Development of 3D invariants using linear algebra and tensor theory

Gilles Burel, Hugues Henocq, Jean-Yves Catros

Europe-China Workshop on Geometrical modelling and Invariants for Computer Vision,
April 27-29th, 1995, Xi'an, China

DEVELOPMENT OF 3D INVARIANTS USING LINEAR ALGEBRA AND TENSOR THEORY

Gilles BUREL, Hugues HENOCQ and Jean-Yves CATROS
(Thomson CSF/LER, Av. de Belle Fontaine, 35510 Cesson-Sévigné, France)

Abstract

A framework for development of 3D invariants is proposed. Linear algebra (projection on the basis of spherical harmonics) and tensor theory are used to derive invariants with respect to 3D rotations. These invariants can be used to recognize 3D objects using 3D information, as shown in the experiments.

1. Introduction

Object recognition is a key point in computer vision and automatic scene analysis. Many works have been dedicated to 2D pattern recognition, either by structural or statistical approach. 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 statistical approach, a feature vector of fixed length is computed from the 2D image of an object. The components of the feature vector may be, for instance, the moments of Hu [5], or Fourier descriptors [3] [4]. Recognition is then achieved by a classifier, such as a k-nearest-neighbours, or a neural network.

Few studies have been devoted to 3D pattern recognition, by using full 3D information. However, it becomes nowadays easier and easier to obtain 3D information of objects, thanks to the development of computer tomographic reconstruction, X-ray scanners, etc. Recent new and interesting research has been devoted to the development of 3D moment invariants [7] [8]. These moments, which are invariant under 3D translation and rotation of the object, can be used as the components of a feature vector for a statistical approach. However, these moments are difficult to derive, and only a small number of them is explicitly derived. Thus, it does not allow long feature vectors, which may be useful for discriminating close shapes. Furthermore, due to their definition, that involves multiplications by powers of the coordinates, the impact of digitizing errors on the moments is important. In fact, only the low-order moments are little sensitive to noise. Higher order moments are less reliable.

We propose an approach based on some results of quantum mechanics and group theory [10] [6] to derive 3D invariants that are not moments. These invariants present the advantages to be derived in a simple and systematic way, as well as to avoid the drawbacks of moments, such as multiplication by high powers of the coordinates. Furthermore, hundreds of invariants can be derived, allowing long features vectors. Although we have not achieved to prove it at present, such vectors may contain enough information to define the 3D shape without ambiguity (i.e. it would be possible to reconstruct the 3D shape from the knowledge of the feature vector).

The approach can be used for automatic object recognition in domains where 3D information is available. For example, in the medical field, it could be used for automatic classification of bones (figure 2). In robotics, a robot equipped with a camera could take multiple views to built a 3D representation of an object, and then use this approach to recognize the object. The proposed approach may also be used for associative access to a database of 3D objects: such a system would provide the shapes from the database which are the "most similar" to a shape selected by an operator.

The approach consists in decomposing the object onto an orthonormal basis composed of the eigenvectors of the angular momentum operator of quantum mechanics (§2). Then, using Clebsch-Gordan coefficients, contravariant tensors of order 1 are built, and 3D invariants are obtained by tensor contraction (§3). The approach offers an alternative to structural methods for 3D objects description and recognition. A generalization is presented in §4, and experimental results are provided in §5 to illustrate the method.

2. Decomposition onto the Basis of Spherical Harmonics

In this section, we consider objects represented by surface information. The generalization to volume information is presented in section 4. Let us note $\mathcal{F_S}$ the space of differentiable functions from $[0, \pi]$x$[0, 2\pi]$ to $\mathcal{C}$, with finite energy ($\int d\phi \int sin\theta \, d\theta |\Psi(\theta, \phi)|^2 < \infty$). To each 3D object, we associate a function $|\Psi\rangle$ such that $\Psi(\theta, \phi)$ is the distance between the center of gravity of the object and the farthest point still belonging to the object in the direction given by the spherical coordinates (θ, ϕ).

The spherical harmonics are functions of $\mathcal{F_S}$ which can be computed by recurrence equations. They are the eigenvectors common to two operators of quantum mechanics [10] [9]: $\mathbf{L}^2$ (square of angular momentum) and L_Z (projection of angular momentum along Z axis). The set of spherical harmonics $\{|Y_{lm}\rangle;\ l = 0, ..., \infty;\ m = -l, ..., l\}$ is an orthonormal basis of $\mathcal{F_S}$. Hence, any function $\Psi(\theta, \phi)$ can be described by its coordinates in this basis:

$$c_l^m = \langle Y_{lm}|\Psi\rangle = \int_0^{2\pi} d\phi \int_0^{\pi} sin\theta \ d\theta \ Y_{lm}^*(\theta, \phi) \ \Psi(\theta, \phi) \tag{1}$$

and the reconstruction equation is:

$$|\Psi\rangle = \sum_{l,m} c_l^m |Y_{lm}\rangle \tag{2}$$

The effect of a rotation of the object on these coordinates is given by:

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

This equation shows the interest of reasoning in the basis of spherical harmonics instead of the canonical basis: $\mathcal{F}_{\mathcal{S}}$ can be decomposed in a direct sum of orthogonal subspaces, each of which is globally invariant by rotation:

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

3. Computation of Invariants using Tensor Theory

At first, let us recall some tensor definitions and properties. Consider a vector space $\mathcal{V}$ and a basis $\{e_i\}$. Any vector x of $\mathcal{V}$ may be represented as

$$x = \sum_i x^i e_i = x^i e_i \tag{3}$$

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:

$$\tilde{e}_i = \alpha_i^j e_j \quad \text{or} \quad e_i = \beta_i^j \tilde{e}_j \tag{4}$$

where α_i^j denotes a linear transformation, and β_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:

$$\tilde{A}_{k_1 \ldots k_p}^{n_1 \ldots n_q} = \alpha_{k_1}^{l_1} \ldots \alpha_{k_p}^{l_p} \beta_{m_1}^{n_1} \ldots \beta_{m_q}^{n_q} A_{l_1 \ldots l_p}^{m_1 \ldots m_q} \tag{5}$$

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

$$\mathcal{A}_{lm}^{ijk} = \eta_l^{ij} \mu_m^k \tag{6}$$

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

$$\mathcal{A}_m^{ij} = \eta_n^{ij} \mu_m^n \tag{7}$$

and

$$\mathcal{A}^i = \eta_n^{ij} \mu_j^n \tag{8}$$

preserve tensor properties.

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

$$a_i = \langle a|e_i\rangle \tag{9}$$
$$a = a^i e_i \tag{10}$$

Hence,

$$\begin{aligned} a_i &= \langle a^j e_j|e_i\rangle \\ &= (a^j)^* \langle e_j|e_i\rangle \end{aligned} \tag{11}$$

If the basis is orthonormal, then:

$$a_i = (a^i)^* \tag{12}$$

c_l^m is a contravariant tensor of order 1 of $\mathcal{E}_l$. It is 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 corresponding covariant tensor is $c_{lm} = (c_l^m)^*$

We will take profit of the theorem below to build new tensors:

<u>Theorem [1][2]:</u>
If T_l^m is a tensor of $\mathcal{E}_l$, then:

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

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 can be computed using recurrent equations [1] [9] [10]. Thanks to this result, we can build new tensors. In the sequel, we will use:

$$\Pi(l_1, l_2)_l^m = \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} \langle l_1\ m_1\ l_2\ m_2|l\ m\rangle c_{l_1}^{m_1} c_{l_2}^{m_2} \tag{13}$$

By multiplying a contravariant tensor of order 1 with a covariant tensor of order 1, and equaling the indexes, we obtain a tensor of order 0 (an invariant). Hence, we can build, for example, the invariants below:

$$\begin{aligned} N(l) &= c_l^m c_{lm} \\ P(l, l_1, l_2) &= \Pi(l_1, l_2)_l^m c_{lm} \\ Q(l, l_1, l_2, l_3, l_4) &= \Pi(l_1, l_2)_l^m \Pi(l_3, l_4)_{lm} \end{aligned}$$

Obviously, the process we have followed to build the invariants could be continued as far as we want. For instance, we could build another tensor:

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

and use T to build other invariants.

4. Extension to Full 3D Description

Method so far only captures the outer boundary because, in previous sections, the object was described by a function $\Psi(\theta,\phi)$. However, for objects of complex shapes, such a description may be ambiguous (i.e. two different objects may have the same description). The objective in 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). For example $\Psi(x,y,z)$ could be 1 inside the object and 0 outside. It could also represent the local density, or anything else related to the object.

Since any real object is limited, it can always be included into a sphere of radius 1 by a suitable choice of the unit of length. Let us call $\mathcal{F}_{\mathcal{V}}$ the space of functions from $\mathcal{V}$ to $\mathcal{C}$, where $\mathcal{V}$ is the interior of the unit sphere of $\mathcal{R}^3$.

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:

$$c_{kl}^m = \langle\Omega_{klm}|\Psi\rangle = \int_0^\infty r^2dr \int_0^{2\pi} d\phi \int_0^\pi sin\theta \ d\theta \ \Omega_{klm}^*(r,\theta,\phi) \ \Psi(r,\theta,\phi) \tag{14}$$

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.

Below, we prove that, if we take:

$$\Omega_{klm}(r,\theta,\phi) = R_k(r)Y_{lm}(\theta,\phi) \tag{15}$$

where the $Y_{lm}(\theta,\phi)$ are the spherical harmonics, and

$$R_k(r) = \sqrt{2}\frac{\sin(\pi kr)}{r} \tag{16}$$

where $k\geq 1$ and $l\geq 0$, then $\{|\Omega_{klm}\rangle\}$ is an orthonormal basis of $\mathcal{F}_{\mathcal{V}}$. Then, once the c_{kl}^m have been computed, the approach is exactly as previously. One has only to replace each l_i by (k_i,l_i) in the equations.

We have to check that this is an orthonormal system, and that it generates $\mathcal{F}_{\mathcal{V}}$. The orthonormality is proved by developing the scalar product:

$$\begin{aligned}\langle\Omega_{klm}|\Omega_{k'l'm'}\rangle &= \int_0^1 r^2dr \int_0^{2\pi} d\phi \int_0^\pi \sin\theta d\theta R_k^*(r)Y_{lm}^*(\theta,\phi)R_{k'}(r)Y_{l'm'}(\theta,\phi) \\ &= \langle Y_{lm}|Y_{l'm'}\rangle \int_0^1 r^2drR_k(r)R_{k'}(r) \\ &= \delta_{kk'}\delta_{ll'}\delta_{mm'}\end{aligned} \tag{17}$$

Then, using the properties of Fourier Series, and the definition of $R_k(r)$, it is straightforward to prove that $\{|\Omega_{klm}\rangle\}$ generates $\mathcal{F}_{\mathcal{V}}$.

5. Experimental Results

Figure 1 shows some spherical harmonics Y_{lm} represented as a function $r = |Y_{lm}(\theta, \phi)|$. Figure 2 shows the reconstruction of a vertebra at various resolutions. The original 3D image of the vertebra comes from scanner data. A simple thresholding has been performed to isolate the bone. The c_l^m are computed for $0 \leq l \leq L$, and then the shape of the vertebra is reconstructed using equation (2). Increasing the value of L provides a better representation of the shape.

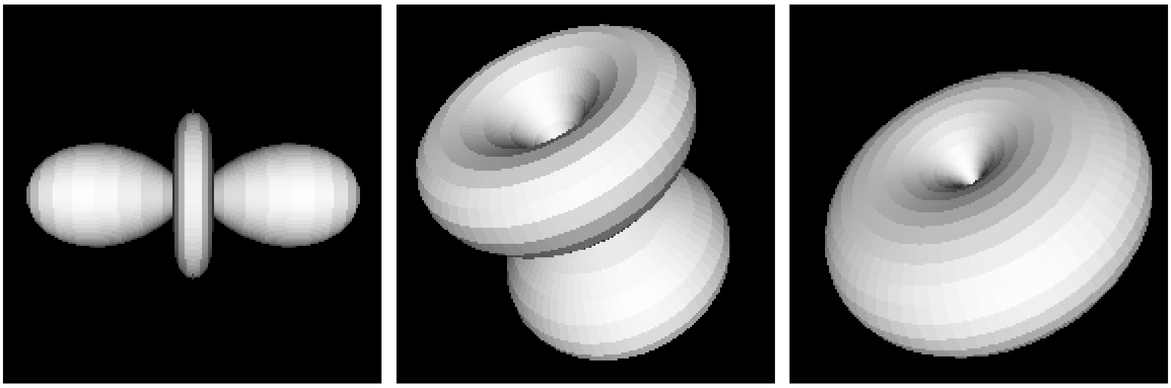

Figure 1: Module of the spherical harmonics Y20, Y21, Y22

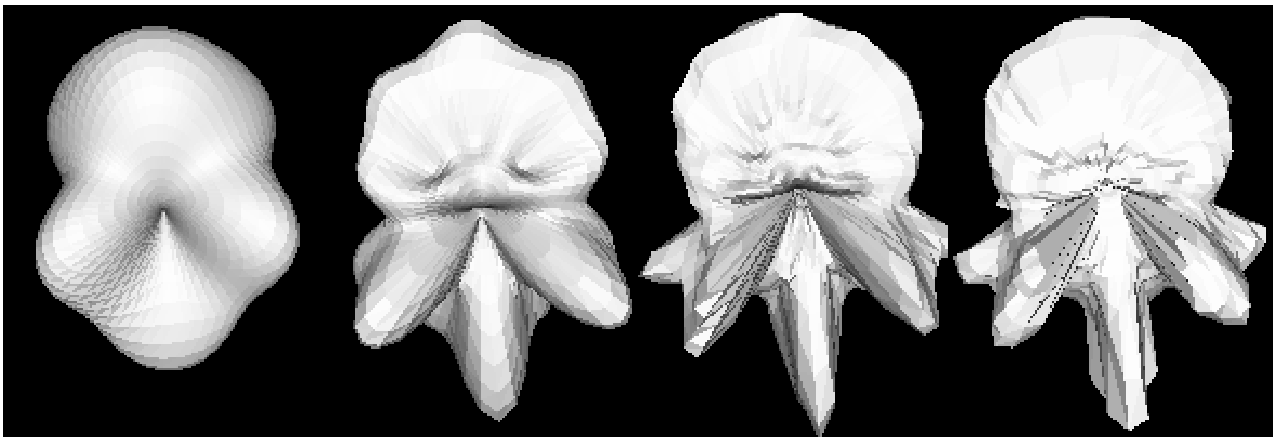

Figure 2: Reconstruction of a vertebra at various resolutions (from left to right: L= 4, 10, 20, and original)

The sampling step size has an impact on the values of the invariants because, in practice, equation (1) is evaluated by discretizing θ and ϕ. Experimental evaluations show that the error on the values of the invariants becomes noticeable only for rough sampling (more than 10^o).

Nearest-neighbour classification consists in computing the feature vector (here the 3D invariants) of the object to classify and comparing it with the feature vectors of models. The comparison is performed via the computation of a distance:

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

where $I_{object,i}$ stands for the i^{th} invariant of the object and $I_{model,i}$ stands for the i^{th} invariant of the model. The model which provides the minimal distance determines the class. Let us consider the three models on the right side of figure 3 (F14-1, X29-1, F15-1). An airplane to recognize is shown on the left side of the figure. The 3D-invariants of this airplane are computed and compared with the invariants of the models. The distances are mentioned on the arrows. As expected, the minimal distance is obtained when the object is compared with the X29 model.

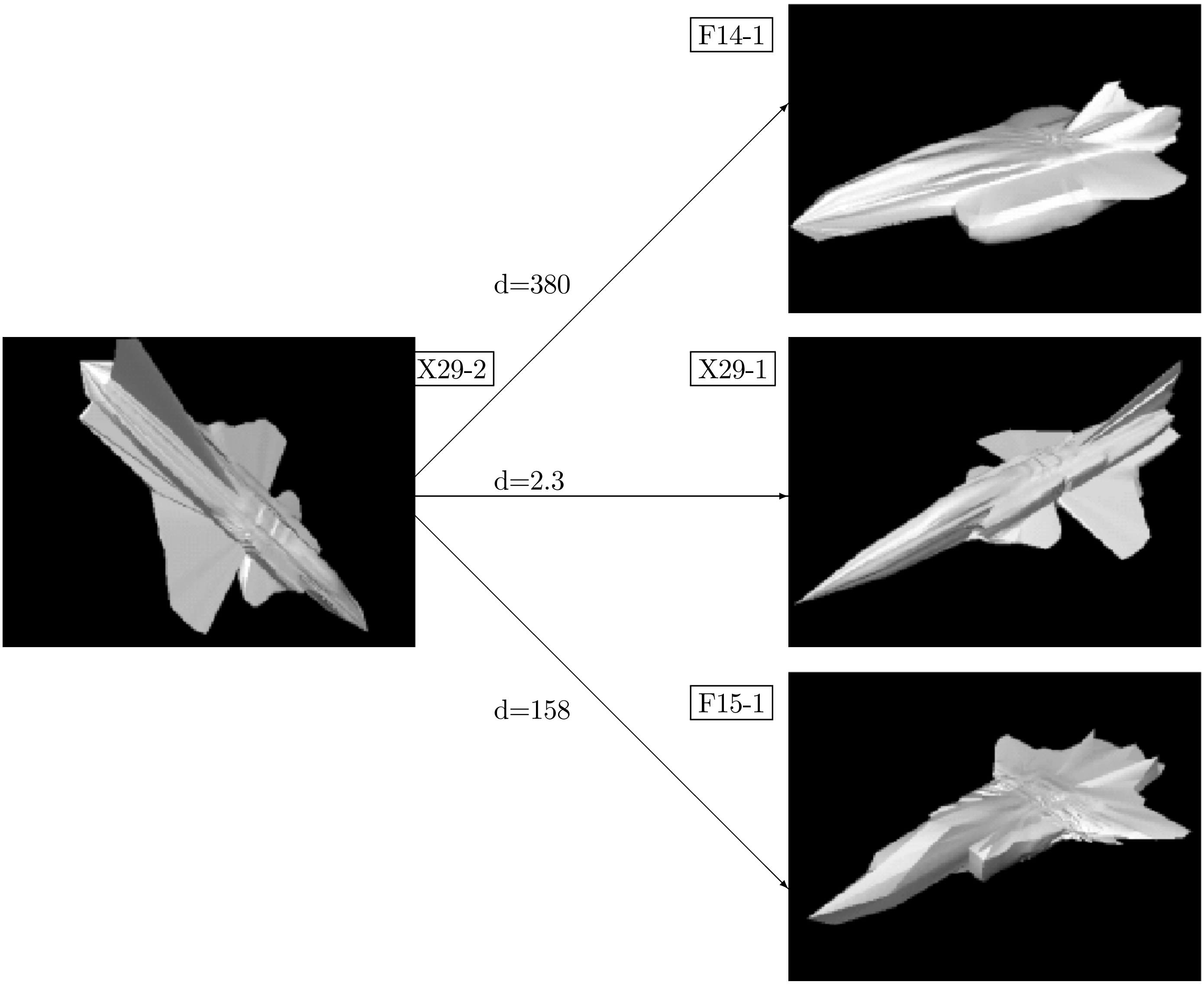

Figure 3: Comparison of an unknown object with three models

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 right class).

	X29-1	F14-1	F15-1
X29-2	2.3	380	158
F14-2	360	3	130
F15-2	167	122	3.4

These results illustrate an interesting feature of the approach: the ability to discriminate very close shapes.

6. Conclusion

A theoretical framework to derive 3D invariants has been proposed in this paper. The approach consists in decomposing 3D shapes onto an orthonormal basis composed of the eigen-vectors of the angular momentum operator, building contravariant tensors of order 1, and computing invariants by tensor contraction. The approach provides many invariants. Experimental results show that these invariants can be used to classify 3D shapes.

Such invariants offer an alternative to structural methods for description and recognition of 3D shapes, and the framework proposed in this paper allows the derivation of many invariants. An interesting feature of the approach is also the ability to discriminate very close shapes. Further work will include evaluation of the discrimination capabilities on larger data-bases, study of reversibility (i.e. does any subset of invariants characterize without ambiguity the shape of an object at a given resolution?), and application of the tensors to 3D object positioning. Some properties of the invariants may also be useful for automatic detection of symmetries of a 3D object.

Acknowledgements: The scanner images of vertebras have been provided by the LTSI Laboratory, University of Rennes.

References

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

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

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

[4] 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

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

[6] R. Lenz, "Group theoretical methods in image processing", Lecture notes in computer science, n^o 413, 1987

[7] 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

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

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

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