, June 15-18, Champaign, IL, pp. 703-706 (1992)

[44] C. T. Zahn and R. Z. Roskies *Fourier descriptors for plane closed curves* IEEE-T on computers, Vol. C-21, No. 3, pp. 269-281, (1972)

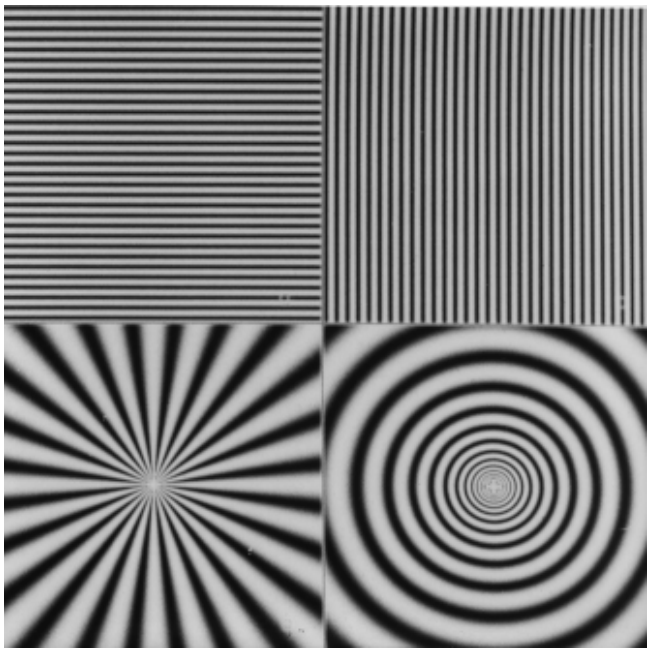

Figure 1: The figure illustrates the HFPs used in Example 1 (top) and in Example 2 (bottom) respectively.

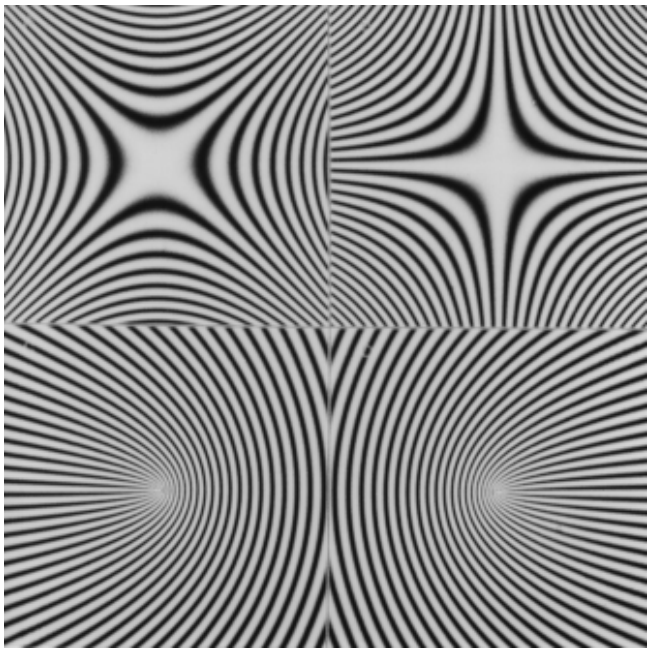

Figure 2: The figure illustrates the HFPs used in Example 3 (top) and in Example 4 (bottom) respectively.

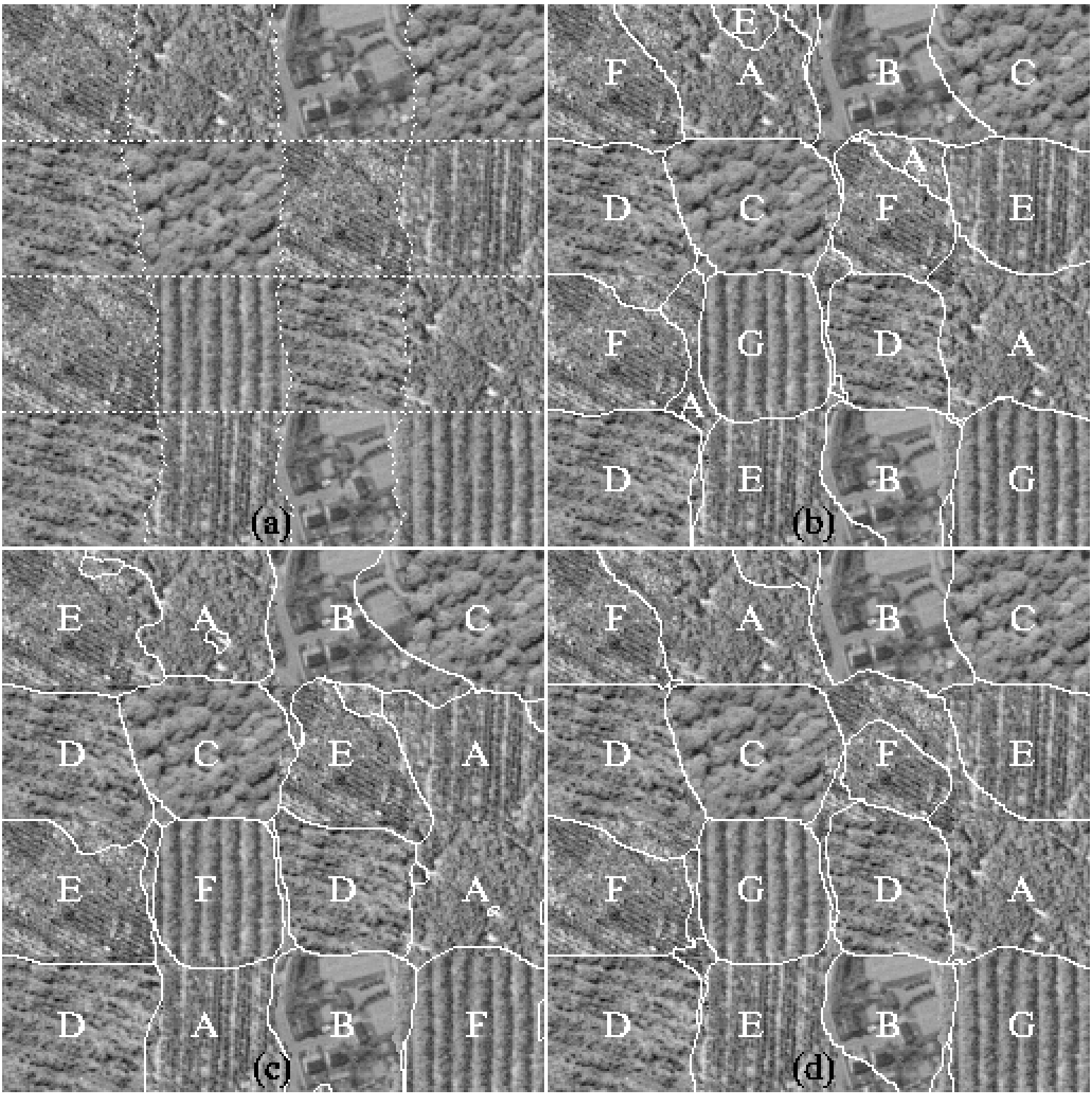

Figure 3: The texture image to be segmented in an unsupervised manner (a), and the segmentation results (b), (c), and (d), when 7,6, and 8 classes are requested from the unsupervised segmentation.

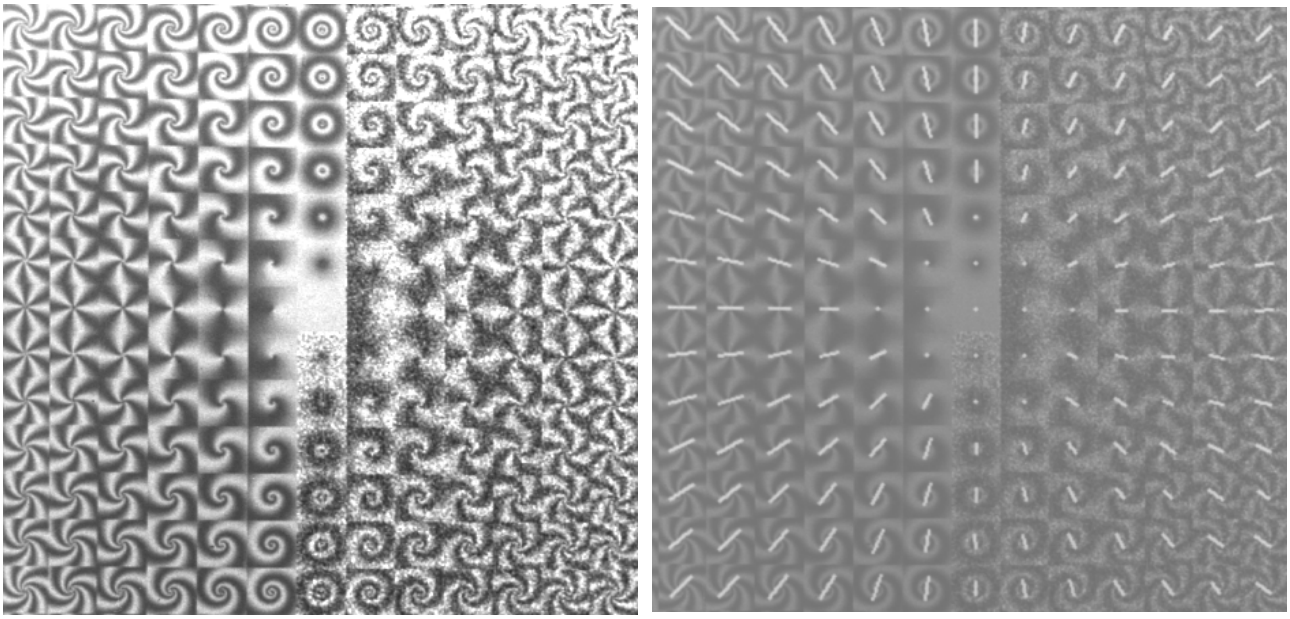

Figure 4: (Left) The figure illustrates some neighborhoods included in the harmonic polar isocurve model. The orientation of the model corresponds to the "twistedness" of a neighborhood. (Right) The "symmetry" of the spiral family are shown as lines. The orientations of the lines permit classification between, circular, radial, left handed, and right handed objects.

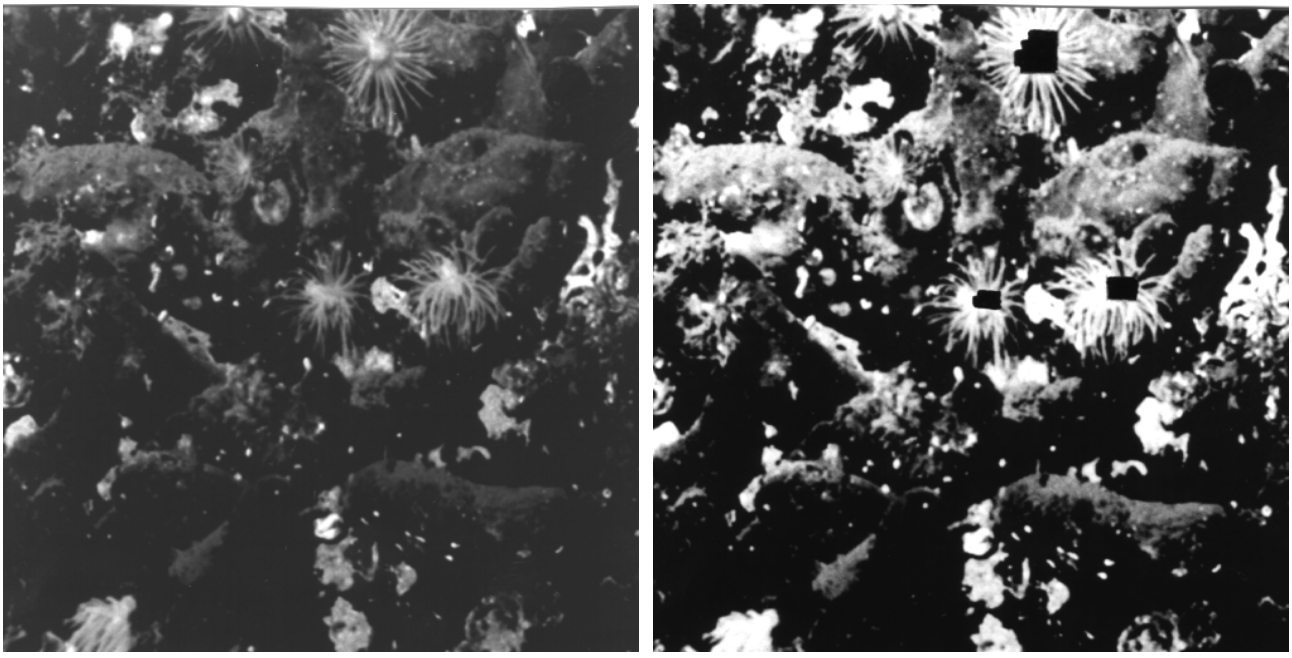

Figure 5: (Left) The image is a sea bottom photograph. The objective is to identify the sea anemones. (Right) The labels are obtained as a result of box classification.

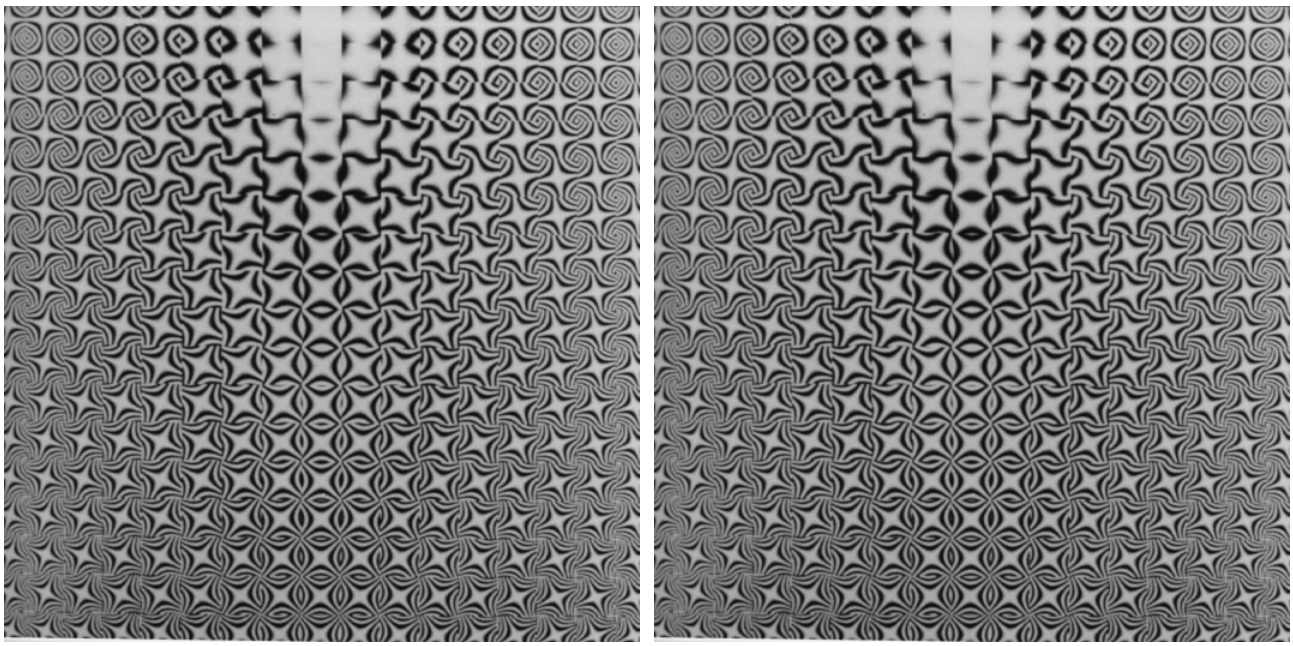

Figure 6: (Left) The figure illustrates neighborhoods included in the hyperbolic curve model. The orientation of the symmetry parameter represents the orientation of the two orthogonal asymptotes. (Right) The crosses which correspond to the asymptotes are overlayed the original. The orientations and the sizes of the crosses are the result of modulation with $\arg I_{20}$ and $|I_{20}|$ respectively.

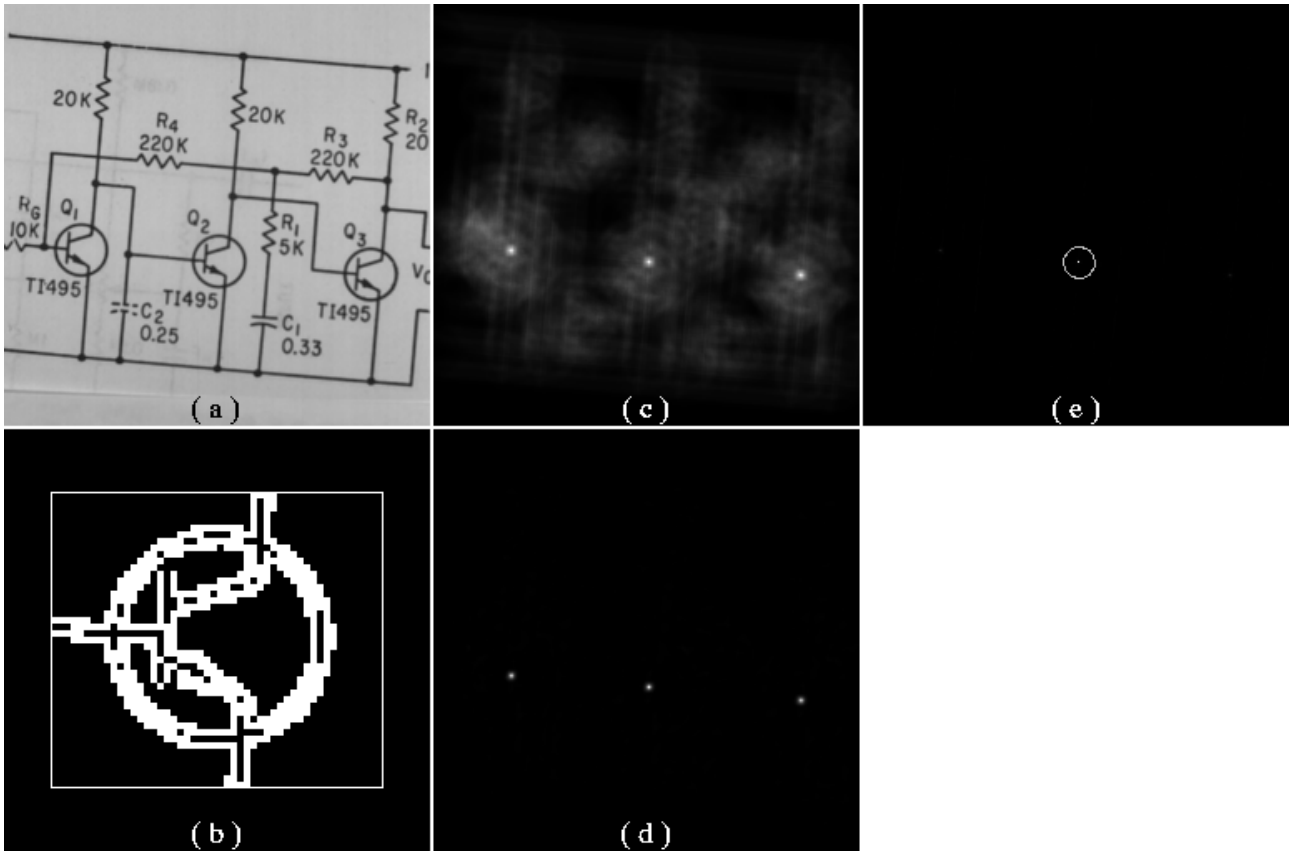

Figure 7: The electronic circuit and transistor detection. a,b: the original and the prototype. c,d: $|I_{20}|$ and I_{11}. e: GHT accumulator with $2\pi/256$ angle resolution. The circle is inserted to show the location of the peak.

Figure 8: The original and D, measuring lack of symmetry ($\xi = x$) which can be used to identify corner, junction, high curvature, ...etc points.