Symmetry-Based Image Coding

Pierangela Cicconi, Riccardo Leonardi*and Murat Kunt
Signal Processing Laboratory
EPFL-Ecublens (DE)
CH-1015 Lausanne, Switzerland

Abstract

A novel coding technique which proposes the use of symmetry to reduce redundancy in images is presented. Axes of symmetry are extracted using the Principal Axes of Inertia theory and the technique is extended to non-symmetric images by the introduction of a Coefficient of Symmetry. One part of the images is then linearly predicted with respect to the chosen axis. The method is implemented in a block-based fashion in order to adapt to local symmetries on the image data. An image representation and a coding strategy is illustrated, and results are presented on real static images.

1 Introduction

Recent efforts towards reduction of redundancy in images showed the relevance of modeling them as a combination of nonstationary visual primitives [1]. This lead to significant improvement in compression. As an example, segmentation-based or contour-texture-based coding demonstrated superior performance with respect to classical waveform coding whose fundamental assumption remains the stationarity of the source.
If contours or areas with uniform textural characteristics seemed adequate to be chosen as primitives, more elaborate patterns or properties that can be found in natural scenes may be considered for better performance as well.
In this work, we propose a novel technique which suggests the use of symmetry. The introduction of this geometric property is dictated by the fact that natural objects often give rise to the human sensation of symmetry. This sense of symmetry is so strong that most man-made objects are symmetric, and this concept is more general than the strict mathematical notion [2]. For instance, a picture of a human face is considered highly symmetric, although it is not symmetric in the mathematical sense.
The outline of the paper is the following. Section 2 explains how to find axes of symmetry in images; section 3 generalizes this technique to non-symmetric images and introduces a way to measure the degree of symmetry. In section 4 a scheme is proposed to build a linear prediction model of one part of the image with respect to an axis of symmetry. This scheme is then used to code the luminance information of an image by limiting it to the location of its symmetry axes and a limited amount of the original information. In section 6 coding results are presented on real static images, while section 7 concludes with an outlook on future improvements.

2 Axes of Symmetry Extraction

Various techniques for extracting axes of symmetry of objects have been proposed, mainly for pattern recognition applications [3, 4].
In [3], Freidberg describes a technique which finds the axes of skewed symmetry. In [4], Marola presents an algorithm which is based on the identification of the centroids of a given object and other related sets of points, followed by a maximization of a specially defined coefficient of symmetry.
In the following, we propose a simple and efficient technique, which identifies axes of symmetry to the Principal Axes of Inertia (PAI) of a rigid body [5]. In this section we present a technique for extracting the PAIs. This method is

*Currently at Signals & Communications Lab., Dept. of Electronics for Automation, University of Brescia, Italy

 0-8194-1018-7/92/$4.00

then applied to extract symmetry axes on real test images.

Def 2.1 *A set of three mutually perpendicular axes, fixed with respect to the body, rotating with it and such that the products of inertia with respect to them are zero, are called Principal Axes of Inertia (PAI) of the body.*

Def 2.2 *A PAI is such that, if a rigid body rotates around it, the direction of the Movement Quantity Momentum* Ω *is equal to the direction of the Angular Velocity* ω.

$$\Omega = \lambda\omega \tag{1}$$

If $\tilde{\mathbf{I}}$ is the symmetric inertia matrix

$$\tilde{\mathbf{I}} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \tag{2}$$

with

$$\left\{ \begin{array}{l} I_{xx} = \sum m_v (y^2{}_v + z^2{}_v) \\ I_{yy} = \sum m_v (x^2{}_v + z^2{}_v) \\ I_{zz} = \sum m_v (x^2{}_v + y^2{}_v) \\ I_{xy} = I_{yx} = -\sum m_v x_v y_v \\ I_{xz} = I_{zx} = -\sum m_v x_v z_v \\ I_{yz} = I_{zy} = -\sum m_v y_v z_v \end{array} \right\} \tag{3}$$

and m_v is the mass of point (x_v, y_v, z_v) in 3D space, then the Movement Quantity Momentum Ω can be expressed as

$$\Omega = \tilde{\mathbf{I}}\omega \tag{4}$$

Using (1) and (4), we obtain the eigenvalue decomposition system of three equations

$$(\tilde{\mathbf{I}} - \lambda\mathbf{I})\omega = 0 \; . \tag{5}$$

where $\mathbf{I}$ is the identity matrix. In order to have solutions different from the trivial one $\begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}^T = 0$ we impose that

$$det(\tilde{\mathbf{I}} - \lambda\mathbf{I}) = 0 \; . \tag{6}$$

which results in a cubic equation with roots $I_{1,2,3}$ that are called *Principal Momenta of Inertia.* Substituting these solutions in Eq.(5), we obtain the directions of the PAIs.

The previously described technique to find the PAIs of a rigid body can then be used to detect the PAIs of the 3D object defined by the image plane and the luminance function, with m_v set to 1. These PAIs can then be labelled as likely axes of symmetry. In fact, it can be demonstrated that, if an object has an axis of symmetry, this axis is also a PAI. Conversely, if an object does not have a real symmetry, the PAI partitions the body "optimally" into two quasi-symmetric parts. [5]

As an example of application, Fig. 1 shows the results of the PAI extraction to different block partitions of the test image "Lena" (256 × 256 pixels, 8 bit/pixel).

3 Non-Symmetric Images and Coefficient of Symmetry

In section 2, we described a technique which finds the PAIs of the object defined by the image plane and the luminance function. In this section, a Coefficient of Symmetry is defined, which allows to measure the degree of symmetry associated to each PAI [4].

Let F be an image with an axis of symmetry and $g(x, y)$ its associated luminance function. If $P(x, y)$ and $\bar{P}(\bar{x}, \bar{y})$ are two points symmetric with respect to a certain axis d (see Fig. 2), we define the quantity $\beta \geq 0$ as follows

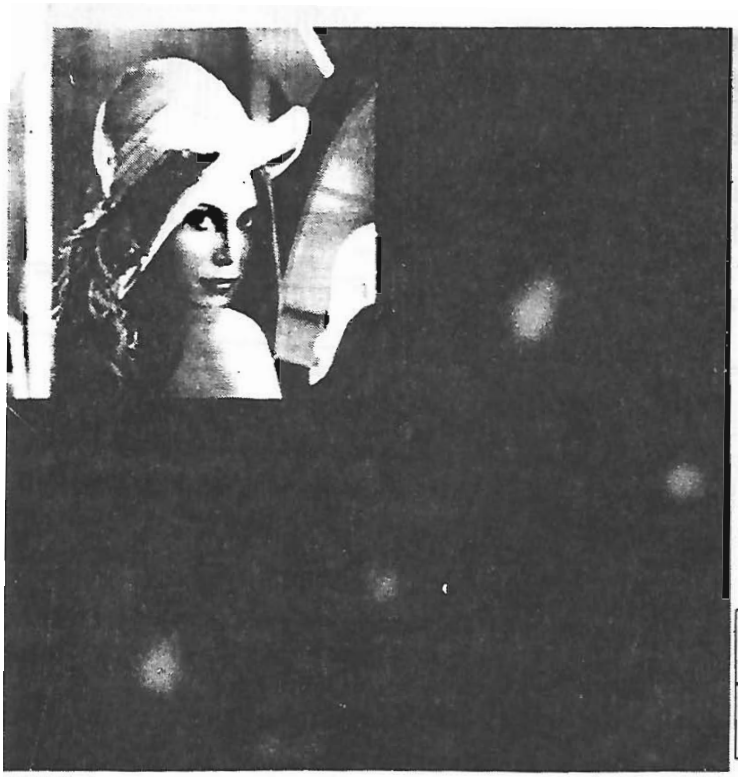

Figure 1: Principal Axes of Inertia Extraction a) original image b) PAIs on 8 × 8 blocks c) PAIs on 64 × 64 blocks

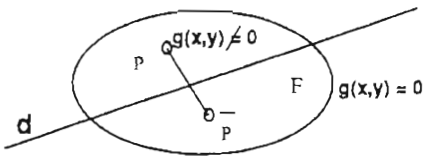

Figure 2:

Def 3.1

$$\beta = 1 - \frac{\frac{1}{2}\iint [g(x,y) - g(\bar{x},\bar{y})]^2 dxdy}{\iint g^2(x,y)dxdy} = \frac{\iint g(x,y)g(\bar{x},\bar{y})dxdy}{\iint g^2(x,y)dxdy} \quad (7)$$

with

$$0 \leq \beta \leq 1 \quad (8)$$

if $\beta = 0$ the axis is outside F
if $\beta = 1$ F is symmetric
β is called Coefficient of Symmetry associated to the axis.

By retaining the PAI associated to the highest symmetry measure, a given image will be separated into the two most quasi-symmetric parts, according to the coefficient of symmetry measure. Figure 3 shows the graphs of the Coefficients of Symmetry associated to the PAIs evaluated for different block partitions of the original test image "Lena".

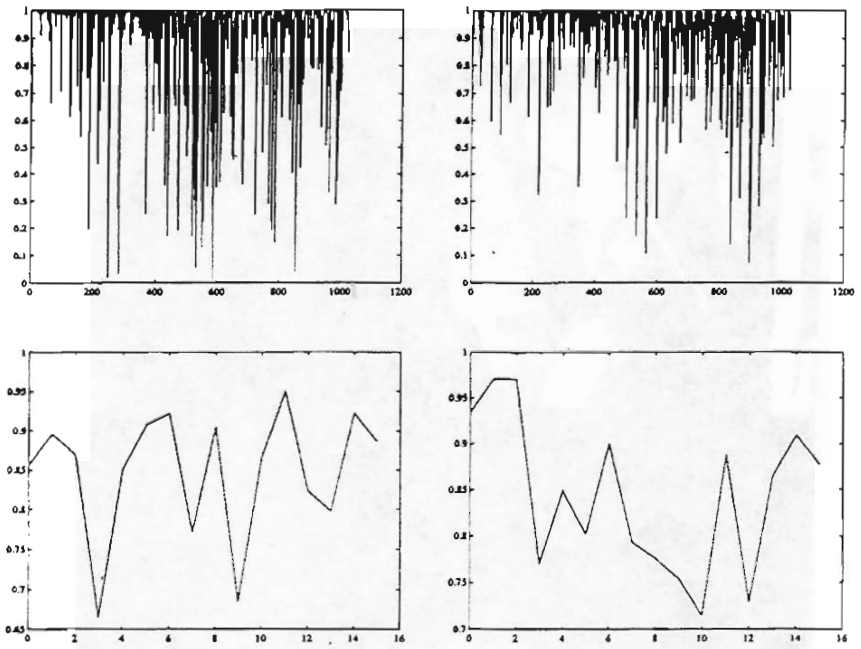

Figure 3: Graphs of Coefficients of Symmetry; from left to right: first & second row of 8 x 8 blocks; first & second row of 64 x 64 blocks

4 Linear Prediction

Once a symmetry axis has been found, we propose to build the linear prediction of one part of the image from the opposite part measured symmetrically from the chosen PAI. In other words, each pixel belonging to one part of the image is linearly predicted from pixels symmetrically chosen on the other side of the selected PAI.
In general, Linear Prediction [6] is a model in which a signal s_n is considered to be the output of a system with some unknown input u_n such that the following relation holds:

$$s_n = -\sum_{k=1}^{p} a_k s_{n-k} + G \sum_{l=0}^{q} b_l u_{n-l} \,. \tag{9}$$

$$b_0 = 1 \tag{10}$$

where a_k, $1 \leq k \leq p$, b_l, $1 \leq l \leq q$, and the gain G are the parameters of the prediction. Equation (9) states that the "output" s_n is a linear function of past outputs and present and past inputs. That is, the signal s_n is predictable from linear combinations of past outputs and inputs.
The problem is to determine the predictor coefficients a_k and the gain G in some manner. Using an intuitive least squares approach and assuming that the input u_n is totally unknown, the signal s_n can be predicted only approximately from a linearly weighted summation of past samples. Let this approximation of s_n be $\tilde{s}_n$, where

$$\tilde{s}_n = -\sum_{k=1}^{p} a_k s_{n-k} \tag{11}$$

In the method of least squares the parameters a_k are obtained as a result of the minimization of the mean or total error with respect to each of the parameters. In this way, we obtain the following set of equations:

$$\sum_{k=1}^{p} a_k \sum_{n} s_{n-k} s_{n-i} = -\sum_{n} s_n s_{n-i}, \quad \text{for} \quad 1 \leq i \leq p \,. \tag{12}$$

The predictor coefficients a_k, for $1 \leq k \leq p$, can be computed by solving a set of p equations with p unknowns. There exist several standard methods for performing the necessary computations [7, 8]. In our system, we make use of the method attributed to Durbin [9] which is twice as fast as [7]; a pixel can be obtained as a linear combination of a set of pixels which are in quasi-symmetric position with respect to the chosen PAI. In the following sections, results are presented using a Linear Prediction of order zero. That is, a pixel is obtained from the one which is symmetrically located with respect to the chosen PAI.

5 Coding

The method which has been described in the previous sections was implemented in a block-based fashion in order to adapt to local symmetries of the image data. In the following, the scheme of a possible coder, together with a coding strategy specifying the PAIs and the luminance on one side of the PAIs will be described.
Let $I(x,y) = I_1(x,y) \cup ... \cup I_n(x,y)$, $I_i(x,y) \cap I_j(x,y) = 0$, $\forall i,j$, $i \neq j$ be a partition of the original image, where $I_i(x,y)$ is a generic subblock of fixed $M \times M$ size.
Let $I_i(x,y) = I_{i_1}(x,y) \cup I_{i_2}(x,y)$, $I_{i_1}(x,y) \cap I_{i_2}(x,y) = 0$ be a partition of $I_i(x,y)$, where $I_{i_j}(x,y)$ is the luminance function on one side of the PAI defined on the i-th block (see Fig. 4a). Using this notation, the proposed coding scheme is illustrated in Fig. 4b.

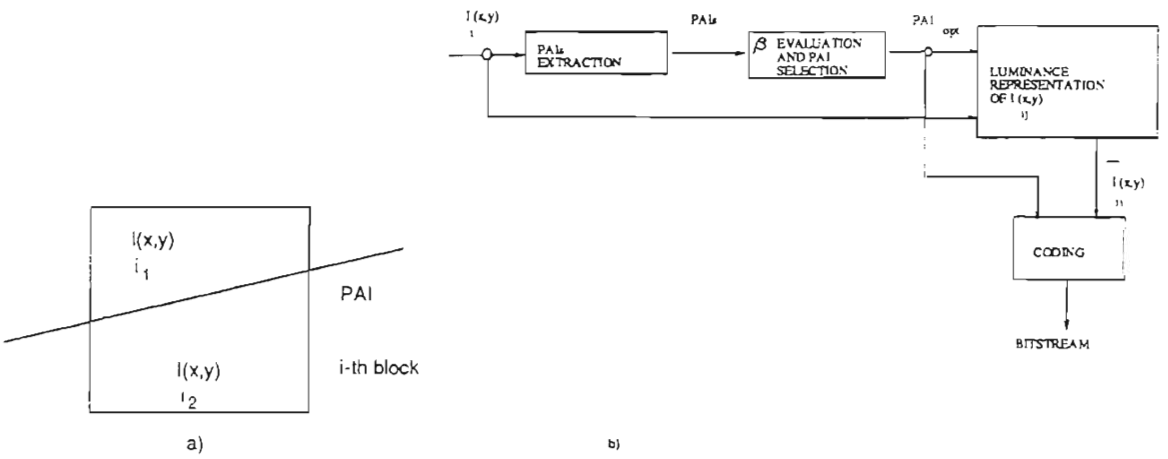

Figure 4: a) generic block and its axis of symmetry; b) encoder block diagram

For each block the technique described in section 2 is applied, which finds the two PAIs on the basis of the luminance function. Then the Coefficient of Symmetry associated to each PAI is evaluated; according to this measure, the PAI corresponding to the highest value is selected.
The luminance representation of I_{i_j} is obtained by coding separately the PAI location and orientation, the luminance on one side of it, $I_{i_1}(x,y)$, and the prediction error with respect to $I_{i_2}(x,y)$, using the PAI and $I_{i_1}(x,y)$ coded information.
In the following subsections, more details on the luminance representation and the coding strategy will be outlined.

5.1 Luminance Representation

The problem at hand is the following: we are looking for a representation of $I(x,y)$ which approximates the given gray values and describes them in a compact form allowing efficient data compression.
In other words,

$$I(x,y) \rightsquigarrow \hat{I}(x,y)\ , \tag{13}$$

where

$$\hat{I}(x,y) = a_1\varphi_1(x,y) + \cdots + a_n\varphi_n(x,y)\ . \tag{14}$$

The functions $\varphi_i(x,y)$ are called the *Basis Functional Set*, which for the task of approximation must be complete, according to the *Weierstrass Approximation Theorem.*
One possible choice is to use the *Polynomial Basis Functions* [10]:

$$\varphi_i(x,y) = x^{k(i)}y^{l(i)}, \quad \text{for} \;\; k(i)+l(i) \leq n \; . \tag{15}$$

A polynomial representation has, at least, the following advantages:

1. Images predominantly consist of slowly varying surfaces which are well represented by polynomials.
2. Polynomials correspond to very simple mathematical expressions.

In order to find the coefficients of the approximation, it is necessary to define an error minimization criterion. The most commonly used is the one attributed to Gauss (1795) which minimizes the L_2 - norm:

$$d(I,\hat{I}) = \sum_x \sum_y \left[I(x,y) - \hat{I}(x,y)\right]^2 \; . \tag{16}$$

The solution to this problem involves a pseudo-inverse formulation and can directly be obtained. The minimization of $d(I,\hat{I})$ leads to the so called system of *Normal Equations*:

$$\sum_{n=1}^{N} a_n \sum_x \sum_y \varphi_n(x,y)\varphi_q(x,y) = \sum_x \sum_y \hat{I}(x,y)\varphi_q(x,y), \quad q = 1,2,\cdots,N \tag{17}$$

the geometric interpretation of which is illustrated in Fig. 5.

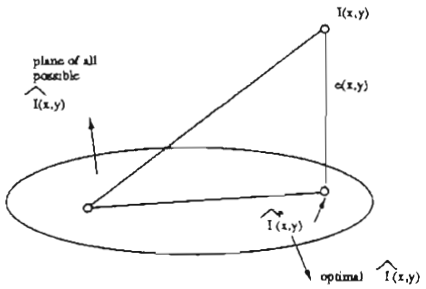

Figure 5:

In most cases, when the number of pixels is greater than the dimension of the Basis Functional Set, an elegant solution to the problem is given by the *Gauss-Jordan Algorithm*, which distributes the fitting error to minimize the squared error.
In Fig. 6, the reconstructed version of the test image "Lena" is shown for different block dimensions when using a 2-dimensional second order polynomial function:

$$g(x,y) = a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 + a_5xy \; . \tag{18}$$

5.2 Coding Strategy

In this subsection, a coding strategy specifying the selected PAIs and the luminance on one side of the PAIs is defined.
The information to be coded are the coefficients of the polynomial approximation, together with two parameters specifying the PAI location and orientation.

Figure 6: Second-order Polynomial Approximation a) original image b) Approximation on 8 × 8 blocks c) Approximation on 64 × 64 blocks

5.2.1 Quantization of Coefficients

A transform allows to map the original pixel domain in a transformation domain. However, a transform alone does not result in any data compression. Only in connection with a quantization of the transformed coefficients reduction in the amount of information is obtained.
Originally, the polynomial approximation is computed using a floating point format for all values; any quantization of the coefficients leads to an increase in the reconstruction error. The simplest and most common form of quantizer that could be used is the zero-memory quantizer [11]. Its output value is determined by the quantizer only from one corresponding input sample, independently of the values taken on by earlier (or later) samples applied to the quantizer input. Rigorously speaking,

Def 5.1 *A zero-memory N-point quantizer Q is defined by N+1 decision levels* $x_0, \cdots, x_N$ *and N output points* $y_1, \cdots, y_N$. *If an input sample* $x_i \in R_i = \{x_{i-1}, x_i\}$. *the quantizer produces* $y_i \in R_i$ *(see Fig. 7b).*

In general, uniform quantization is not the most effective way to achieve optimal performance. The reason is that quantization with minimum square error (optimum quantization) depends on the probability density function of the coefficients to be quantized. Both the decision and the representation levels should then be optimized [12], with computationally expensive methods. A simplification consists in assuming a uniform distribution of the coefficients within the quantization intervals [13]. The representation levels are centered inside the intervals, and the quantization intervals are adapted to the probability density function. To decide on possible quantization schemes we need to investigate the amplitude distribution of the coefficients [14]. Figure 7a shows the highly peaked distribution of the coefficient amplitudes for the test image "Lena", 8 × 8 blocks.
For a given number of quantization intervals, taking into account the input probability density, nonuniform spacing of the decision levels can yield lower quantizing noise and less sensitivity to variations in input signal statistics. An effective technique for studying nonuniform quantization, used in [15], is to model the quantizer as a memoryless nonlinearity $F(x)$, the "compressor", followed by a uniform quantizer, as shown in Fig. 8a. The effect is to allocate more quantized levels to the high probability lower amplitudes, and fewer levels to the less frequently occurring

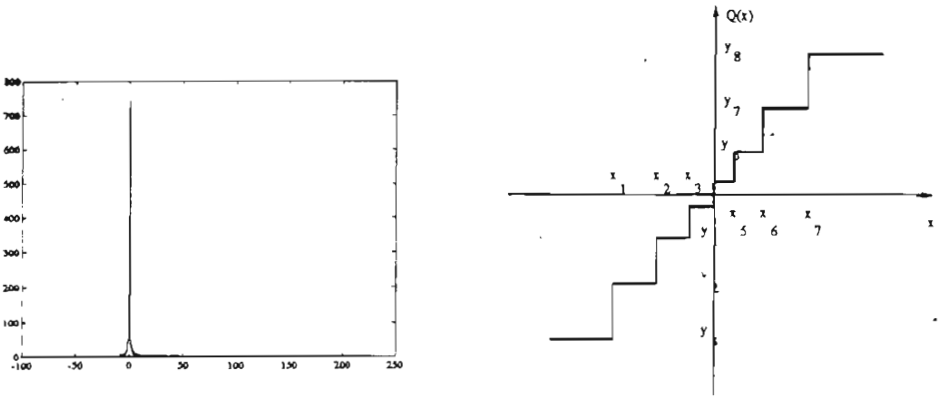

Figure 7: a) Distribution of the coefficient amplitudes; b) Input-output characteristic of a quantizer

higher amplitudes.
The coding strategy applied in the proposed system is the following. In the range which contains 90% of the data, the coefficients are linearly quantized. The peaked shape of the histogram is exploited by variable length coding (Huffmann Coding in our case). Outside the interval, nonuniform quantization, achieved by using a compander, takes into account the dynamic range of the data without introducing an excessive number of representation levels. The quantization scheme is shown in Fig. 8b.

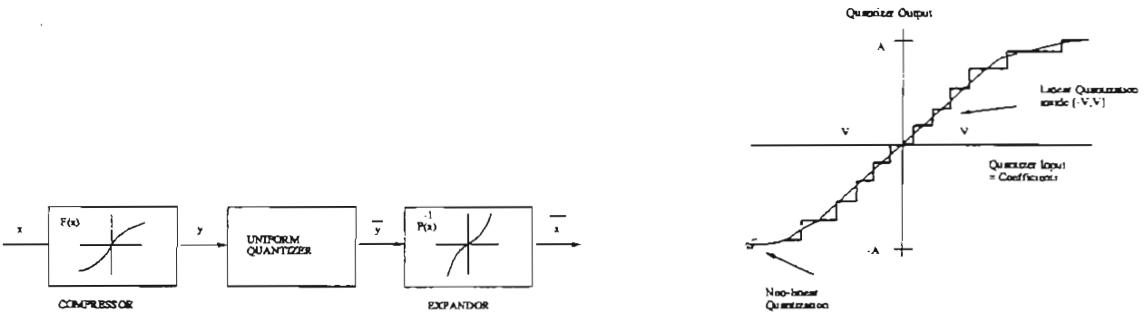

Figure 8: a) Companding model of nonuniform quantization; b) Quantization scheme

5.2.2 Coding of the PAIs

A set of image points (x, y), which lie on a straight line, can be defined by a relation f, such that

$$f[(m, c), (x, y)] = y - mx - c = 0 \ . \tag{19}$$

where m and c are two parameters, the slope and intercept, which describe the straight line.
In [16], it is suggested that straight lines might better be parameterized by the length ρ, and orientation θ of the normal vector to the line from the image origin (see Fig. 9a), where

$$\rho = x \cos\theta + y \cos\theta \tag{20}$$

This representation has advantages over the (m, c) parameterization which has a singularity for lines with large slopes, that is, for $m \to \infty$. Another parameterization [17] of straight lines consists of a pairing of the two coordinates (s_1, s_2) where the line intersects the border of the block boundaries (see Fig. 9b). These points are given by their distance along the perimeter of the block, where the distance is measured counterclockwise along the block starting at the origin. It is clear that, for a block of size $x_{max} \times y_{max}$, the values of ρ and θ are bounded by

$$-R < \rho < R \quad and \quad 0 < \theta < \pi \tag{21}$$

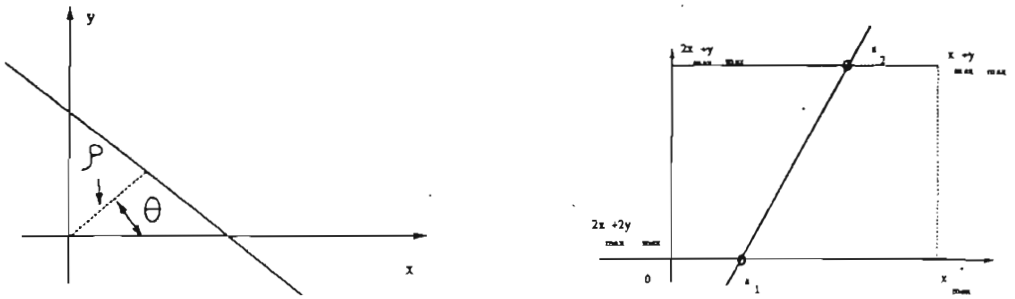

Figure 9: a) ρ-θ line representation; b) $x_{max} \times y_{max}$ block perimeter intersection line representation

where

$$R = \sqrt{x_{max}^2 + y_{max}^2} \tag{22}$$

The histograms of θ and ρ for the test image "Lena", 8 x 8 blocks, are shown in Fig. 10 when the origin of the coordinates coincides with the center of each block. In this case, a uniform quantization, confined to the region $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$, $-\frac{R}{2} \leq \rho \leq \frac{R}{2}$, of the two line parameters (ρ, θ) has been used.

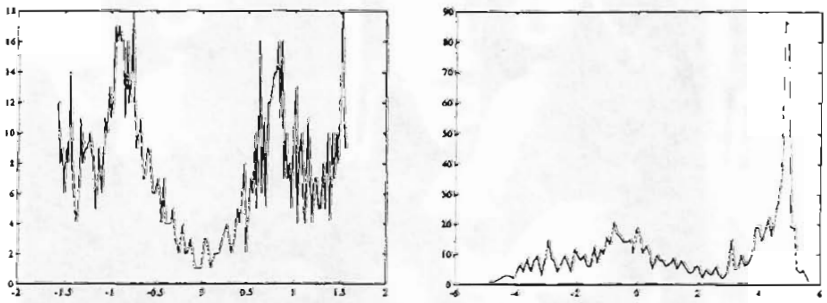

Figure 10: a) histogram of θ; b) histogram of ρ

6 Results

In this section, results of the described coding technique are presented.
According to (18) and to Fig. 8b, the values of the parameters specifying the quantization strategy are the following. a_0 is uniform quantized without companding, because of its almost uniform probability distribution and its wide dynamic range. As far as the coefficients $a_1, ..., a_5$, the values of V are shown in Table 1. Outside the interval $[-V, V]$, the following companding curve was used:

$$F(x_i) = \frac{V \ln(1 + \mu x_i / V)}{\ln(1 + \mu)}, \quad \mu = 255 . \tag{23}$$

In Fig. 11, the performance of the system on the test image "Lena", 8 x 8 blocks, is shown, where 130 and 100 quantization levels have been used for coding the coefficients (Fig.11b, Fig.11c respectively). In both cases, 100 quantization levels have been used for the uniform quantization of ρ and θ. Huffman coding has been used to code the quantization information, obtaining a compression rate of 10:1 and 15:1 respectively (Fig.11b, Fig.11c), without coding of the prediction error (see section 4).
It should be pointed out that, both the polynomial approximation and quantization contribute, though to different extents, to the degradation of the reconstructed image. The values of PSNR (Peak Signal-to-Noise Ratio) evaluated for the two reconstructed images considered before and for a non-quantized image are shown in Table 2.

Coefficients	V value
1	14.34
2	8.46
3	1.62
4	3.63
5	2.82

Table 1: The values of V for the coefficients $a_1, \cdots, a_5$

Figure 11: Symmetry-based image coding results a) original image b) reconstructed image, compression ratio=10:1 c) reconstructed image, compression ratio=15:1

Reconstructed Image	PSNR value
non-quantized	27.10
quantized, comp. ratio=10:1	25.41
quantized, comp. ratio=15:1	24.49

Table 2: The values of PSNR for differently reconstructed images

7 Conclusions

In this work, we presented a novel technique which suggests the use of symmetry to reduce the redundancy in images. Among the several techniques which have been proposed to extract axes of symmetry of objects, we propose to use the Principal Axes of Inertia (PAI) of a rigid body. The results were used to separate an image domain into two quasi symmetric parts with the definition of a Coefficient of Symmetry. The method was implemented in a block-based fashion in order to adapt to local symmetries on the image data. An image representation and coding strategy was defined. The luminance function on one side of the PAI is approximated by means of a second order bidimensional polynomial function. The PAIs are parameterized by their spatial position and orientation. As far as the coding strategy is concerned, the wide dynamic range of the coefficients of the polynomial approximation is reduced by means of a nonuniform companding; the companded coefficients, together with the line parameters, are then uniformly quantized.
Coding by symmetries is a novel and open area of research; several aspects of the proposed system and its possible extensions are currently under investigation. In particular, a more sophisticated coding strategy is under investigation and will soon be presented [18]. The extension of this technique for a symmetry-based image segmentation system will be the object of future research.

References

[1] M. Kunt, A. Ikonomopoulos, and M. Kocher. Second-generation image-coding techniques. *Proceedings of the IEEE*, 73(4), April 1985.

[2] F. Attneave. Informational aspects of visual perception. *Psychological Review*, 61:183–193, 1954.

[3] S.A. Freidberg. Finding axis of skewed symmetry. *Comput. Graphics Image Processing*, 34:138–155, 1986.

[4] G. Marola. On the detection of the axes of symmetry of symmetric and almost symmetric planar images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-11(1), January 1989.

[5] M.S. Spiegel. *Meccanica Razionale*. Collana Schaum - Etas Libri, 1974.

[6] J. Makhoul. Linear prediction: a tutorial review. *Proceedings of the IEEE*, 63(4), April 1975.

[7] N. Levinson. The Wiener rms (root mean square) error criterium in filter design and prediction. *J. Math. Phys.*, 25(4):261–278, 1947.

[8] E.A. Robinson. *Statistical Communication and Detection*. New York: Hafner, 1967.

[9] J. Durbin. The fitting of time-series models. *Rev. Inst. Int. Statist.*, 28(3):233–243, 1960.

[10] M. Eden, M. Unser, and R. Leonardi. Polynomial representation of pictures. *Signal Processing*, 10:385–393, 1986.

[11] A. Gersho. Principles of quantization. *IEEE Transactions on Circuits and Systems*, CAS-25(7), July 1978.

[12] J. Max. Quantization for minimum distorsion. *IRE Trans. Inf. Theory*, 6:7–12, 1960.

[13] N.S. Jayant and P. Noll. *Digital coding of waveforms*. Prentice-Hall, Englewood Cliffs, NJ, 1984.

[14] M. Gilge, T. Engelhardt, and R. Mehlan. Coding of arbitrarily shaped image segments based on a generalized orthogonal transform. *Signal Processing: Image Communication*, 1:153–180, 1989.

[15] W.R. Bennett. Spectrum of quantized signals. *Bell System Technical Journal*, 27:446–472, July 1948.

[16] J. Illingworth and J. Kittler. A survey of the Hough Transform. *Computer Vision, Graphics, and Image Processing*, 44:87–116, 1988.

[17] R.S. Wallace. A modified hough transforn for lines. In *IEEE Computer Vision and Pattern Recognition (CVPR) Conference San Francisco, California*, pages 665–667, June 19-23 1985.

[18] P. Cicconi, M. Mattavelli, A. Nicoulin, and M. Kunt. Image compression by local symmetries extraction, and entropy coding with vector quantization and arithmetic coding. Submitted to IEEE International Conference on Communications, Geneva 1993.