# 3D invariants and their application to object recognition

Gilles Burel, Hugues Henocq

Signal Processing, Vol. 45, No. 1, July 1995

# 3D invariants and their application to object recognition

Gilles BUREL & Hugues HENOCQ

Thomson Broadband Systems, Centre de Rennes
Avenue de Belle Fontaine, 35510 Cesson-Sévigné, France

## Abstract

*Although recognition of objects from* 2D *projections (i.e. images) has been widely studied among the image processing community, little research has been devoted to recognition using* 3D *information. A general approach for deriving* 3D *invariants is proposed in this paper. These invariants can be used as input to a statistical classifier, such as a k-nearest-neighbours algorithm or a neural network. The approach consists of decomposing the object onto an orthonormal basis composed of the eigenvectors of the angular momentum operator from quantum mechanics. Then, using Clebsch-Gordan coefficients, contravariant tensors of order* 1 *are constructed, and* 3D *invariants are obtained by tensor contraction. The approach offers an alternative to structural methods for 3D object description and recognition. Experimental results are provided to illustrate the method.*

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



# Résumé

*La reconnaissance d'objets par l'intermédiaire de projections 2D (c'est-à-dire d'images) du monde tridimensionnel a fait l'objet de nombreuses études au sein de la communauté scientifique. Par contre, peu de travaux ont été consacrés à la reconnaissance de formes à partir d'une information 3D. Ceci est dû à la complexité nettement accrue des transformations que peut subir un objet en trois dimensions. Dans cet article, une approche générale pour l'obtention d'invariants 3D est proposée. L'approche consiste à décomposer la forme 3D sur une base orthonormée formée des vecteurs propres de l'opérateur moment cinétique de la mécanique quantique. Ensuite, grâce aux coefficients de Clebsch-Gordan, des tenseurs contravariants d'ordre 1 sont construits, ce qui permet, par contraction tensorielle, d'obtenir des invariants. Des résultats expérimentaux sont fournis pour valider et illustrer la méthode.*

# Keywords

*Pattern Recognition, Rotation Invariance,* 3D *Invariance, Tensor Theory, Spherical Harmonics, Angular Momentum.*

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



# 1 Introduction

Object recognition is a key topic in computer vision and automatic scene analysis. A great deal of work has been dedicated to 2D pattern recognition, either using structural or statistical approaches. By "2D pattern recognition", we refer to methods that use 2D information (i.e. an image) of an object. The object itself may be 2D or 3D.

In the structural approach, features such as holes and corners are detected on the 2D image of the object and linked to provide a structural description of it [1]. The structural description is generally a graph whose nodes represent the features and whose arcs represent the spatial organization of these features (and more generally the relationships between them). Recognition is then achieved by graph matching techniques using a database of models.

In the statistical approach, a feature vector of fixed length is computed from the 2D image of the object. The components of the feature vector may be, for instance, the moments of Hu [8], or Fourier descriptors [6] [7]. Recognition is then performed by a classifier, such as a k-nearest-neighbours algorithm, or a neural network [11].

Few studies have been devoted to 3D pattern recognition by using full 3D information. However, nowadays it is becoming easier and easier to obtain 3D object information, thanks to the development of techniques such as computed tomography, cineangiography [15], nuclear magnetic resonance imaging, and active laser range finders.

Recent work has been devoted to the development of 3D moment invariants [9] [12]. These moments, which are invariant with respect to 3D translation and rotation of the object, can be used as the components of a feature vector for a statistical approach. However, they are difficult to derive, and only a small number of them is explicitly derived. Thus, the moment-based approach does not provide long feature vectors, which may be useful for discriminating close shapes.

We propose an approach based on the results of quantum mechanics for deriving 3D invariants. These invariants are derived in a simple and systematic way, and it is shown that moment invariants can be obtained as a degenerate case of our approach. Furthermore, hundreds of invariants can be derived, allowing the obtention of long feature vectors. The approach may be used for automatic object recognition in domains where 3D information is available. For example, in the medical field, it could be used for the automatic classification of bones (figure 7) or for detection of pathological heart deformations. In robotics, a robot equipped

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



with a camera could use computer tomographic reconstruction to construct a 3D representation of an object, and then use 3D invariants to recognize it. The proposed approach could also be used for associative access to a database of 3D objects: it would search for the shapes which are the closest to a shape selected by an operator.

# 2 Principle of the approach

A 3D object can be described by a function $\Psi(x, y, z)$, where $x, y, z$ are the coordinates in the 3D world. This function may be either binary (1 inside the object and 0 outside), real valued (for instance it may represent the local density if such information is available), or complex valued (for instance it may represent radar backscatter coefficients).

A good feature vector should be invariant when the object is translated or rotated. Translational invariance is easy to obtain, for instance by placing the origin of the 3D coordinates at the center of gravity of the object. If required, volume normalization can also be performed, for instance by choosing as unit of length the distance between the origin and the farthest point of the object. Hence, as for moments, we have to focus on rotational invariance. Of course, for some objects, the rotation may be fixed by diagonalizing the inertia matrix. However, many real objects have similar inertia values on at least two principal axes (for example, this is the case for the vertebrae shown in figure 11). Hence, the determination of orientation using this method does not provide results accurate enough to discriminate between similar shapes (average estimation error of 6 degrees for the vertebrae).

In this paper, we will use the following notation:

- $i = \sqrt{-1}$
- $\mathcal{R}$: the set of real numbers
- $\mathcal{C}$: the set of complex numbers
- $\mathcal{F}$: the space of differentiable functions from $\mathcal{R}^3$ to $\mathcal{C}$, with finite energy (i.e. $\int\int\int |\Psi(x, y, z)|^2 \, dx \, dy \, dz < \infty$)
- $\mathcal{F}_\mathcal{S}$: the space of differentiable functions from $[0, \pi]$x$[0, 2\pi]$ to $\mathcal{C}$, with finite energy (i.e. $\int d\phi \int sin\theta \, d\theta |\Psi(\theta, \phi)|^2 < \infty$)
- $A^H$: the Hermitian transpose of matrix $A$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



- $\mathcal{O}_{\mathcal{V}}$: the space of linear operators from $\mathcal{V}$ to $\mathcal{V}$
- $L_x, L_y, L_z$: angular momentum operator along $X, Y$, and $Z$ axes
- $\mathbf{L}^2$: "square of angular momentum" operator
- $\langle u|v\rangle$: the scalar product of vectors $u$ and $v$

Figure 1 shows the conventions for spherical coordinates.

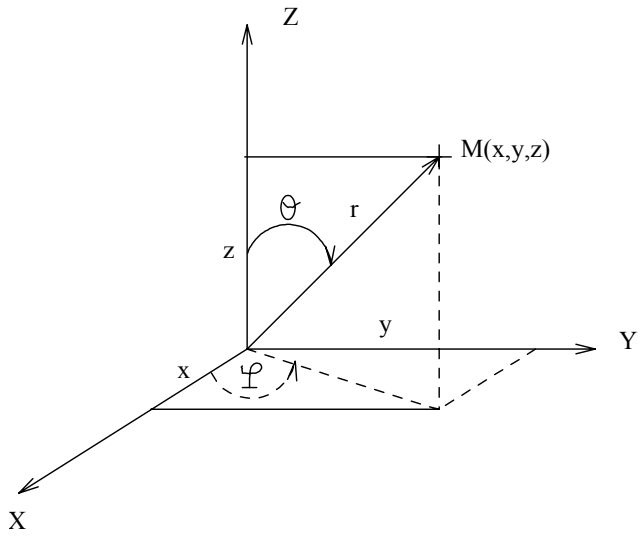


Figure 1: Conventions for spherical coordinates

Then we have:

$$\begin{aligned} x &= r\ sin\theta\ cos\phi \\ y &= r\ sin\theta\ sin\phi \\ z &= r\ cos\theta \end{aligned}$$

with:

$$\begin{aligned} &r \geq 0 \\ &0 \leq \theta \leq \pi \\ &0 \leq \phi < 2\pi \end{aligned}$$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



Any 3D object can be considered a vector $|\Psi\rangle$ of $\mathcal{F}$ (using Dirac notation for vectors). $\mathcal{F}$ can also be considered a space of wave functions, as they are defined in quantum mechanics. A rotation in $\mathcal{R}^3$ is a linear operator $R_3$, transforming $(x, y, z)$ into $(x', y', z')$. This operator is unitary ($R_3^H = R_3^{-1}$). To each operator $R_3$ of $\mathcal{O}_{\mathcal{R}^3}$, we can associate a linear operator $R$ of $\mathcal{O}_{\mathcal{F}}$. $R$ transforms $|\Psi\rangle$ in $|\Psi'\rangle = R|\Psi\rangle$ such that $\forall(x, y, z), \Psi'(x', y', z') = \Psi(x, y, z)$. It can be proved that $R$ is unitary [14] [16].

There exists a strong link between the rotational operator and the angular momentum operator used in quantum mechanics [4] [14]:

$$R = e^{-i\alpha(u_x L_x + u_y L_y + u_z L_z)} \tag{1}$$

where $\alpha$ is the angle of rotation, and $[u_x\ u_y\ u_z]^T$ is the axis of rotation (unitary vector).

Due to its major role in quantum mechanics, the angular momentum has been widely studied. Hence, equation 1 suggests that it could be interesting to use some results of quantum mechanics to deal with rotations of 3D objects.

Figure 2 provides an overview of the proposed approach. The object $|\Psi\rangle$ is decomposed on the orthonormal basis of $\mathcal{F}$ composed of the eigenvectors of the angular momentum (section 3). In this basis, the rotational operator becomes irreducible, and it is possible to construct contravariant tensors of order 1 using the components of $|\Psi\rangle$. Then, using tensor contraction, we obtain invariants (section 4). Some properties of the invariants are given in section 5.

We first study the objects represented by a vector of $\mathcal{F}_{\mathcal{S}}$. For example, $\Psi(\theta, \phi)$ may be the distance between the origin and the farthest point of the object in the direction $(\theta, \phi)$. Of course, for some objects of complex shape (objects with holes, for instance), there is a certain loss of shape information in this kind of description. Afterwards, we extend our method to objects described by a vector of $\mathcal{F}$ (section 6), and we show that moment invariants can be derived as a particular case of the approach. Finally, in section 7, we provide experimental results to illustrate the method.

# 3 Decomposition on eigenvectors of angular momentum

In quantum mechanics, observations concerning angular momentum are defined by the action of two operators [4] [14]: $\mathbf{L}^2$ (square of angular momentum) and $L_Z$ (projection of angular momentum along the $Z$ axis). These operators are defined

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



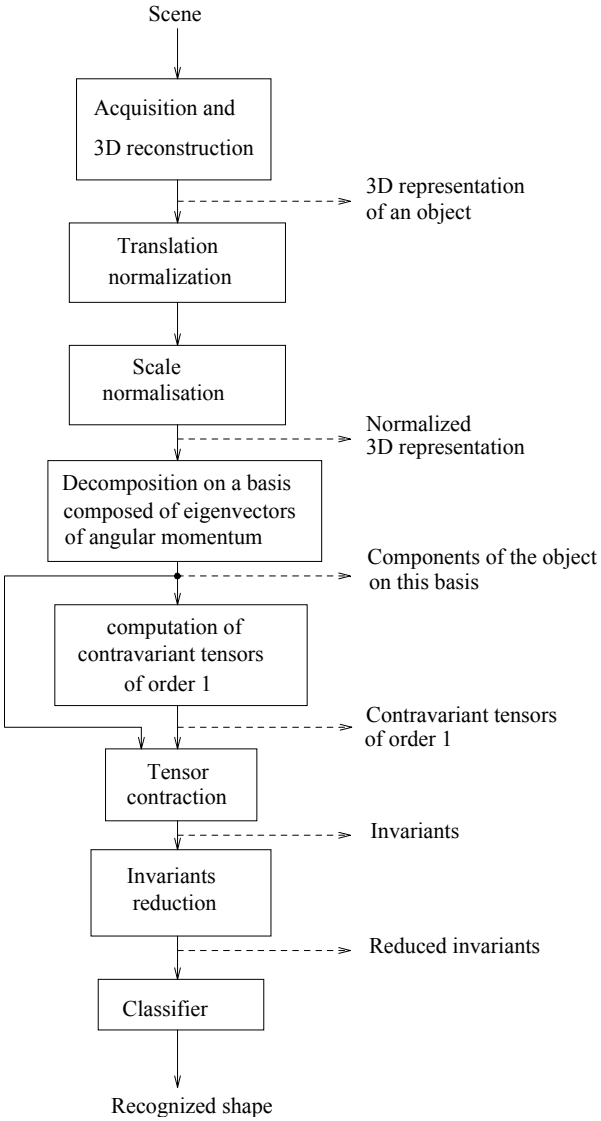


Figure 2: Overview of the approach

below:

$$\mathbf{L}^2 = -\left(\frac{\partial^2}{\partial\theta^2} + \frac{1}{tg\theta}\frac{\partial}{\partial\theta} + \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \quad (2)$$

$$L_Z = \frac{1}{i}\frac{\partial}{\partial\phi} \quad (3)$$

where the Planck constant has been omitted (since it is only a multiplicative factor, it is not useful for our purpose).

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



$\mathbf{L}^2$ and $L_Z$ are commuting Hermitian operators. Hence, there exists an orthonormal basis composed of their common eigenvectors. These eigenvectors are known as the "spherical harmonics": $|Y_{lm}\rangle$. They can be computed using recurrent equations (see appendix 1), and are defined for $l = 0, 1, 2, ..., \infty$ and $m = -l, ..., l-1, l$.

The decomposition of the shape $|\Psi\rangle$ onto the basis $\{|Y_{lm}\rangle\}$ is obtained by the scalar products of the shape with the vectors of the basis:

$$c_l^m = \langle Y_{lm}|\Psi\rangle = \int_0^{2\pi} d\phi \int_0^{\pi} sin\theta \; d\theta \; Y_{lm}^*(\theta,\phi) \; \Psi(\theta,\phi) \tag{4}$$

We compute these coordinates for $l = 0, 1, 2, ..., L$ and $m = -l, ..., l-1, l$, where $L$ is a constant. Increasing the value of $L$ provides a more precise description of the shape, but also increases computation time. Our experiments (section 7) have shown that a small value of $L$ is sufficient in practice. Since, the object is represented by a discrete array, the components corresponding to high values of $l$ represent mainly sampling noise.

Clearly, it is possible to reconstruct the shape if we know $c_l^m$:

$$|\Psi\rangle = \sum_{l,m} c_l^m |Y_{lm}\rangle \tag{5}$$

We will now show that the matrix representation of the rotational operator is very simple in the basis of spherical harmonics. Let us denote:

- $L_+ = L_x + iL_y$
- $L_- = L_x - iL_y$
- $u_+ = \frac{u_x + iu_y}{2}$
- $u_- = \frac{u_x - iu_y}{2}$

Using equation (1) and Taylor expansion, the action of a rotational operator on a spherical harmonic is:

$$\begin{aligned} R|Y_{lm}\rangle &= e^{-i\alpha(u_-L_+ + u_+L_- + u_zL_z)}|Y_{lm}\rangle \\ &= |Y_{lm}\rangle \\ &\quad -i\alpha(u_-L_+ + u_+L_- + u_zL_z)|Y_{lm}\rangle \\ &\quad -\frac{\alpha^2}{2!}(u_-L_+ + u_+L_- + u_zL_z)^2|Y_{lm}\rangle \\ &\quad -... \end{aligned} \tag{6}$$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



But we have [4] [14]:

- $\mathbf{L}^2|Y_{lm}\rangle = l(l+1)|Y_{lm}\rangle$
- $L_z|Y_{lm}\rangle = m|Y_{lm}\rangle$
- $L_+|Y_{lm}\rangle = \sqrt{l(l+1) - m(m+1)}|Y_{l,m+1}\rangle$
- $L_-|Y_{lm}\rangle = \sqrt{l(l+1) - m(m-1)}|Y_{l,m-1}\rangle$

$R|Y_{lm}\rangle$ is therefore a linear combination of $\{|Y_{l,-l}\rangle, |Y_{l,-l+1}\rangle, ..., |Y_{ll}\rangle\}$. Hence, $\mathcal{F}_\mathcal{S}$ can be decomposed as a direct sum of orthogonal subspaces, each of which is globally invariant to rotation:

- $\mathcal{E}_0$, a basis of which is $\{|Y_{00}\rangle\}$
- $\mathcal{E}_1$, a basis of which is $\{|Y_{1,-1}\rangle, |Y_{10}\rangle, |Y_{11}\rangle\}$
- $\mathcal{E}_2$, a basis of which is $\{|Y_{2,-2}\rangle, |Y_{2,-1}\rangle, |Y_{20}\rangle, |Y_{21}\rangle, |Y_{22}\rangle\}$
- etc

If a vector belongs to $\mathcal{E}_l$, its image resulting from any rotation will also belong to $\mathcal{E}_l$. Then, in the basis of spherical harmonics, the rotational operator takes the special form below:

$$\begin{pmatrix} \tilde{c}_0^0 \\ \tilde{c}_1^{-1} \\ \tilde{c}_1^0 \\ \tilde{c}_1^1 \\ \tilde{c}_2^{-2} \\ \tilde{c}_2^{-1} \\ \tilde{c}_2^0 \\ \tilde{c}_2^1 \\ \tilde{c}_2^2 \\ . \\ . \\ . \end{pmatrix} = \begin{pmatrix} D_0 & & & & & \\ & \begin{pmatrix} & D_1 & \end{pmatrix} & & & & \\ & & \begin{pmatrix} & & \\ & D_2 & \\ & & \end{pmatrix} & & & \\ & & & . & & \\ & & & & . & \\ & & & & & . \end{pmatrix} \begin{pmatrix} c_0^0 \\ c_1^{-1} \\ c_1^0 \\ c_1^1 \\ c_2^{-2} \\ c_2^{-1} \\ c_2^0 \\ c_2^1 \\ c_2^2 \\ . \\ . \\ . \end{pmatrix} \tag{7}$$

Since they form a basis of $\mathcal{F}_\mathcal{S}$, the spherical harmonics are able to represent any closed 3D shape. Moreover, the size of the set of coefficients can be chosen so that

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



the spherical harmonics characterize the desired details of the shape.

It is worth noting that spherical harmonics have been used in a recent paper by Wen Chen and Huang [15] to model the time-varying deformations of the left ventricle of the heart. The 3D coordinates of coronary artery bifurcation points are obtained by cineangiography. Then, the coefficients of the spherical harmonic interpolation are computed by an optimization algorithm. According to [15], an order of approximation $L = 3$ was enough to provide satisfactory modelling of the heart deformations. The method we propose here could be used for automatic recognition of pathological heart deformations.

# 4 Computation of invariants

## 4.1 Tensor theory

Some tensor definitions and properties are presented here. A more complete introduction to tensor theory may be found in [13] or [10]. A short introduction to tensor theory, and an example of application to planar object orientation determination may also be found in [3].

Let us consider a vector space $\mathcal{V}$ and a basis $\{e_i\}$. Any vector $x$ of $\mathcal{V}$ may be represented as

$$x = \sum_i x^i e_i = x^i e_i \tag{8}$$

where, in the rightmost expression, the Einstein summation convention has been assumed (implicit summation over any repeated index which is simultaneously in an upper and a lower position).

A new basis $\{\tilde{e}_i\}$ is related to the original basis by:

$$\tilde{e}_i = \alpha_i^j e_j \quad \text{or} \quad e_i = \beta_i^j \tilde{e}_j \tag{9}$$

where $\alpha_i^j$ denotes a linear transformation, and $\beta_i^j$ its inverse. Any form $A_{l_1 \ldots l_p}^{m_1 \ldots m_q}$ is said to be a tensor of covariant rank $p$ and contravariant rank $q$ if it transforms itself according to the equation:

$$\tilde{A}_{k_1 \ldots k_p}^{n_1 \ldots n_q} = \alpha_{k_1}^{l_1} \ldots \alpha_{k_p}^{l_p} \beta_{m_1}^{n_1} \ldots \beta_{m_q}^{n_q} A_{l_1 \ldots l_p}^{m_1 \ldots m_q} \tag{10}$$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



The outer product of two tensors is defined as the product of elements:

$$\mathcal{A}_{lm}^{ijk} = \eta_l^{ij} \mu_m^k \tag{11}$$

The inner product (or tensor contraction) is defined as the operation of pairing covariant and contravariant indices from two tensors, and summing individually over each such pair. For example, the operations

$$\mathcal{A}_m^{ij} = \eta_n^{ij} \mu_m^n \tag{12}$$

and

$$\mathcal{A}^i = \eta_n^{ij} \mu_j^n \tag{13}$$

preserve tensor properties.

A vector is a tensor of order 1. It can be represented either by its covariant or its contravariant components:

$$a_i = \langle a | e_i \rangle \tag{14}$$

$$a = a^i e_i \tag{15}$$

Hence,

$$\begin{aligned} a_i &= \langle a^j e_j | e_i \rangle \\ &= (a^j)^* \langle e_j | e_i \rangle \end{aligned} \tag{16}$$

If the basis is orthonormal, then:

$$a_i = (a^i)^* \tag{17}$$

## 4.2 Construction of tensors

Let us consider a vector $c_l$ of $\mathcal{E}_l$. We have $c_l = c_l^m Y_{lm}$. Therefore, according to equation (15), $c_l^m$ is a contravariant tensor of order 1. It must be pointed out that only $m$ is a tensorial index, while $l$ is just a parameter indexing the subspace $\mathcal{E}_l$ in which the tensor is defined. Since $\{Y_{lm}, -l \leq m \leq l\}$ is an orthonormal basis of $\mathcal{E}_l$, the covariant components of $c_l$ are $c_{lm} = (c_l^m)^*$.

The theorem below allows constructing new tensors:

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



Theorem [4][5]:
*If $T_l^m$ is a tensor of $\mathcal{E}_l$, then:*

$$\Pi(l_1,l_2)_l^m = \sum_{m_1,m_2} \langle l_1\ m_1\ l_2\ m_2|l\ m\rangle\, T_{l_1}^{m_1} T_{l_2}^{m_2} \tag{18}$$

*is also a tensor of $\mathcal{E}_l$.*

The coefficients $\langle l_1\ m_1\ l_2\ m_2|l\ m\rangle$ are Clebsch-Gordan coefficients. They are coupling coefficients between the angular momenta of two particles. Clebsch-Gordan coefficients can be computed using the recurrent equations given in appendix 2. Thanks to this result, we can construct new tensors. In what follows, we will use:

$$\Pi(l_1,l_2)_l^m = \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} \langle l_1\ m_1\ l_2\ m_2|l\ m\rangle c_{l_1}^{m_1} c_{l_2}^{m_2} \tag{19}$$

$$= \sum_{m_1=-l_1}^{l_1} \langle l_1,\ m_1,\ l_2,\ m-m_1|l,\ m\rangle c_{l_1}^{m_1} c_{l_2}^{m-m_1} \tag{20}$$

Some properties of Clebsch-Gordan coefficients are given in appendix 2. To transform (19) in (20), we used property (C1).

## 4.3 Using tensor contraction to derive invariants

By multiplying a contravariant tensor of order 1 with a covariant tensor of order 1 and equating the indices, we obtain a tensor of order 0 (an invariant). Since we have shown (17) that the variance of a tensor can be modified by conjugation of its components (because $\{|Y_{lm}\rangle\}$ is orthonormal), we can construct the following invariants:

$$N(l) = c_l^m c_{lm} = \sum_{m=-l}^{l} c_l^m [c_l^m]^* \tag{21}$$

$$P(l,l_1,l_2) = \Pi(l_1,l_2)_l^m c_{lm} = \sum_{m=-l}^{l} \Pi(l_1,l_2)_l^m [c_l^m]^* \tag{22}$$

$$Q(l,l_1,l_2,l_3,l_4) = \Pi(l_1,l_2)_l^m \Pi(l_3,l_4)_{lm} = \sum_{m=-l}^{l} \Pi(l_1,l_2)_l^m [\Pi(l_3,l_4)_l^m]^* \tag{23}$$

The invariants $N(l)$ are computed for $l = 0,1,...,L$. Many symmetry relations between the other invariants allow to restrict the computation to a few index values,

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



as demonstrated in the next section.

Obviously, the process we have followed to construct these invariants could be continued as far as desired. For instance, we could define another tensor:

$$T(l_1,l_2,l_3,l_4,l_5,l_6)_l^m = \sum_{m_1,m_2} \langle l_1 m_1 l_2 m_2 | l m \rangle \Pi(l_3,l_4)_{l_1}^{m_1} \Pi(l_5,l_6)_{l_2}^{m_2} \tag{24}$$

and use $T$ to construct other invariants. However, we already have a large number of invariants, so here we will restrict the study of properties to invariants $N$, $P$, and $Q$.

# 5 Symmetry relations between invariants

Many properties of the invariants are established. The proofs which are most difficult to derive are given in appendix 3. Some properties are valid for objects described by real valued functions only and are identified by the ($\mathcal{R}$) symbol.

## Results concerning the tensors

**P1:** if $\Psi(\theta,\phi)$ is a real function, then $c_l^m = (-1)^m (c_l^{-m})^*$

**P2:** *Permutation*
$\Pi(l_2,l_1)_l^m = (-1)^{l_1+l_2+l} \Pi(l_1,l_2)_l^m$

**P3** ($\mathcal{R}$)**:** *Conjugation*
$[\Pi(l_1,l_2)_l^m]^* = (-1)^{l_1+l_2+l} (-1)^m \Pi(l_1,l_2)_l^{-m}$

## Results concerning invariants $N$

**P4:** All $N(l)$ are real and positive

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



## Results concerning invariants $P$

**P5:** *Permutation*

a) $P(l, l_2, l_1) = (-1)^{l_1+l_2+l} P(l, l_1, l_2)$

b) ($\mathcal{R}$) $P(l_2, l, l_1) = (-1)^{l+l_1} \sqrt{\frac{2l+1}{2l_1+1}} P(l, l_1, l_2)$

Consequence: if $\{l', l'_1, l'_2\}$ is a permutation of $\{l, l_1, l_2\}$, then $P(l', l'_1, l'_2)$ is fully defined by the knowledge of $P(l, l_1, l_2)$. Hence, the computations can be reduced to $l_2 \leq l_1 \leq l$.

**P6** ($\mathcal{R}$)**:** *Conjugation*

$P(l, l_1, l_2)^* = (-1)^{l_1+l_2+l} P(l, l_1, l_2)$

**P7** ($\mathcal{R}$)**:**

a) If $l_1 + l_2 + l$ is odd, then $P(l_1, l_2, l)$ is pure imaginary

b) If $l_1 + l_2 + l$ is even, then $P(l_1, l_2, l)$ is real

**P8:** *Special cases*

a) ($\mathcal{R}$) $P(l, l_1, l_1)$ is real

b) If $l$ is odd, then $P(l, l_1, l_1) = 0$

## Results concerning invariants $Q$

**P9:** *Permutation*

a) $Q(l, l_2, l_1, l_3, l_4) = (-1)^{l_1+l_2+l} Q(l, l_1, l_2, l_3, l_4)$

b) $Q(l, l_1, l_2, l_4, l_3) = (-1)^{l_3+l_4+l} Q(l, l_1, l_2, l_3, l_4)$

c) ($\mathcal{R}$) $Q(l, l_3, l_4, l_1, l_2) = (-1)^{l_1+l_2+l_3+l_4} Q(l, l_1, l_2, l_3, l_4)$

**P10** ($\mathcal{R}$)**:** *Conjugation*

$Q(l, l_1, l_2, l_3, l_4)^* = (-1)^{l_1+l_2+l_3+l_4} Q(l, l_1, l_2, l_3, l_4)$

**P11** ($\mathcal{R}$)**:**

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



a) If $l_1 + l_2 + l_3 + l_4$ is odd, then $Q(l, l_1, l_2, l_3, l_4)$ is pure imaginary
b) If $l_1 + l_2 + l_3 + l_4$ is even, then $Q(l, l_1, l_2, l_3, l_4)$ is real

**P12:** *Special cases*

a)If $l$ is odd, then:

- $Q(l, l_1, l_1, l_3, l_4) = 0$
- $Q(l, l_1, l_2, l_3, l_3) = 0$

b) $(\mathcal{R})$ $Q(0, l_1, l_1, l_3, l_3) = \dfrac{(-1)^{l_1+l_3}}{\sqrt{(2l_1+1)(2l_3+1)}} N(l_1)N(l_3)$

c) $Q(l, l_1, l_2, 0, l) = c_{00}P(l, l_1, l_2)$
(Hence, we can impose $l, l_1, l_2, l_3, l_4 \geq 1$)

## Reduced invariants

Due to the properties above, the computation of the invariants can be restricted to some values of the indexes.

The $N(l)$ are computed for $l = 0, 1, ..., L$.

The $P(l, l_1, l_2)$ are computed for the indexes that verify simultaneously the conditions below:

$$\begin{array}{l} 1 \leq l_2 \leq l_1 \leq l \leq L \\ l_1 - l_2 \leq l \leq l_1 + l_2 \\ l_2 \neq l_1 \quad \text{or} \quad l \quad \text{even} \\ l_1 \neq l \quad \text{or} \quad l_2 \quad \text{even} \end{array}$$

The $Q(l, l_1, l_2, l_3, l_4)$ are computed for indexes that verify simultaneously the conditions below:

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



$$
\begin{array}{l}
0 \leq l_2 \leq l_1 \leq L \\
0 \leq l_4 \leq l_3 \leq l_1 \\
1 \leq l \leq L \\
l_1 - l_2 \leq l \leq l_1 + l_2 \\
l_3 - l_4 \leq l \leq l_3 + l_4 \\
l_2 \neq l_1 \quad \text{or} \quad l \quad \text{even} \\
l_4 \neq l_3 \quad \text{or} \quad l \quad \text{even}
\end{array}
$$

Then we define a set of invariants, which we call "reduced invariants":

$$
\begin{aligned}
n(l) &= [\Re\{N(l)\}]^{\frac{1}{2}} \\
p(l,l_1,l_2) &= H_3(\Re\{P(l,l_1,l_2)\}) \text{ if } l+l_1+l_2 \text{ is even} \\
&= H_3(\Im\{P(l,l_1,l_2)\}) \text{ if } l+l_1+l_2 \text{ is odd} \\
q(l,l_1,l_2,l_3,l_4) &= H_4(\Re\{Q(l,l_1,l_2,l_3,l_4)\}) \text{ if } l_1+l_2+l_3+l_4 \text{ is even} \\
&= H_4(\Im\{Q(l,l_1,l_2,l_3,l_4)\}) \text{ if } l_1+l_2+l_3+l_4 \text{ is odd}
\end{aligned}
$$

where $H_n(a) = sign(a).|a|^{\frac{1}{n}}$, and $\Re$ and $\Im$ stand for the real part and the imaginary part, respectively.

For classification, the $n, p, q$ coefficients are provided to the input of a classifier. The classifier could well be a k-nearest-neighbours, a Bayesian classifier, or a neural network.

# 6 Extension to full 3D description

## 6.1 Principle

Until now, the object has been described by a function $\Psi(\theta, \phi)$. However, for objects of complex shapes, this description may be ambiguous (i.e. two different objects may have the same description). The objective of this section is to extend the method to objects described by a function $\Psi(x, y, z)$ (or equivalently $\Psi(r, \theta, \phi)$ in spherical coordinates).

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



To follow the same approach as before, we first have to find an orthonormal basis of $\mathcal{F}$ in which the rotational operator has an irreducible form. As before, we will try to find a basis of eigenvectors common to $\mathbf{L}^2$ and $L_Z$. The eigen-equations are:

$$\mathbf{L}^2|\Omega_{klm}\rangle = l(l+1)|\Omega_{klm}\rangle \quad (25)$$
$$L_Z|\Omega_{klm}\rangle = m|\Omega_{klm}\rangle \quad (26)$$

where $k$ is an index we have added to take into account possible degeneracy. Since $\mathbf{L}^2$ and $L_Z$ do not depend on $r$, $k$ can be considered a parameter. Hence, the general form of the solution is:

$$\Omega_{klm}(r,\theta,\phi) = R_{kl}(r)Y_{lm}(\theta,\phi) \quad (27)$$

where $Y_{lm}(\theta,\phi)$ is a spherical harmonic, and $R_{kl}(r)$ is the integration constant (it can be proved [4][14] that $R_{kl}(r)$ does not depend on $m$).

The decomposition of a shape $|\Psi\rangle$ onto the basis $\{|\Omega_{klm}\rangle\}$ is obtained from the scalar products of the shape with the vectors of the basis:

$$c_{kl}^m = \langle\Omega_{klm}|\Psi\rangle = \int_0^\infty r^2 dr \int_0^{2\pi} d\phi \int_0^\pi sin\theta \; d\theta \; \Omega_{klm}^*(r,\theta,\phi) \; \Psi(r,\theta,\phi) \quad (28)$$

It is possible to reconstruct the shape if we know the $c_{kl}^m$:

$$|\Psi\rangle = \sum_{k,l,m} c_{kl}^m |\Omega_{klm}\rangle \quad (29)$$

Once the $c_{kl}^m$ have been computed, invariants are constructed using the tensor-based method described in previous sections. One has only to replace each $l_i$ by $(k_i, l_i)$ in the equations.

There are many solutions for the functions $R_{kl}(r)$. The only condition is that $\{|\Omega_{klm}\rangle\}$ be an orthonormal basis of $\mathcal{F}$. The more natural solution would be to add an operator to avoid the degeneracy, as may be done by a field in quantum mechanics. This approach is described in appendix 4. However, we prefer to look for a basis composed of harmonic functions, since with a harmonic basis, determining an upper bound for $k$ is easy (it only requires determining up to which frequency we want to analyse the shape).

To obtain a discrete spectrum, the transformation domain must be limited in

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



space (the spectrum should be discrete because, otherwise, $k$ would be a continuous index). Since any real object is finite, it can always be enclosed in a sphere of radius 1 by a suitable choice of the unit of length. Let $\mathcal{F}_\mathcal{V}$ be the subspace of $\mathcal{F}$ composed of functions from $\mathcal{V}$ to $\mathcal{C}$, where $\mathcal{V}$ is the volume of the unit sphere of $\mathcal{R}^3$. We are looking for an orthonormal basis of $\mathcal{F}_\mathcal{V}$. Let us check that:

$$R_k(r) = \sqrt{2}\frac{\sin(\pi k r)}{r} \tag{30}$$

where $k \geq 1$, is a solution. We have to check that this is an orthonormal system, and that it generates $\mathcal{F}_\mathcal{V}$.

Orthonormality:

$$\begin{aligned}
\langle \Omega_{klm} | \Omega_{k'l'm'} \rangle &= \int_0^1 r^2 dr \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta R_k^*(r) Y_{lm}^*(\theta,\phi) R_{k'}(r) Y_{l'm'}(\theta,\phi) \\
&= \langle Y_{lm} | Y_{l'm'} \rangle \int_0^1 r^2 dr R_k(r) R_{k'}(r) \\
&= \delta_{kk'}\delta_{ll'}\delta_{mm'}
\end{aligned} \tag{31}$$

because

$$\begin{aligned}
\int_0^1 r^2 R_k(r) R_{k'}(r) dr &= 2\int_0^1 \sin(\pi k r)\sin(\pi k' r) dr \\
&= \int_0^1 \{\cos(\pi(k-k')r) - \cos(\pi(k+k')r)\} dr \\
&= \delta_{kk'}
\end{aligned} \tag{32}$$

Generation of $\mathcal{F}_\mathcal{V}$:

We must check that any vector of $\mathcal{F}_\mathcal{V}$ can be decomposed onto this basis. We can write:

$$\Psi(r,\theta,\phi) = \sum_{l,m} a_{lm}(r) Y_{lm}(\theta,\phi) \tag{33}$$

with:

$$a_{lm}(r) = \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta Y_{lm}^*(\theta,\phi)\Psi(r,\theta,\phi) \tag{34}$$

because $\{|Y_{lm}\rangle\}$ is an orthonormal basis of $\mathcal{F}_\mathcal{S}$. Let us define a function $\tilde{a}(r)$:

$$\tilde{a}_{lm}(r) = \tilde{a}_{lm}(r-2) \tag{35}$$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



$$\tilde{a}_{lm}(-r) = -\tilde{a}_{lm}(r) \quad (36)$$

$$\tilde{a}_{lm}(r) = r\, a_{lm}(r) \text{ for } 0 \leq r \leq 1 \quad (37)$$

$\tilde{a}(r)$ is an odd function of period 2. Hence, its decomposition in Fourier series is:

$$\tilde{a}_{lm}(r) = \sum_{k=1}^{\infty} b_{klm} \sin(\pi k r) \quad (38)$$

where the $b_{klm}$ are unique. Let $c_{kl}^m = \frac{b_{klm}}{\sqrt{2}}$. Then:

$$a_{lm}(r) = \sum_{k=1}^{\infty} c_{kl}^m R_k(r) \quad (39)$$

where the $c_{kl}^m$ exist and are unique. Therefore,

$$|\Psi\rangle = \sum_{k=1}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{kl}^m |\Omega_{klm}\rangle \quad (40)$$

## 6.2 Determination of the decomposition level

In practice, the coefficients $c_{kl}^m$ are calculated for $k = 1, 2, ..., K$. In this section, we want to determine a suitable value of $K$. At first, let us consider objects described by a binary function $\Psi(r, \theta, \phi)$ (1 inside the object and 0 outside). Let us consider that, in any direction $(\theta, \phi)$, there are never more than $P$ intersections with the surface and let us denote the corresponding distances as $(r_1, r_2, \ldots, r_P)$. Figure 4 shows a function $a_{\theta,\phi}(r) = \Psi(r, \theta, \phi)$ for the object of figure 3.

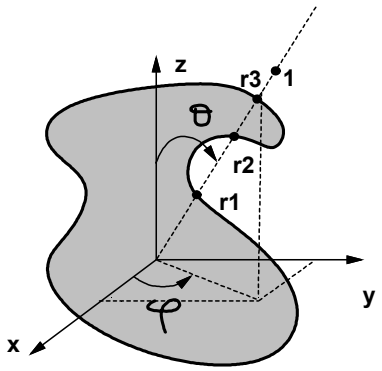


Figure 3: Example of object: P=3

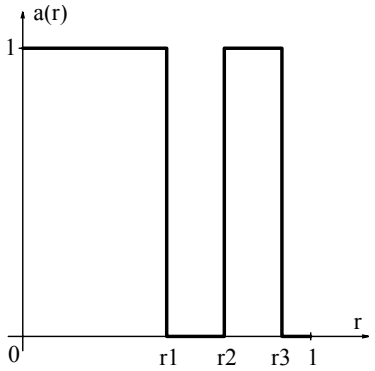


Figure 4: Example of function $a_{\theta,\phi}(r)$

Using (27) and (29) we can write:

$$\Psi(r,\theta,\phi) = \sum_{k=1}^{\infty} R_k(r) c_k(\theta,\phi) \tag{41}$$

where

$$c_k(\theta,\phi) = \sum_{l,m} c_{kl}^m Y_{lm}(\theta,\phi) \tag{42}$$

Then, using (41) and (32), we have:

$$c_k(\theta,\phi) = \int_0^1 r^2 dr R_k(r) \Psi(r,\theta,\phi) \tag{43}$$

$$= \sqrt{2} \int_0^1 r sin(\pi k r) \Psi(r,\theta,\phi) dr \tag{44}$$

Since $\Psi(r,\theta,\phi)$ is a binary function, equation (44) can be written:

$$c_k(\theta,\phi) = \sum_{p=1}^{P} (-1)^{p+1} F_k(r_p) \tag{45}$$

where

$$\frac{1}{\sqrt{2}} F_k(r) = \int r sin(\pi k r) dr = \frac{1}{(\pi k)^2} sin(\pi k r) - \frac{r}{(\pi k)} cos(\pi k r) \tag{46}$$

The coefficients $c_{kl}^m$ are computed for $k = 1, ..., K$, then, using (42), $c_k(\theta,\phi)$ is known for $k = 1, ..., K$. Hence (45) provides $K$ equations, the unknowns of which are $(r_1, r_2, \ldots, r_P)$. Therefore, using the a priori knowledge that the object is originally described by a binary function, the shape is determined without ambiguity when $K \geq P$. For most natural objects, except those with very complex shapes, we have $P \leq 3$. We can conclude that $K = 3$ is usually enough to represent the shape without losing information. Of course, in applications were objects have very complex shapes,

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



the value of $K$ might be increased slightly.

When the object is originally represented by a non-binary function (for example, it may be a function which depends on the local density of water, such as in nuclear magnetic resonance), $K$ should be higher if we want $c_{kl}^m$ to represent the object without losing information. Figure 5 shows the reconstruction of $a_{\theta,\phi}(r)$ for some values of $K$. However, it must be stressed that a lossless representation of the shape is not always necessary to discriminate between shapes.

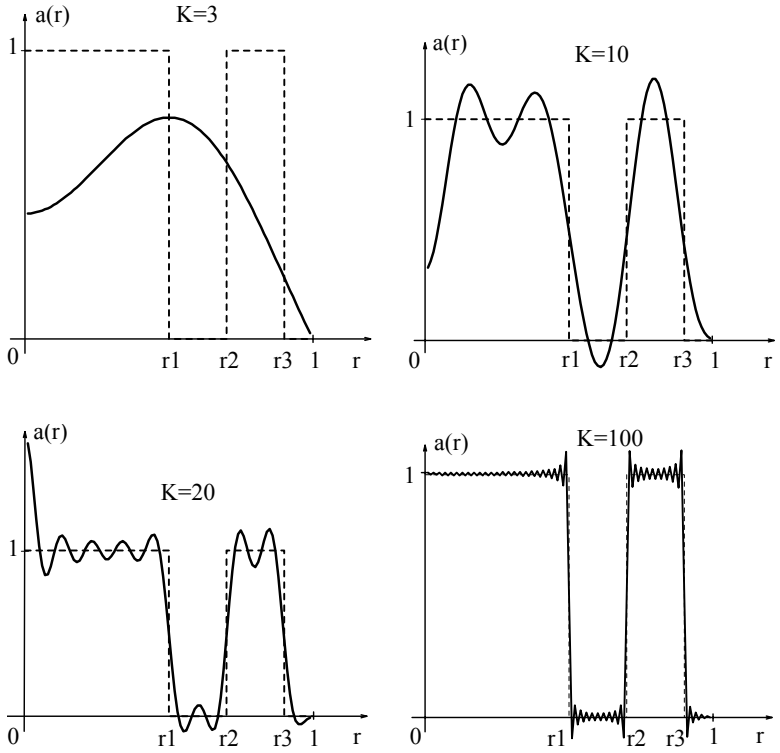


Figure 5: Reconstruction of $a_{\theta,\phi}(r)$ for some values of $K$

## 6.3 Link with moments

In this paragraph, we show that 3D moment invariants [9] [12] can be derived from a degenerate case of our approach. Let us consider the normalized set of functions:

$$\Omega_{klm}(r,\theta,\phi) = \frac{1}{\sqrt{2k+3}} r^k Y_{lm}(\theta,\phi) \tag{47}$$

where $k = l + 2n$ and $n \geq 0$. The interest of the constraint on $k$ will appear

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



later: without it, we could not obtain the moments. This set of functions is not orthogonal, nor is it a basis. However, the subset corresponding to a given $k$ is orthogonal. Another interesting orthogonal subset is the subset of functions with $k = l$. In quantum mechanics, these functions are known as l-pole moments, and they are usually denoted as $Q_{lm}(r, \theta, \phi)$. The quadrupole moments $Q_{2m}$ are often used to study systems of particles. The set of l-pole moments is not a basis of $\mathcal{F}$: it generates only the subspace of functions with null Laplacian. This can be easily proved, since the Laplacian of $Q_{lm}$ is given by:

$$\begin{aligned}\sqrt{2l+3}\,\Delta Q_{lm}(r,\theta,\phi) &= \left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\mathbf{L}^2}{r^2}\right)\left(r^l Y_{lm}(\theta,\phi)\right)\\ &= \frac{1}{r}\left(l(l+1)r^{l-1}Y_{lm}(\theta,\phi) - l(l+1)r^{l-1}Y_{lm}(\theta,\phi)\right)\\ &= 0\end{aligned}$$

In appendix 1, we show that the spherical harmonic $Y_{lm}$ is a linear sum of terms in $sin^m\theta\, cos^a\theta\, sin^{m-b}\phi\, cos^b\phi$, where $a \in \{l-m, l-m-2, ...\}$, $a \geq 0$, and $0 \leq b \leq m$. Due to the relation between $(x, y, z)$ and $(r, \theta, \phi)$, given in section 2, we obtain:

$$sin^m\theta\, cos^a\theta\, sin^{m-b}\phi\, cos^b\phi = x^b y^{m-b} z^a / r^{m+a} \tag{48}$$

Hence, the scalar product $c^m_{kl} = \langle \Omega_{klm} | \Psi \rangle$ is a linear sum of terms in:

$$\int\int\int r^{k-m-a} x^b y^{m-b} z^a \Psi(x, y, z) dx dy dz \tag{49}$$

Since $a = l - m - 2q$ and $k = l + 2n$, $r^{k-m-a} = r^{2(q+n)}$ can be replaced by $(x^2 + y^2 + z^2)^{q+n}$. The development of the integral (49) provides moments of order $k$ only (because the sum of the exponents is [2q + 2n] + b + [m-b] + [l-m-2q] = l + 2n = k). Finally, it appears that $c^m_{kl}$ is a linear sum of moments of order $k$. Then, we can obtain moment invariants by using the tensor-based method described in section 4 (for a given $k$, the functions (47) are orthonormal, and their angular dependence is given by the spherical harmonics, hence the tensor properties used to derive invariants still apply).

As an example, let us show that $N(2, 2)$ is a function of two moment invariants of Sadjadi. Using the explicit expressions of the $l = 2$ spherical harmonics given in appendix 1, one can check easily that for $l = 2$, we have the following quadrupole moments:

$$Q_{2,\pm 2}(x, y, z) = \sqrt{\frac{3}{2}}\gamma(x \pm iy)^2 \tag{50}$$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



$$Q_{2,\pm1}(x,y,z) = \mp\sqrt{6}\gamma z(x \pm iy) \tag{51}$$

$$Q_{2,0}(x,y,z) = \gamma(2z^2 - x^2 - y^2) \tag{52}$$

where $\gamma = \sqrt{\frac{5}{112\pi}}$. Then, since $c_{ll}^m = \int\int\int Q_{lm}^*(x,y,z)\Psi(x,y,z)dxdydz$, we have:

$$c_{22}^{\pm2} = \sqrt{\frac{3}{2}}\gamma(M_{x^2} - M_{y^2} \mp 2iM_{xy}) \tag{53}$$

$$c_{22}^{\pm1} = \mp\sqrt{6}\gamma(M_{xz} \mp iM_{yz}) \tag{54}$$

$$c_{22}^{0} = \gamma(2M_{z^2} - M_{x^2} - M_{y^2}) \tag{55}$$

where $M_{x^2}, M_{y^2}, ...$ are the second order moments. Finally, the invariant $N(2,2)$ is:

$$N(2,2) = \sum_{m=-2}^{2} |c_{22}^m|^2 \tag{56}$$

$$= 4\gamma^2(J_1^2 - 3J_2) \tag{57}$$

where

$$J_1 = M_{x^2} + M_{y^2} + M_{z^2} \tag{58}$$

$$J_2 = M_{x^2}M_{y^2} + M_{x^2}M_{z^2} + M_{y^2}M_{z^2} - M_{xz}^2 - M_{yz}^2 - M_{xy}^2 \tag{59}$$

In fact, $J_1$ and $J_2$ are the moment invariants of Sadjadi ([12], p. 134). $N(2,2)$ is also proportional to the moment invariant $\nu_{2,2,0,0}$ proposed in [9]. Similarly, we can show that $N(2,0) = \frac{1}{28\pi}J_1^2$.

Hence, the moment invariants can be derived using our approach just by choosing a special set of functions for the decomposition of the object. Furthermore, this approach is more simple than previous techniques proposed to derive 3D moment invariants, in the sense that here a generic set of equations is derived, unlike previous methods where programming a specific formula for each moment invariant was required. However, when only a small number of features is needed, the approaches proposed by Sadjadi [12] and Lo [9] are easier to implement.

It should be pointed out that, unlike our standard 3D invariants, moment invariants use partial shape information only, since the set (47) is not a basis of $\mathcal{F}$. Hence, it is clear that many different shapes can produce exactly the same moment invariants (which occurs when the shapes have the same projections on the subspace generated by the set (47)), in which cases these shapes cannot be discriminated by a moment-based approach.

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



Furthermore, there are more standard 3D invariants than moment invariants because, for moments, the index $k$ is restricted to be equal to $l + 2n$. Hence, to provide a given number of invariants, the moment-based approach requires higher values of $k$ than our approach. This is a drawback of moment-based methods, since, as $k$ increases, the calculations become less reliable. Finally, since the functions which are used to generate the moments are not orthogonal, moment invariants are likely to provide more correlated information than the standard 3D invariants.

# 7 Experimental Results

## 7.1 The spherical harmonics and the decomposition of an object

Figure 6 shows some spherical harmonics $Y_{lm}$ represented as a function $r = |Y_{lm}(\theta, \phi)|$. Figure 7 shows the reconstruction of a vertebra at various resolutions (here, and in all the experiments described below, the objects are represented by vectors of $\mathcal{F}_{\mathcal{S}}$). The original 3D image of the vertebra comes from scanner data. A simple thresholding has been performed to extract the bone. The $c_l^m$ are computed for $0 \leq l \leq L$, after which the shape of the vertebra is reconstructed using equation (5). Increasing the value of $L$ provides a better representation of the shape.

## 7.2 Verification of 3D invariance

Figure 8 shows the values of the invariants for an F14 aircraft (fig. 10). The resolution for decomposition is $L = 9$. For this value of $L$, we have 10 invariants $n$, 80 invariants $p$, and we add 60 invariants $q$. These invariants are indexed from 0 to 149, and their values are represented on the diagrams. The top diagram represents the superposition of the invariants for 3 rotations of the object. This confirms 3D invariance, as the curves are almost perfectly superposed. The residual difference (bottom diagram) is due to sampling.

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



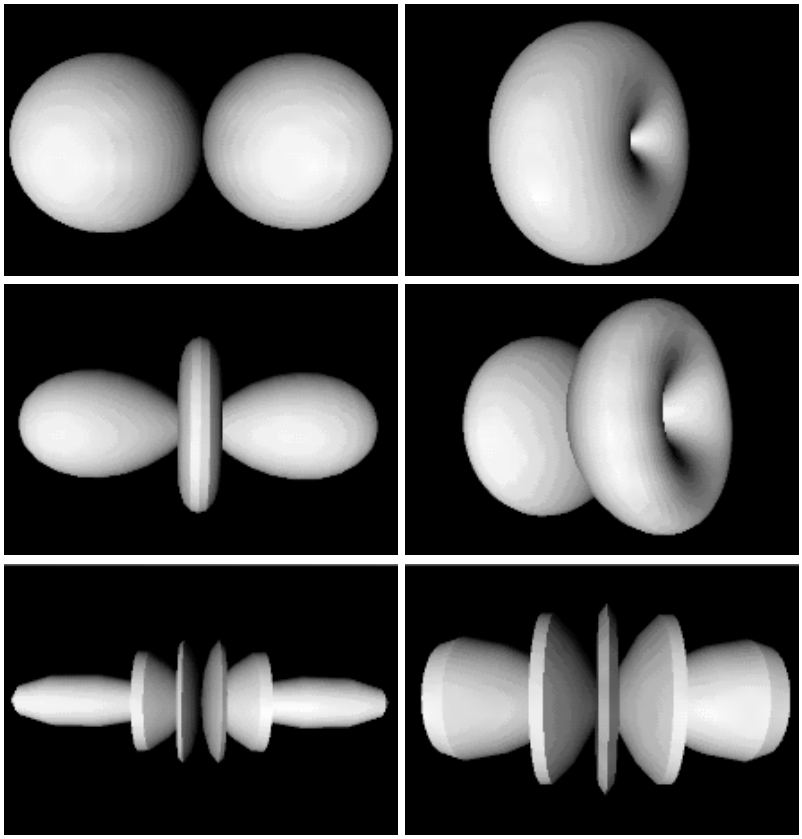

Figure 6: Some spherical harmonics (from left to right and top to bottom: $Y_{10}, Y_{11}, Y_{20}, Y_{21}, Y_{50}, Y_{51}$)

## 7.3 Influence of sampling

Figure 9 shows the influence of the sampling step size on the values of the invariants. In practice, equation (4) is evaluated by discretizing $\theta$ and $\phi$. Angular sampling step sizes of $2^o$, $5^o$ and $20^o$ have been tested. The top diagrams show the values of the invariants for sampling step sizes of $2^o$ and $5^o$, and $2^o$ and $20^o$, and the bottom diagrams show the differences. These diagrams show that the error on the values of the invariants becomes noticeable only for very rough sampling ($20^o$).

## 7.4 Recognition results using a nearest-neighbour classifier

Nearest-neighbour classification consists of comparing the feature vector of the object with feature vectors of models. The comparison is performed with the help of a

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



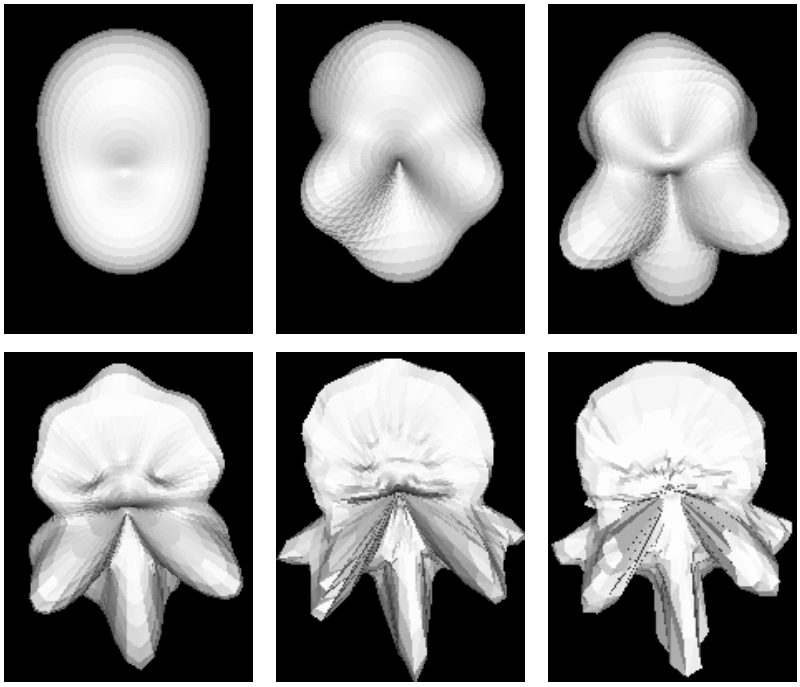

Figure 7: Reconstruction of a vertebra at various resolutions (from left to right and top to bottom: L= 2, 4, 6, 10, 20, and original)

distance measure:

$$d(object, model) = \sqrt{\sum_i (I_{object,i} - I_{model,i})^2} \tag{60}$$

where $I_{object,i}$ stands for the $i^{th}$ invariant of the object and $I_{model,i}$ stands for the $i^{th}$ invariant of a model. The model providing the minimal distance determines the object class.

Let us consider the three models on the right side of figure 10 (F14-1, X29-1, F15-1). The aircraft to be recognized is shown on the left side of the figure. 3D invariants of this aircraft are computed and compared with the invariants of the models. The distances are labelled on the arrows. The minimal distance is obtained when the object is compared with the X29 model. Hence, nearest-neighbour classification would provide the correct result, as the object on the left side is in fact an X29 aircraft.

The table below shows the distances between three objects to recognize and three models. As expected, the minimal distances are always obtained on the diagonal (i.e. when the object is compared with the model of the correct class).

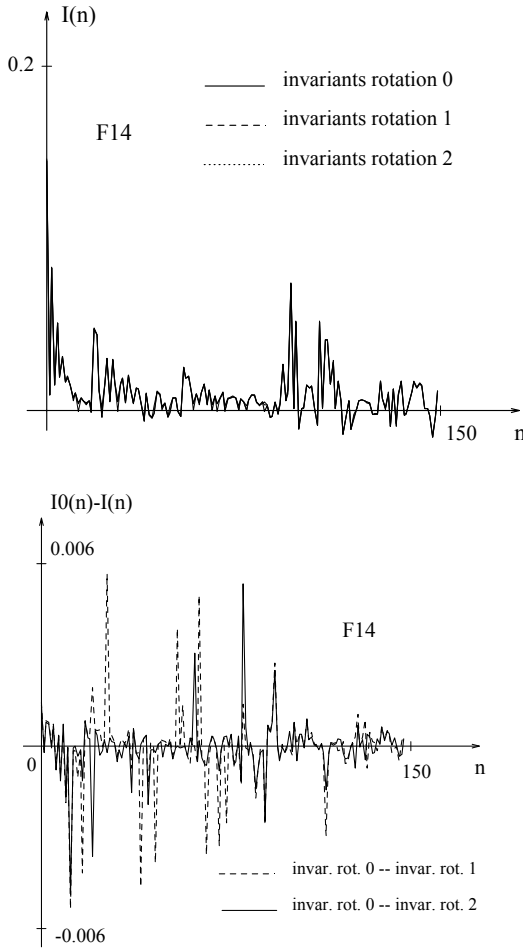


Figure 8: Values of the invariants for various orientations of the F14

| | X29-1 | F14-1 | F15-1 |
|---|---|---|---|
| X29-2 | 2 | 380 | 158 |
| F14-2 | 360 | 3 | 130 |
| F15-2 | 167 | 122 | 3 |

For these experiments, the resolution level was L=9. The invariants used were the 10 invariants $n$, the 80 invariants $p$ plus the 60 first invariants $q$. The results illustrate an interesting feature of the approach: the ability to discriminate between very similar shapes.

Under the same conditions, the method is used to discriminate between two kinds of vertebrae (class A and class B, whose shapes slightly differ: they are at different locations in the spinal column). An example of each class is shown in figure 11.

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



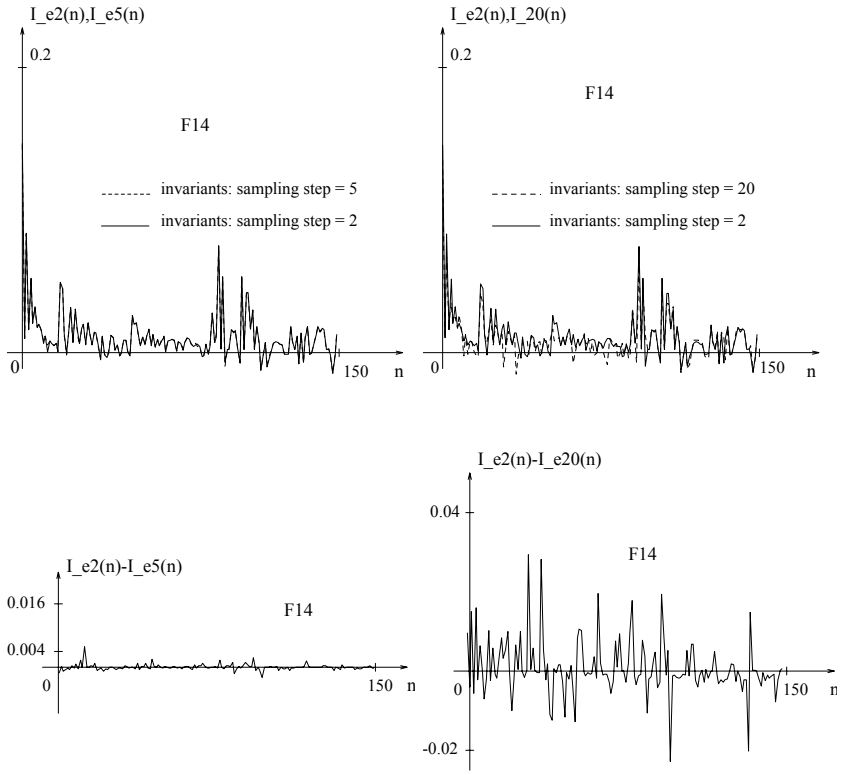


Figure 9: Influence of the sampling step size on the values of the invariants (F14)

The table below shows the average distance between examples of both classes and two models.

| | model-A | model-B |
|---|---|---|
| class-A | 5 | 37 |
| class-B | 38 | 7 |

Figure 12 shows a vertebra of class B obtained with poor tuning of the acquisition process. This has created a hole through the vertebra. To evaluate the robustness of the discrimination, the table below gives the distance between this 3D image and the models of both classes.

| | model-A | model-B |
|---|---|---|
| Poor acquisition, class-B | 48 | 23 |

Hence, the image would still be correctly classified, but the numerical values show that further degradation in the acquisition could hardly be tolerated. While with today scanner equipment it is easy to obtain good data (for the experimentation above, the tuning was intentionally degraded), such degradations could occur in

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



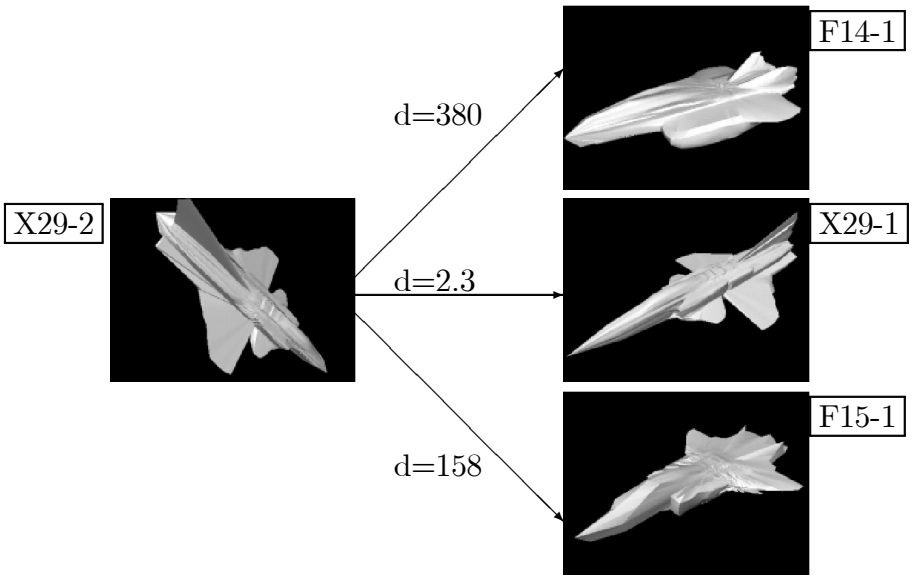


Figure 10: Comparison of an unknown object with three models

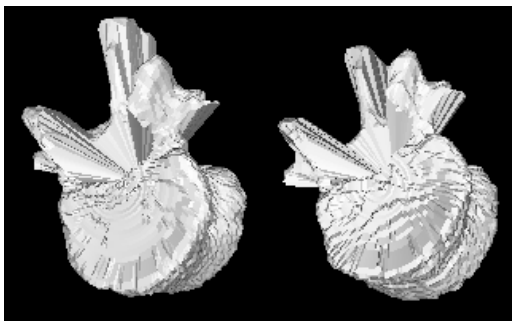

Figure 11: Two classes of vertebrae

other applications (for example, when using a laser range finder), where self-occlusion of the object causes a severe degradation of its 3D description. A solution to this problem could be to improve the classification process by using a learning technique to train a classifier (a neural network with the invariants as inputs, for example) to become robust with respect to self-occlusion. This kind of approach was proposed and evaluated in a previous paper [2] for objects described by 2D Fourier descriptors, and produced good results on real-world data.

Finally, let us consider a problem of discrimination of bones: figure 13 shows some scanner slices of various kinds of bones (radii, ulnae, humerus, hip bone, vertebra). Ten classes were defined: the right (R) and left (L) radius, ulna, humerus

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



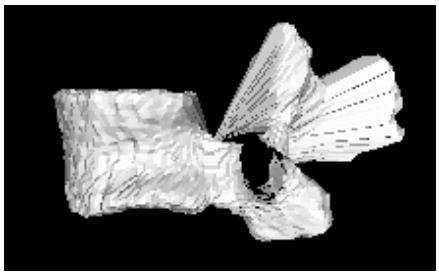

Figure 12: Poorly segmented vertebra of class B

and hip bone, plus the 2 classes of vertebrae (A and B) mentioned above. The invariants were first computed on one example of each class to create the model vector of the class. Then, using a data-base containing 4 other examples of each class, the average distance between the examples of the class and the models was computed. The result is shown in the table below (Only one half is shown, for ease of reading. Anyway, the table is symmetrical as far as 2 significant digits are considered).

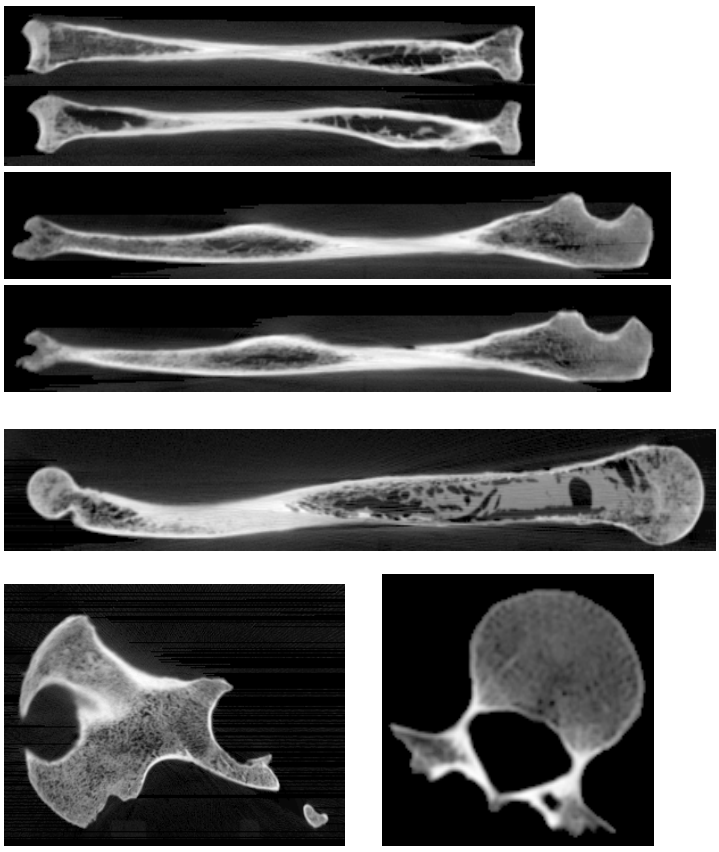

Figure 13: Scanner slices of bones: 2 radii, 2 ulnae, 1 humerus, 1 hip bone, and 1 vertebra

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



| model / class | | radius L | radius R | ulna L | ulna R | humerus L | humerus R | vertebra A | vertebra B | hip bone L | hip bone R |
|---|---|---|---|---|---|---|---|---|---|---|---|
| radius | L | 4 | 14 | 50 | 130 | 150 | 140 | 540 | 490 | 560 | 480 |
| | R | | 3 | 48 | 130 | 150 | 140 | 540 | 490 | 560 | 480 |
| ulna | L | | | 3 | 120 | 150 | 150 | 540 | 490 | 560 | 470 |
| | R | | | | 5 | 230 | 220 | 470 | 420 | 520 | 430 |
| humerus | L | | | | | 4 | 31 | 590 | 540 | 600 | 540 |
| | R | | | | | | 3 | 590 | 540 | 600 | 530 |
| vertebra | A | | | | | | | 5 | 37 | 460 | 500 |
| | B | | | | | | | | 7 | 420 | 450 |
| hip bone | L | | | | | | | | | 4 | 330 |
| | R | | | | | | | | | | 2 |

This table shows that, as expected, the distance is low when the example is compared with the correct model (diagonal). Furthermore, the method is able to discriminate between the left and right versions of the bones (especially, the hip bones and the ulnae produce high distances between their left and right versions). Finally, as expected, long bones (radius, ulna, humerus) are closer to other long bones than to more compact bones (vertebra, hip bone).

### 7.5 Study of redundancies using PCA

Figure 14 shows the eigenvalues obtained by Principal Component Analysis (PCA) on the set composed of the 10 invariants $N$ plus the 80 invariants $P$. The decomposition level was $L = 9$ and a set of 2000 random objects was generated by random choice of the $c_l^m$ values (under the constraint provided by property **P1**). This figure shows that there is no null eigenvalue, so there are no remaining implicit linear equations between the invariants considered here.

## 8 Conclusion

A theoretical framework to derive 3D invariants has been proposed. The approach consists of decomposing 3D shapes onto an orthonormal basis composed of the eigenvectors of the angular momentum operator, deriving contravariant tensors of order 1, and then computing invariants by tensor contraction. Numerous invariants are obtained, and we have established many symmetry properties between the invariants. Experimental results show that these invariants are robust with respect

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



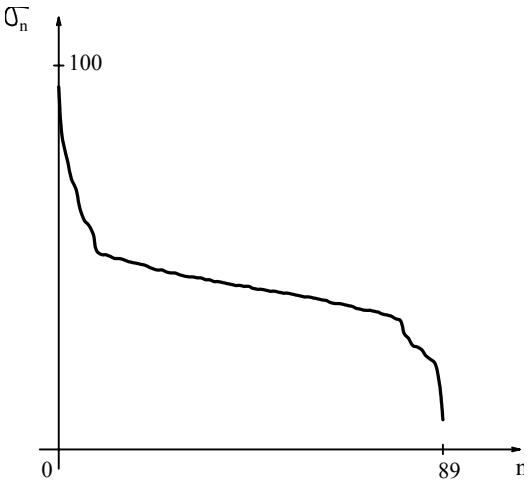


Figure 14: Eigenvalues obtained by PCA (90 invariants)

to sampling noise, and that they can be used to classify 3D shapes.

Such invariants offer an alternative to structural methods for description and recognition of 3D shapes. Unlike 3D moment invariants, they do not require the use of high orders, as the proposed framework allows the derivation of a large number of invariants. Another interesting feature of the approach is the ability to discriminate between very similar shapes. Further work will include the evaluation of the discrimination power on larger data-bases, a study of reversibility (i.e. does any subset of invariants characterize unambiguously the shape of an object at a given resolution?), and application of tensors to 3D object positioning. Some properties of the invariants may also be useful for automatic detection of symmetries of 3D objects.

# Acknowledgments

The 3D images of vertebrae have been provided by the University of Rennes, LTSI Laboratory. The 3D images of radii, ulnae, humerus, and hip bones have been provided by ENST Brest, Image and Information Processing Laboratory.

The authors would like to thank Gordon Wells, Institut de Cibernética, Barcelona, Spain, Nadine Rondel, Thomson Broadband Systems, Rennes, France, and the referees for their valuable comments for improving the quality and contents of this paper.

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



# References

[1] R.C. Bolles, R.A. Cain, "Recognizing and Locating Partially Visible Objects: The Local-Feature-Focus Method", The International Journal of Robotics Research, vol 1, $n°3$, Fall 1982

[2] G. Burel, "Reconnaissance d'objets 3D par réseau d'automates", Proceedings of RFIA conference, Lyon, France, November 25th-29th, 1991

[3] D. Cyganski, J.A. Orr, "Applications of tensor theory to object recognition and orientation determination", IEEE-PAMI, vol 7, $n^o6$, November 1985

[4] A.R. Edmonds, "Angular Momentum in Quantum Mechanics", Princeton, NJ: Princeton University Press, 1974

[5] E. Elbatz, "Mécanique Quantique", Editions Marketing, Paris, 1985

[6] G.H. Granlund, "Fourier processing for hand print character recognition", IEEE Trans. Computers, vol 21, pp 195-201, 1972

[7] C.S. Lin, C.L. Hwang, "New forms of shape invariants from elliptic Fourier Descriptors", Pattern Recognition, vol 20, $n^o5$, pp 535-545, 1987

[8] M.K. Hu, "Visual Pattern Recognition by moment invariants", IEEE Transactions on Information Theory, vol IT-8, pp 179-187, Feb. 1962

[9] C.H. Lo, H.S. Don, "3D moment forms: their construction and application to object identification and positioning", IEEE-PAMI, vol 11, $n^o10$, October 1989

[10] D. Lovelock, H. Rund, "Tensors, Differential Forms, and Variational Principles", New York: Wiley-Interscience, 1975

[11] D.E. Rumelhart, G.E. Hinton, R.J. Williams, "Learning internal representations by error backpropagation", Parallel Distributed Processing, D.E. Rumelhart and J.L. Mc Clelland, Chap 8, Bradford Book - MIT Press - 1986

[12] F.A. Sadjadi, E.L. Hall, "Three-dimensional moment invariants", IEEE-PAMI, vol 2, $n^o$ 2, march 1980

[13] J.L. Synge, A. Schild, "Tensor Calculus", New York: Dover, 1978

[14] C. Cohen-Tanoudji, B. Diu, F. Laloë, "Mécanique Quantique", tomes 1 et 2, Hermann, Paris, 1973

[15] C. Wen Chen, S. Huang, M. Arrott, "Modeling, Analysis, and Visualization of Left Ventricle Shape and Motion by Hierarchical Decomposition", IEEE-PAMI, vol. 16, $n^o$ 4, pp. 342-356, April 1994

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



[16] E. Wigner, "Group Theory and its application to Quantum Mechanics of Atomic Spectra", New-York: Academic, 1959

# Appendix 1: Properties of spherical harmonics

The spherical harmonics can be computed using recurrent equations [4] [14]:

$$\begin{aligned}
Y_{lm}(\theta,\phi) &= Z_{lm}(\theta)e^{im\phi} \text{ where } Z_{lm}(\theta) \text{ is real} \\
Z_{l+1,m}(\theta) &= \sqrt{\frac{(2l+1)(2l+3)}{(l+m+1)(l-m+1)}}\left\{\cos\theta Z_{lm}(\theta) - \sqrt{\frac{(l+m)(l-m)}{(2l+1)(2l-1)}}Z_{l-1,m}(\theta)\right\} \\
Z_{ll}(\theta) &= d_l(\sin\theta)^l \text{ with } d_l = \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}} \\
Z_{l,-m}(\theta) &= (-1)^m Z_{lm}(\theta)
\end{aligned}$$

A useful property is $Y_{l,-m}(\theta,\phi) = (-1)^m Y^*_{lm}(\theta,\phi)$

One can also prove [14] that there exists a link between the spherical harmonics and Legendre Polynomials:

$$Y_{lm}(\theta,\phi) = (-1)^m\sqrt{\frac{2l+1}{4\pi}}\sqrt{\frac{(l-m)!}{(l+m)!}}P_{lm}(cos\theta)e^{im\phi}$$

This relation is true for $m \geq 0$ (for $m < 0$, one can use the property above).

$$P_{lm}(u) = \sqrt{(1-u^2)^m}\frac{d^m}{du^m}P_l(u)$$

and $P_l(u)$ is the Legendre Polynomial of order $l$:

$$P_l(u) = \frac{(-1)^l}{2^l l!}\frac{d^l}{du^l}(1-u^2)^l$$

$P_l$ is a linear sum of $u^l, u^{l-2}, u^{l-4}, ...$ Hence, $P_{lm}(cos\theta)$ is the product of $sin^m\theta$ with a polynomial of degree $l-m$ in $cos\theta$ (which is a linear sum of $cos^{l-m}\theta, cos^{l-m-2}\theta, cos^{l-m-4}\theta, ...$).

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



Therefore, since $cos(m\phi)$ and $sin(m\phi)$ are homogeneous polynomials of degree $m$ in $(cos\phi, sin\phi)$, the spherical harmonic $Y_{lm}$ is a linear sum of terms in $sin^m\theta\, cos^a\theta\, sin^{m-b}\phi\, cos^b\phi$, where $0 \leq a \leq l-m$ and $0 \leq b \leq m$. This property will be used to derive the relation between our invariants and moments.

The explicit formulas for the first spherical harmonics are given below:

$$Y_{00}(\theta,\phi) = \frac{1}{\sqrt{4\pi}}$$

$$Y_{1,\pm1}(\theta,\phi) = \mp\sqrt{\frac{3}{8\pi}} sin\theta e^{\pm i\phi}$$

$$Y_{1,0}(\theta,\phi) = \sqrt{\frac{3}{4\pi}} cos\theta$$

$$Y_{2,\pm2}(\theta,\phi) = \sqrt{\frac{15}{32\pi}} sin^2\theta e^{\pm 2i\phi}$$

$$Y_{2,\pm1}(\theta,\phi) = \mp\sqrt{\frac{15}{8\pi}} sin\theta cos\theta e^{\pm i\phi}$$

$$Y_{2,0}(\theta,\phi) = \sqrt{\frac{5}{16\pi}}(3cos^2\theta - 1)$$

# Appendix 2: Properties of Clebsch-Gordan coefficients

Clebsch-Gordan coefficients are coupling coefficients between angular momenta of two particles. They are real numbers, and some properties are given below [4] [14]:

**C1:** *Condition of existence*

$\langle l_1 m_1 l_2 m_2 | lm \rangle = 0$ if the conditions below are not simultaneously verified:

a) $m = m_1 + m_2$

b) $|l_1 - l_2| \leq l \leq l_1 + l_2$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



c) $-l_1 \leq m_1 \leq l_1$
d) $-l_2 \leq m_2 \leq l_2$
e) $-l \leq m \leq l$

**C2:** *Orthogonality*
$\sum_{m_1,m_2} \langle l_1 m_1 l_2 m_2 | lm \rangle \langle l_1 m_1 l_2 m_2 | l' m' \rangle = \delta_{ll'} \delta_{mm'}$
$\sum_{l,m} \langle l_1 m_1 l_2 m_2 | lm \rangle \langle l_1 m_1' l_2 m_2' | lm \rangle = \delta_{m_1 m_1'} \delta_{m_2 m_2'}$

**C3:** *Recurrence*

$$\begin{aligned} \sqrt{l(l+1) - m(m-1)} \langle l_1 m_1 l_2 m_2 | lm \rangle &= \sqrt{l_1(l_1+1) - m_1(m_1-1)} \langle l_1, m_1 - 1, l_2, m_2 | l, m-1 \rangle \\ &+ \sqrt{l_2(l_2+1) - m_2(m_2-1)} \langle l_1, m_1, l_2, m_2 - 1 | l, m-1 \rangle \end{aligned}$$

$$\langle l, m, l, -m | 0, 0 \rangle = \frac{(-1)^{l-m}}{\sqrt{2l+1}} \quad \text{and} \quad \langle 00lm | lm \rangle = 1$$

**C4:** *Symmetries*
a) $\langle l_1, -m_1, l_2, -m_2 | l, -m \rangle = (-1)^{l_1+l_2+l} \langle l_1 m_1 l_2 m_2 | lm \rangle$
b) $\langle l_2 m_2 l_1 m_1 | lm \rangle = (-1)^{l_1+l_2+l} \langle l_1 m_1 l_2 m_2 | lm \rangle$
c) $\langle l_1 m_1 l_2 m_2 | lm \rangle = (-1)^{l_2+m_2} \sqrt{\frac{2l+1}{2l_1+1}} \langle l_2, -m_2, l, m, | l_1, m_1 \rangle$

# Appendix 3: Proof of some properties of the invariants

**P1:**

$$\begin{aligned} (c_l^m)^* &= \langle Y_{lm} | \Psi \rangle^* \\ &= \langle (Y_{lm})^* | \Psi^* \rangle \\ &= \langle (-1)^m Y_{l,-m} | \Psi \rangle \\ &\qquad \text{because } \Psi^* = \Psi \text{ and } (Y_{lm})^* = (-1)^m Y_{l,-m} \\ &= (-1)^m c_l^{-m} \end{aligned}$$

**P2:**

$$\begin{aligned}
\Pi(l_2,l_1)_l^m &= \sum_{m_1,m_2} \langle l_2 m_1 l_1 m_2 | lm\rangle c_{l_2}^{m_1} c_{l_1}^{m_2} \\
&= \sum_{m_1,m_2} (-1)^{l_1+l_2+l} \langle l_1 m_2 l_2 m_1 | lm\rangle c_{l_2}^{m_1} c_{l_1}^{m_2} \qquad \text{(using C4-b)} \\
&= (-1)^{l_1+l_2+l} \Pi(l_1,l_2)_l^m
\end{aligned}$$

**P3:**

$$\begin{aligned}
[\Pi(l_1,l_2)_l^m]^* &= \sum_{m_1,m_2} \langle l_1 m_1 l_2 m_2 | lm\rangle (c_{l_1}^{m_1})^* (c_{l_2}^{m_2})^* \\
&= \sum_{m_1,m_2} (-1)^{l_1+l_2+l} \langle l_1,-m_1,l_2,-m_2 | l,-m\rangle (-1)^{m_1+m_2} c_{l_1}^{-m_1} c_{l_2}^{-m_2} \\
&\qquad \text{(using C4-a and P1)} \\
&= (-1)^{l_1+l_2+l+m} \Pi(l_1,l_2)_l^{-m} \\
&\qquad \text{(using C1-a)}
\end{aligned}$$

**P5:**
a)

$$\begin{aligned}
P(l,l_2,l_1) &= \Pi(l_2,l_1)_l^m c_{lm} \\
&= (-1)^{l_1+l_2+l} \Pi(l_1,l_2)_l^m c_{lm} \qquad \text{(using P2)} \\
&= (-1)^{l_1+l_2+l} P(l,l_1,l_2)
\end{aligned}$$

b)

$$\begin{aligned}
P(l,l_1,l_2) &= \sum_{m,m_1,m_2} \langle l_1 m_1 l_2 m_2 | lm\rangle c_{l_1}^{m_1} c_{l_2}^{m_2} c_{lm} \\
&= \sum_{m,m_1,m_2} (-1)^{l_2+m_2} \sqrt{\frac{2l_1+1}{2l+1}} \langle l_2,-m_2,l,m | l_1,m_1\rangle c_{l_2}^{m_2} (-1)^m c_l^{-m} (-1)^{m_1} c_{l_1,-m_1} \\
&\qquad \text{(using C4-c and P1)} \\
&= \sum_{m,m_1,m_2} (-1)^{l_2+m_2} \sqrt{\frac{2l_1+1}{2l+1}} (-1)^{l_1+l_2+l} \langle l_2,m_2,l,-m | l_1,-m_1\rangle c_{l_2}^{m_2} (-1)^m c_l^{-m} (-1)^{m_1} c_{l_1,-m_1} \\
&\qquad \text{(using C4-a)}
\end{aligned}$$

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



$$= (-1)^{l+l_1}\sqrt{\frac{2l_1+1}{2l+1}}P(l_2,l,l_1)$$

(using C1-a)

**P6:**

$$\begin{aligned} P(l,l_1,l_2)^* &= \sum_m [\Pi(l_1,l_2)_l^m]^* c_l^m \\ &= \sum_m (-1)^{l_1+l_2+l+m}\Pi(l_1,l_2)_l^{-m}(-1)^m(c_l^{-m})^* \\ &\qquad \text{(using P3 and P1)} \\ &= (-1)^{l_1+l_2+l}P(l,l_1,l_2) \end{aligned}$$

# Appendix 4: Suppression of degeneracy using the Hamiltonian

Here, we propose to suppress the degeneracy mentioned in section 6.1 using the Hamiltonian operator corresponding to the electric field. In quantum mechanics, the eigenfunctions of the electron in the hydrogen atom are the solutions to the system below:

$$\begin{aligned} H|\Omega_{klm}\rangle &= E_{kl}|\Omega_{klm}\rangle \\ \mathbf{L}^2|\Omega_{klm}\rangle &= l(l+1)|\Omega_{klm}\rangle \\ L_Z|\Omega_{klm}\rangle &= m|\Omega_{klm}\rangle \end{aligned}$$

where $H$ is the Hamiltonian (energy operator), and $E_{kl}$ its eigenvalues.

$$H = -\frac{\bar{h}}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{2\mu r^2}\mathbf{L}^2 - \frac{q_e^2}{r}$$

where $\bar{h}$ is the Planck constant, $\mu$ is the magnetic constant, and $q_e$ is the electric quantum.

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



The solutions are:

$$R_{kl}(r) = r^l e^{-r/(k+l)} \sum_{q=0}^{k-1} c_q r^q$$

where

$$c_q = -\frac{2(k-q)}{q(q+2l+1)(k+l)} c_{q-1}$$

and $c_0$ is determined by the normalization condition. The practical problem with this kind of solution is that it is difficult to determine an upper bound for $k$. Thence, in section 6.1, we propose to use harmonic functions instead.

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



# List of Figures

Signal Processing, vol 45, no 1, pp. 1-22, July 1995



# Contents