

Representing Local Structure Using Tensors.

Hans Knutsson

Linköping University
Computer Vision Laboratory
S-581 83 Linköping Sweden

ABSTRACT

The fundamental problem of finding a suitable representation of the orientation of 3D surfaces is considered. A representation is regarded suitable if it meets three basic requirements: *Uniqueness, Uniformity and Polar separability*. A suitable *tensor* representation is given.

At the heart of the problem lies the fact that orientation can only be defined mod 180° , i.e the fact that a 180° rotation of a line or a plane amounts to no change at all. For this reason representing a plane using its normal vector leads to ambiguity and such a representation is consequently not suitable. The ambiguity can be eliminated by establishing a mapping between R_3 and a higher-dimensional tensor space.

The uniqueness requirement implies a mapping that map all pairs of 3D vectors x and $-x$ onto the same tensor T . Uniformity implies that the mapping implicitly carries a definition of distance between 3D planes (and lines) that is rotation invariant and monotone with the angle between the planes. Polar separability means that the *norm* of the representing tensor T is rotation invariant. One way to describe the mapping is that it maps a 3D sphere into 6D in such a way that the surface is uni-

formly stretched and all pairs of antipodal points maps onto the same tensor.

It is demonstrated that the above mapping can be realized by sampling the 3D space using a specified class of symmetrically distributed *quadrature filters*. It is shown that 6 quadrature filters are necessary to realize the desired mapping, the orientations of the filters given by lines through the vertices of an *icosahedron*. The desired tensor representation can be obtained by simply performing a weighted summation of the quadrature filter outputs. This situation is indeed satisfying as it implies a simple implementation of the theory and that requirements on computational capacity can be kept within reasonable limits.

Noisy neighborhoods and/or linear combinations of tensors produced by the mapping will in general result in a tensor that has no direct counterpart in R_3 . In an adaptive hierarchical signal processing system, where information is flowing both up (increasing the level of abstraction) and down (for adaptivity and guidance), it is necessary that a meaningful inverse exists for each level-altering operation. It is shown that the point in R_3 that corresponds to the best approximation of a given tensor is given by the largest eigenvalue times the corresponding eigenvector of the tensor.

INTRODUCTION

A meaningful application of most signal processing concepts requires that the data to be processed represent some aspect of the real world in an orderly way. More precisely, it is generally required that an increased difference between real world events results in an increased distance between the data points that represent these events. Working with representations where these requirements are not met would make most signal processing concepts meaningless, e.g. averaging and differentiation. Thus, it must be concluded that a suitable representation of events to be analysed is the basis for successful signal processing. For many real-world aspects, however, establishing well behaved representations is a non-trivial task and in these cases a first and necessary step of any analysis is to find such a representation.

In this paper the problem of finding a suitable representation of 3-D surface orientation is considered. As an introductory example the equivalent 2-D case will be discussed. The problem then of course reduces to finding a suitable representation for 2-D lines.

In the following a *line* will be taken to mean a symmetric separator of two regions such that no relation between the regions can be inferred, whereas by an *edge* will be understood an asymmetric separator such that an ordering of the two regions is always implied. Note that a consequence of this definition is that lines can only be unambiguously associated with an angle over an interval of 180° , while edges can be unambiguously associated with angles spanning an interval of 360° . Another way to view the situation is that both lines and edges are associated with angles spanning a 180° interval and that in addition, edges are '*signed*' but lines are '*unsigned*'.

Considering the above, it is argued that a general partitioning element must have the properties of a line. The reason for this is that

it is not possible to consistently define an ordering of regions if the data constituting the different regions are intrinsically more than 1-dimensional and consequently the concept of an edge is not applicable. To see why the edge concept does not work for multi-dimensional data, consider the case of differently colored adjacent patches (to a human observer color is intrinsically 2-dimensional). It does not seem to make much sense to say that one color has a higher value than any other (equal brightness and saturation assumed) and furthermore it is easy to convince oneself that all attempts to impose an ordering of the color space will result in unreasonable distances between at least some of the colors.

Considering the fact that a substantial part of the concepts used in image processing are of multi-dimensional nature (color, texture, shape etc.) it becomes apparent that, in any image processing theory claiming generality, a partitioning element must have the above mentioned line properties.

Another and perhaps more fundamental view of the situation is the following. The representation should be invariant to *what* is *in* the regions and only represent the geometric qualities of the separator itself i.e. in this case, its orientation.

In trying to find a representation for line orientation vectors defined over a 2-dimensional half plane immediately suggests itself, the vector angle being identical to the line orientation and the vector length being a positive valued function of some other property of the line (e.g. its energy, probability or some other definition of its 'magnitude'). A space resulting from this type of direct mapping will be referred to as the original space. It is obvious in the present case that, for a given vector length, the original space corresponds to a half circle. The problem with this original space is that it leads to an unacceptable distance measure as vectors representing two lines differing by a small angle can end up being represented by

vectors that are very far apart, i.e. one line is represented close to one end of the half circle and the other line is represented close to the opposite end.

MAPPING REQUIREMENTS

To remove this discontinuity a mapping is required that maps the original space into a suitable representation space. To guarantee a suitable mapping three basic requirements should be met.

The ‘uniqueness’ requirement:

It is evident that the discontinuity will be removed if the mapping is such that it maps the vectors \mathbf{x} and $-\mathbf{x}$ of the original space onto the same vector in the representation space i.e.

$$\mathbf{T}(\mathbf{x}) = \mathbf{T}(-\mathbf{x}) \quad (1)$$

where:

\mathbf{x} is a vector in the original space.

\mathbf{T} is the map of \mathbf{x} .

The ‘uniform stretch’ requirement:

The mapping should locally preserve the angle metric of the original space, i.e

$$\|\delta\mathbf{T}\| = c \|\delta\mathbf{x}\|_{r=const.} \quad (2)$$

where:

$$r = \|\mathbf{x}\|$$

c is a ‘stretch’ constant.

The ‘polar separability’ requirement:

As the information carried by the magnitude of the original vector \mathbf{x} does not normally depend on the vector angle, it is reasonable to require that:

$$\|\mathbf{T}\| = f(\|\mathbf{x}\|) \quad (3)$$

i.e the norm of \mathbf{T} is independent of the direction of \mathbf{x} .

THE MAPPING

A mapping M that maps the vector \mathbf{x} onto the tensor \mathbf{T} and meets all the above criteria is given by:

$$M : \quad \mathbf{T} \equiv r^{-1} \mathbf{x}\mathbf{x}^T \quad (4)$$

The *norm* of \mathbf{T} is taken to be the Frobenius norm and is given by:

$$\|\mathbf{T}\|^2 \equiv \sum_{ij} t_{ij}^2 = \sum_n \lambda_n^2 \quad (5)$$

where:

t_{ij} are the components of \mathbf{T} and

λ_n are the eigenvalues of \mathbf{T} .

Since \mathbf{T} is a quadratic form a change in the sign of \mathbf{x} will have no effect and it follows directly that the ‘uniqueness’ requirement is met. That the ‘polar separability’ requirement is met is easily demonstrated by calculating the norm of \mathbf{T} .

$$\|\mathbf{T}\|^2 = \sum_{ij} r^{-2} x_i^2 x_j^2 = r^2 \quad (6)$$

where $x_{i(j)}$ are the components of \mathbf{x} .

Showing that the norm of \mathbf{T} is equal to the norm of \mathbf{x} .

That the ‘uniform stretch’ requirement is met is shown in appendix **A**.

A FILTERING REALISATION

Having found a suitable mapping the question arises: Can the representation implied by the mapping be realized using measurements on actual image data, where lines (or other structures) are represented as local grayscale correlations? It will be shown that by combining the outputs from polar separable quadrature filters, it is possible to produce a representation corresponding exactly to eqn. (4).

The exactness relies on the image data being locally 1-dimensional, i.e. on the existence of a locally well defined orientation. The case where the 1-dimensionality assumption does not hold is discussed in ‘THE INVERSE’ section.

There are four features of the above procedure relevant to the present paper.

1. *The quadrature filter concept.*

The quadrature filter concept forms a basis for obtaining phase invariant information. A quadrature filter can be defined in Fourier space as a filter being positive over half of the Fourier space and zero over the other half, or more precisely defined by:

$$\begin{cases} F_k(\mathbf{u}) > 0 & \text{if } \mathbf{u} \cdot \mathbf{n}_k > 0 \\ F_k(\mathbf{u}) = 0 & \text{otherwise} \end{cases} \quad (7)$$

where:

\mathbf{n}_k is the filter directing vector.

\mathbf{u} is the frequency.

The output q_k of the corresponding quadrature filter will be a complex number. The magnitude $\|q_k\|$ of q_k will be phase invariant (implying local shift invariance) and the argument $\arg(q_k)$ represents the local phase.

2. *The filter shape.*

In addition to the quadrature requirement, eqn.(7), it is required that the frequency response of the filters when $(\mathbf{u} \cdot \mathbf{n}_k) > 0$ can be expressed as:

$$F_k(\mathbf{u}) = g(\|\mathbf{u}\|)(\mathbf{u} \cdot \mathbf{n}_k)^2 \quad (8)$$

In other words the filter shape is polar separable, the radial part of the function is arbitrary but positive (usually some type of bandpass function) and the angular part varies as $\cos^2(\varphi)$, where φ is

the difference in angle between \mathbf{u} and the filter direction \mathbf{n}_k . (In general the filters are spatially localized, but this is a separate issue not discussed in the present paper.)

3. *The minimum number of filters.*

The discussion regarding the minimum number of filters can be found in appendix **B**. The result is that the minimum number of filters is 3 for 2D and 6 for 3D.

4. *The combination of filter outputs.*

The final result \mathbf{T}' can be obtained by linear summation of the quadrature filter output magnitudes as indicated by eqn. 9.

$$\mathbf{T}' = \sum_k \|q_k\| (\mathbf{n}_k \mathbf{n}_k^T - \alpha \mathbf{I}) \quad (9)$$

where:

q_k is the output from quadrature filter k .

\mathbf{n}_k is the orientation of quadrature filter k .

$\alpha = 1/(m - 1)$; m being the dimensionality of \mathbf{T}' .

\mathbf{I} is the identity tensor.

The calculations leading to this result can be found in appendix **C**.

THE INVERSE

If the neighborhood is not 1-dimensional, due to noise or surface curvature, it is necessary to find a best approximation \mathbf{T} to \mathbf{T}' where \mathbf{T} corresponds to a 1-dimensional neighborhood. This is simply done by finding the \mathbf{x} that minimizes:

$$\Delta = \|\mathbf{T}' - r^{-1} \mathbf{x} \mathbf{x}^T\| \quad (10)$$

It is shown in appendix **D** that \mathbf{x} is given by:

$$\mathbf{x} = \lambda_1 \mathbf{e}_1 \quad (11)$$

giving:

$$\mathbf{T} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T \quad (12)$$

where:

λ_1 is the largest eigenvalue of \mathbf{T}' and

\mathbf{e}_1 is the corresponding eigenvector.

The value of Δ indicates how well the 1-dimensionality hypothesis fits the neighbourhood, the smaller the value the better the fit.

Appendix A UNIFORM STRETCH

To show that the *uniform stretch* requirement is met by the mapping is fairly straightforward. Add a small perpendicular vector $\epsilon \mathbf{v}$ to \mathbf{x} and calculate the relative difference in the norm of \mathbf{T} . To start define \mathbf{S} to be:

$$\mathbf{S} = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{T}(\mathbf{x} + \epsilon \mathbf{v}) - \mathbf{T}(\mathbf{x})}{\epsilon} \quad (13)$$

where:

$$\|\mathbf{v}\| = 1 \text{ and}$$

$$\mathbf{v} \cdot \mathbf{x} = 0$$

then

$$\|\delta \mathbf{T}\| = \|\mathbf{S}\| \|\delta \mathbf{x}\|_{r=\text{const.}} \quad (14)$$

Carrying out the limit calculation is simple and yields:

$$\mathbf{S} = r^{-1}(\mathbf{x} \mathbf{v}^T + \mathbf{v} \mathbf{x}^T) \quad (15)$$

Let \mathbf{S} operate on \mathbf{x} and \mathbf{v} .

$$\begin{cases} \mathbf{S} \mathbf{x} = r \mathbf{v} \\ \mathbf{S} \mathbf{v} = r^{-1} \mathbf{x} \end{cases} \quad (16)$$

Combining the results yields:

$$\begin{cases} \mathbf{S}(\mathbf{x} + r \mathbf{v}) = \mathbf{x} + r \mathbf{v} \\ \mathbf{S}(\mathbf{x} - r \mathbf{v}) = r \mathbf{v} - \mathbf{x} \end{cases} \quad (17)$$

showing that $(\mathbf{x} + r \mathbf{v})$ and $(\mathbf{x} - r \mathbf{v})$ are eigenvectors of \mathbf{S} the eigenvalues being 1 and -1 respectively. Since all other eigenvectors of \mathbf{S} are orthogonal to \mathbf{x} and \mathbf{v} it follows from eqn.(15) that all other eigenvalues are zero.

Then, according to eqn.(5), the norm of \mathbf{S} is given by:

$$\|\mathbf{S}\| = \sqrt{\lambda_1^2 + \lambda_2^2} = \sqrt{2} \quad (18)$$

showing that the '*uniform stretch*' requirement is met.

Appendix B FILTER OUTPUT

The analysis in this appendix will deal only with real valued neighbourhoods of *one-dimensional variation* i.e. neighbourhoods that can be expressed as:

$$\xi(\mathbf{s}) = f(\mathbf{s} \cdot \mathbf{n}_\xi) \quad (19)$$

where:

\mathbf{s} is the spatial coordinate and

\mathbf{n}_ξ is a unit vector oriented along the axis of maximal signal variation.

For this type of signal the Fourier transform is non-zero only on the line defined by:

$$\mathbf{u} \propto \mathbf{n}_\xi \quad (20)$$

Thus the situation can be treated as 1-dimensional and, using eqn. (8) and 1-dimensional filter theory, it is not hard to show that the magnitude of the quadrature filter output (as a function of the signal orientation) is given by:

$$\|q_k\| = d(\mathbf{n}_\xi \cdot \mathbf{n}_k)^2 \quad (21)$$

where d can be considered a constant as it is independent of the filter orientation and depends only on the magnitude and radial distribution of the signal spectrum.

Appendix C FILTER COMBINATION

As in appendix **B** the analysis in this appendix will deal only with real valued neighbourhoods of *one-dimensional variation*. In addition it is assumed that the filters axes should be symmetrically distributed over the orientation space. It is felt that this is a reasonable assumption as the final result, \mathbf{T} , by definition is invariant to rotation of the filters.

It is helpful in the following discussion to bear in mind that:

1. The Fourier transform is invariant to rotation of the coordinate system.
2. The Fourier transform of a signal of *one-dimensional variation* is a line through the origin parallel to the signal orienting vector, eqns. (19) and (20).
3. The quadrature filter output is invariant to rotation around its axis (given by \mathbf{n}_k) and also diametrically symmetric so that $\|q(\mathbf{u})\| = \|q(-\mathbf{u})\|$, eqn.(21).

Consider the case of 2^{N-1} quadrature filters, having symmetry axes passing through the corners of a cube in N dimensions, giving a fully symmetric distribution of filters. Consider the contribution to the filters from frequencies on a line through the center of two opposing cube faces. Since the angle between the line and any filter axis will be the same it is clear that all the filters will give the same output. Consequently the filter set is incapable of giving information sufficient to determine which pair of cube faces the line passes through and thus, the orientation of the signal is undecidable. It can be concluded, therefore, that more than 2^{N-1} quadrature filters must be used.

It has been shown in [2] that the minimum number of filters required when $N = 2$ is 3, the filter orientations given by:

$$\begin{aligned}\mathbf{n}_1 &= (1, 0) \\ \mathbf{n}_2 &= (-0.5, \sqrt{3}/2) \\ \mathbf{n}_3 &= (-0.5, -\sqrt{3}/2)\end{aligned}\quad (22)$$

For $N = 3$ the number of filters must be greater than 4 but, since there does not exist a way of distributing 5 filters in 3-D in a fully symmetrical fashion, the next possible number is 6. (In fact the only possible numbers are those given by half the number of vertices (or faces) of a diametrically symmetric regular polyhedron, leaving only the numbers 3, 4, 6 and 10. Note that this is in contrast to the 2-D case where the only symmetry restriction is $K > 2$.) It turns out that the minimum required number of quadrature filters K is 6. To attain the final result for the 3-D case, eqn.(31), a number of steps have to be taken.

The orientations of the filters are given by vectors pointing to the vertices of a *hemicosahedron*, see fig. 1.

The 6 normal vectors are given in cartesian coordinates by:

$$\begin{aligned}\mathbf{n}_1 &= c (a, 0, b)^T \\ \mathbf{n}_2 &= c (-a, 0, b)^T \\ \mathbf{n}_3 &= c (b, a, 0)^T \\ \mathbf{n}_4 &= c (b, -a, 0)^T \\ \mathbf{n}_5 &= c (0, b, a)^T \\ \mathbf{n}_6 &= c (0, b, -a)^T\end{aligned}\quad (23)$$

where:

$$\begin{aligned}a &= 2 \\ b &= (1 + \sqrt{5}) \\ c &= (10 + 2\sqrt{5})^{-1/2}\end{aligned}$$

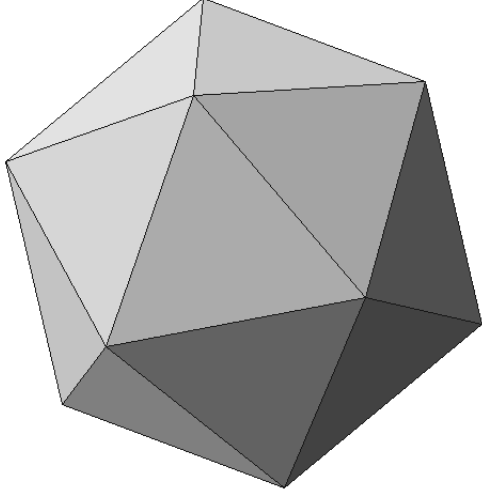


Figure 1: An icosahedron (one of the 5 Platonic polyhedra).

Then the elements of the $\mathbf{n}_k \mathbf{n}_k^T$:s are given by:

$$\begin{aligned}
\mathbf{n}_1 \mathbf{n}_1^T &= c^2 \begin{pmatrix} a^2 & 0 & ab \\ 0 & 0 & 0 \\ ab & 0 & b^2 \end{pmatrix} \\
\mathbf{n}_2 \mathbf{n}_2^T &= c^2 \begin{pmatrix} a^2 & 0 & -ab \\ 0 & 0 & 0 \\ -ab & 0 & b^2 \end{pmatrix} \\
\mathbf{n}_3 \mathbf{n}_3^T &= c^2 \begin{pmatrix} b^2 & ab & 0 \\ ab & a^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\mathbf{n}_4 \mathbf{n}_4^T &= c^2 \begin{pmatrix} b^2 & -ab & 0 \\ -ab & a^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\mathbf{n}_5 \mathbf{n}_5^T &= c^2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & b^2 & ab \\ 0 & ab & a^2 \end{pmatrix} \\
\mathbf{n}_6 \mathbf{n}_6^T &= c^2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & b^2 & -ab \\ 0 & -ab & a^2 \end{pmatrix}
\end{aligned} \tag{24}$$

Let the signal orienting vector be given by:

$$\mathbf{n}_\xi = r^{-1}(x, y, z) \tag{25}$$

then the magnitude of the outputs from the 6 quadrature filters are, according to eqn.(21), given by:

$$\begin{aligned}
\|q_1\| &= dc^2 r^{-2} (a^2 x^2 + 2abxz + b^2 z^2) \\
\|q_2\| &= dc^2 r^{-2} (a^2 x^2 - 2abxz + b^2 z^2) \\
\|q_3\| &= dc^2 r^{-2} (b^2 x^2 + 2abxy + a^2 y^2) \\
\|q_4\| &= dc^2 r^{-2} (b^2 x^2 - 2abxy + a^2 y^2) \\
\|q_5\| &= dc^2 r^{-2} (b^2 y^2 + 2abyz + a^2 z^2) \\
\|q_6\| &= dc^2 r^{-2} (b^2 y^2 - 2abyz + a^2 z^2)
\end{aligned} \tag{26}$$

Next, calculating the sum

$$\mathbf{T}'' = \sum_k \|q_k\| \mathbf{x}_k \mathbf{x}_k^T \tag{27}$$

yields the components of \mathbf{T}'' :

$$\begin{aligned}
t_{11} &= d'(x^2 r^{-2} + \frac{1}{2}) \\
t_{22} &= d'(y^2 r^{-2} + \frac{1}{2}) \\
t_{33} &= d'(z^2 r^{-2} + \frac{1}{2}) \\
t_{12} &= t_{21} = d'xyr^{-2} \\
t_{13} &= t_{31} = d'xzr^{-2} \\
t_{23} &= t_{32} = d'yzr^{-2}
\end{aligned} \tag{28}$$

where $d' = \frac{4}{5}d$.

It is evident that if the quantity $\frac{1}{2}d'$ is subtracted from the diagonal elements of \mathbf{T}'' the result will be of the desired form.

$$\mathbf{T}' = \mathbf{T}'' - \frac{1}{2}d' = d' \mathbf{n}_\xi \mathbf{n}_\xi^T \tag{29}$$

Finally calculate the sum of all quadrature filter output magnitudes.

$$\sum_k \|q_k\| = 2d \tag{30}$$

Combining eqns. (27),(29) and (30) yields the desired result:

$$\mathbf{T}'(d'\mathbf{n}_\xi) = \sum_k \|q_k\|(\mathbf{n}_k\mathbf{n}_k^T - \frac{1}{5}\mathbf{I}) \quad (31)$$

A few comments about higher dimensional spaces are appropriate. For $N = 4$ the number of filters must be greater than 8 and the only possible number is 12. No calculations for the 4-dimensional case have been carried out. For N greater than 4 no regular polyhedron having more vertices than a cube exists.

Appendix D THE INVERSE

The norm of a tensor is invariant under rotation of the coordinate system and eqn.(10) can be rewritten as:

$$\Delta = \|\mathbf{A}^{-1}(\mathbf{T}' - r^{-1}\mathbf{xx}^T)\mathbf{A}\| \quad (32)$$

giving:

$$\Delta = \|\mathbf{A}^{-1}\mathbf{T}'\mathbf{A} - \mathbf{A}^{-1}r^{-1}\mathbf{xx}^T\mathbf{A}\| \quad (33)$$

where \mathbf{A} is an orthogonal matrix.

Let \mathbf{A} be such that $\mathbf{A}^{-1}\mathbf{T}'\mathbf{A}$ is diagonal and note that only one eigenvalue of \mathbf{xx}^T is nonzero. Then, since the norm of $\mathbf{A}^{-1}\mathbf{T}'\mathbf{A}$ is the sum of the squares of its elements, it is clear that Δ is minimized if $\mathbf{A}^{-1}r^{-1}\mathbf{xx}^T\mathbf{A}$ removes the largest of these values, i.e. the largest eigenvalue of $\mathbf{A}^{-1}\mathbf{T}'\mathbf{A}$. Thus, if the eigenvalues are numbered in decreasing order, Δ is given by:

$$\Delta = \sqrt{\sum_{n \neq 1} \lambda_n^2} \quad (34)$$

Then, since \mathbf{T}' and $r^{-1}\mathbf{xx}^T$ are subject to identical rotation, it is clear that the \mathbf{x} which minimizes Δ is given by:

$$\mathbf{x} = \lambda_1\mathbf{e}_1 \quad (35)$$

where \mathbf{e}_1 is the eigenvector corresponding to the largest eigenvalue of \mathbf{T} .

Acknowledgement

The author wishes to thank the members of the Computer Vision group, especially Prof. Gösta Granlund for being a constant source of inspiration and Dr Josef Bigün for showing me the elegant trick involved in appendix A. This work was supported by the Swedish National Board for Technical Development.

References

- [1] G. H. Granlund. "In Search of a General Picture Processing Operator", *Computer Graphics and Image Processing* 8, 155-173, 1978.
- [2] H. Knutsson. *Filtering and Reconstruction in Image Processing*. Diss. No. 88, Linköping University, 1982.
- [3] H. Knutsson, G. H. Granlund. "Texture Analysis Using Two-Dimensional Quadrature Filters". *CAPAIDM Workshop*, Pasadena, California, October 1983.
- [4] H. Knutsson. "Producing a Continuous and Distance Preserving Representation of 3-D Orientation". *CAPAIDM Workshop*, Miami Beach, Florida, November 1985.
- [5] Bigün, J., 1988, "Local Symmetry Features in Image Processing", Diss. No. 179, Linköping University, Sweden.
- [6] R. Wilson, H. Knutsson. "Uncertainty and Inference in the Visual System". *IEEE Systems, Man and Cybernetics* Volume 18, No 2, 1988.
- [7] H. Knutsson, G. H. Granlund. "Compact Association Representation of Structural Information". *Report LiTH-ISY-I-0931*, Linköping University, 1988.