pp. 242-260

Towards a unitary formulation for invariant image description: application to image coding

Faouzi GHORBEL*

Abstract

Here, we introduce a joint topology and harmonic analysis formulation for the extraction of global shape descriptors which are invariant under a given group of geometrical transformations. The topology approach allows the rigorous definition of the notions of shape, shape space, the invariant features space, and a metric between shapes. Therefore a new definition of completeness is given. A stability criterion is defined mathematically. Using harmonic analysis, a unitary operator that is able to separate the shape information and the geometric transformation, allows us to extract a relevant invariant shape descriptors under a given group of transformations. It also gives a robust method for the evaluation ofthe global object motion. In the closed curves, some three-dimensional surfaces and planar gray level image cases, such an operator becomes the Fourier transform on a given group. Therefore, under some assumptions, a complete convergent set of invariant features exists and can be constructed. We derived from this a shape metric. Recent developments in image coding domain for moving pictures offer new perspectives to the application of the image invariant representations of regions and contours. Therefore, we intend to illustrate the importance of our approach in image coding and indexing applications.

Key words : Image coding, Invariance, Pattern recognition, Topology, Harmonic analysis, Geometrical shape, Geometrical transformation, Plane geometry, Three dimensional space.

VERS UNE FOMULATION UNITAIRE DES DESCRIPTIONS INVARIANTES D'IMAGE : APPLICATION AU CODAGE D'IMAGE

Résumé

Dans cet article, une formulation conjointe provenant de la topologie et de l'analyse harmonique du problème de l'extraction de primitives invariantes en vue d'une description de forme dans un cadre général est proposée. Ces descripteurs sont invariants par rapport à un groupe de transformations géométriques donné. L'approche topologique permet la définition rigoureuse des notions de forme, d'espace des formes et d'espace des invariants menant à une nouvelle de la complétude et de la stabilité. L'application de l'analyse harmonique, par la construction d'un opérateur qui est capable de séparer l'information de transformation géométrique de celle de forme, permet l'extraction de descripteurs invariants pertinents par rapport à un groupe de transformations donné. Une méthode d'estimation de mouvement robuste est également possible à partir de cet opérateur. Dans le cas des contours fermés, des images planes à niveaux de gris et de certaines surfaces tridimensionnelles, cet opérateur correspond exactement à la transformation de Fourier sur un groupe. On en déduit, alors, quand il est possible, une famille d'invariants complète et convergente ou stable. Les récents développements en codage d'images animées offrent de nouvelles perspectives aux méthodes de représentations invariantes des régions et des objets contours. Dans cet article l'importance de cette approche pour le codage d'image orienté objet ainsi que dans les applications de type indexation est illustrée.

Mots-clés : Code Image, Invariance, Reconnaissance forme, Topologie, Analyse harmonique, Forme, Transformation, Géométrie plane, Espace tridimensionnel.

Contents

* Groupe de Recherche Images et Formes de Tunisie de l'École Nationale des Sciences de l'Informatique (ENSI). Centre des Études et de Recherche des Télécommunications (CERT)/Direction des Études et de la Recherche-42, rue Asdrubal 1002 Tunis, Tunisie.

INTRODUCTION

The central goal of pattern recognition is to make computer intelligent. The importance of invariance to pattern recognition systems has been recognised. Invariance is a property of geometric configurations which remains unchanged under an appropriate class of transformations. The fundamental difficulty in recognising objects from images is that the appearance of a shape depends on viewpoint. Invariant parameters which can be measured directly from images, are used as shape descriptors. Many applications use invariant features such as robotic vision, shape classification in biologic or medical imaging, speech recognition, radar signature and so on...

The main classes of geometrical transformations of interest for image analysis applications are Euclidean, affine and projective. It is useful to classify these classes of transformations into a hierarchy which is based on the generality of the transformation. For example a projective transformation is more general than a Euclidean transformation because it applies to more situations and consequently there are fewer invariant properties. It is also important to precise the type of function which represents the object, whether it is a curve, a surface or a gray level volume.

Numerous approaches in image analysis domain have been developed for constructing invariant features, as the invariant moments [1, 2, 3] Fourier descriptors [4, 5, 6], Fourier Mellin transform [7, 8], the M-transform [9] and curvatures [10].

Invariant features extracted by differential geometry are local (Euclidean or affine curvature...) and serve to describe curves and surfaces independently of Euclidean, affine and projective transformations [10, 11, 12]. Methods based on harmonic analysis give global invariant descriptors. With such an approach, we can formulate invariance problems for new kind of objects such as planar grey level, three dimensional grey level objects or some kind of parametric three dimensional surfaces [7, 13, 14 ,15]. Within the context of practical applications, the invariant descriptors should necessarily verify a number of criteria. The following is a non exhaustive list:

1. *The fast computation.*

2. *The powerful discrimination* useful for object classification.

3. *The completeness*: All geometric objects related by a transformation will have the same invariant values. However, the converse is not generally true. Namely, two geometric objects with the same values of invariant features need not be related by a transformation. Completeness means that the invariant descriptors characterise the shape uniquely up to a transformation.

4. The *stability*: it guaranties that small shape variations are traduced by a small difference in the values of invariant parameters.

5. The definition of a metric between shapes: it has to be a right physical mean.

6. The *invariance* of the used curves or surfaces algorithms with respect to the considered transformations (Euclidean motion, affine motion, projective...) [19, 20].

The existence of a set of invariant features verifying such criteria depends upon the complexity of the considered scene: planar contours, grey level planar objects, three-dimensional surfaces, grey level three-dimensional volumes (useful in medical applications). It also depends on the considered class of geometric transformations, whether they are projective, affine, Euclidean.

In this paper, we propose a unitary joint topology and harmonic analysis formulation for the construction of global invariant descriptors verifying most of criteria cited above. The topology formulation is introduced to precise the shape notion, the invariant features, the shape space and the invariant features space. Therefore a rigorous definition of completeness, metric shape, stability, is given. By the harmonic analysis, we intend to construct a unitary operator which is able to separate the shape information and the geometric transformation one. This property gives two main results. First, it allows us the extraction of invariant shape descriptors under a given group of transformations. Next, it gives a relevant robust method for the estimation of the global object motion. Under some assumptions, this operator corresponds exactly to the known Fourier transform on a group. This approach gives satisfaction results in some cases such as the closed curves (submitted to the planar Euclidean or planar affine transformations), three dimensional surfaces and the planar grey level images. Here, we intend to give some answers to the completeness in these cases since the Inverse Fourier transform exists. Therefore, a complete and convergent set of invariant features could be constructed. A shape metric can be derived as well.

Recently, pattern recognition offers new perspectives and new methodologies for the object oriented image coding. Then, we illustrate the application of the proposed approaches for invariant features constructions in this context.

The paper is organised as follows: in section I, we introduce the joint topology and Harmonic analysis approach for the mathematical formulation of the geometric invariance in imaging system. In section II the cases of objects represented by a closed planar contour, are presented. The invariance with respect to Euclidean motions is considered. Such restriction allows us to illustrate Stability property by the construction of a complete and stable set of invariants [17]. We also show that the natural shape metric is the Hausdorf one which allows us the estimation of the rigid motion of closed planar contours with uniqueness. Afterwards, affine transformations are considered in section III where a complete and convergent set of invariant features under planar affine motions is proposed. With the same approach, the invariant description of planar grey level object is studied in section IV. A complete and convergent set of invariant descriptors with regard to planar similarities is presented [7]. In section V, shape invariant representations of three dimensional objects (spherical shaped form, Torus-shaped form surfaces and general grey level volumes) are constructed with the proposed approach. In this case, we underline that the obtained descriptors are relevant. However, completeness are not reached in spite of the existence of the Fourier transform.

Finally, in section VI, we describe the application of the model developed in section I in the case of planar closed curves with planar Euclidean motions, to the image coding domain. Object-based coding, for example, uses shape representations that should verify : completeness for the reconstruction, stability for the robustness under errors of transmission and a small shape change, real time computation for the moving pictures and the definition of distances for object matching and estimation of global object rigid motions.

This shape representation approach is very promising for the new coding techniques. Improvements are also foreseen for more general formulations both for the type of object and the type of motion such as :

– *planar grey level images* animated with *planar rigid motion* (Euclidean transformations),

– *planar closed contours* moving in three dimensional with *rigid motion* (planar affine or projective transformations),

– *surfaces* animated with *3D rigid motion* (3D Euclidean transformations), and

– *planar closed contours* moving with three dimensional rigid motion (3D Euclidean motion).

The application of this generalisation in coding will not be considered in this article.

I. A JOINT TOPOLOGY AND HARMONIC ANALYSIS APPROACH

I. 1. Notations and definitions

Throughout this paper, we shall denote the n-dimensional real [res. complex] vector space as R^n, [res. C^n], the multiplicative group of non zero [respectively positive] real numbers as R^* [res. R^{*+}], the group of positive planar similarities having the same centre $SimO^+$ and the multiplicative group of non zero complex numbers as C^*. With their topology, these groups have a topological group structure. We remark that all these groups are locally compact and Abelian.

Let us denote by $GA(n)$ the general affine group which can be seen as the product of the general linear group $GL(n)$ which is formed by all of non singular n-dimensional linear transformations with the translation space $T(n)$ which can be assimilated to R^n. $SA(n)$ is denoted as the special affine group. It is the product of $T(n)$ by the special linear group $SL(n)$ formed by all n-dimensional linear transformations having a unitary determinant. $M(n)$ denotes the group of Euclidean motions which can be seen as the product of $T(n)$ by the special orthogonal group $SO(n)$. S^{n-1} denotes the unit sphere of R^n. All of these sets are topologic groups and locally compact.

We denote the normed vector space of real vector functions f defined on a topological group G, by $L^p_C(G, \mu)$ verifying

$$\|f\|_p = \left(\int_G \|f(g)\|^p \, d\mu(g)\right)^{1/p} < +\infty$$

where p is a real number greater than 1 and μ is a positive measure on G.

We also denote the normed vector space of valued complex functions defined on G, by $L^p_R(G, \mu)$ verifying

$$\|f\|_p = \left(\int_G |f(g)|^p \, d\mu(g)\right)^{1/p} < +\infty$$

$\| f \|_p$ represents a norm of f. We sometimes simply denote by $L^p(G)$ the general case.

Finally, we denote the normed vector space of complex vector sequences $\{U_m\}_{m \in A}$ defined on a sub set A of Z(the set of all integers), by $l^p{}_C{}^n(A)$ verifying :

$$\left\|\{U_m\}_{m \in A}\right\|_p = \left(\sum_{m \in A} \|U_m\|^p\right)^{1/p} < +\infty$$

$\left\|\{U_m\}_{m \in A}\right\|_p$ represents a norm of the sequence $\{U_m\}_{m \in A}$.

I. 2. Topology formulation for shape representation

I. 2. 1. Object and parametrisation, group action :

Image analysis needs different kinds of function to represent planar contours, grey level planar images, three dimensional surfaces or volumes. It is important to distinguish two types of object representation. First, the called implicit or explicit functions are usually used to represent 2D or 3D grey level images. The uniqueness of such representations gives some advantages in image representations. The second type of representation concerns planar contours, curve in space and the surfaces in space. They are generally described by means of a parametrisation which is not unique.

In this case, two groups act on the space of all possible parametrisations. A first group which we denote by G', acts linearly on the left. The second which we denote by G, acts on the right. Such an action can be reduced to a translation in G when an adapted reparametrisation is considered (arclength, affine-arclength, projective-arclength...) [24]. In the proposed formulation, it is important to note that any object representation or parametrisation function has its support contained in G. So the considered objects is represented by a function of $L^2(G)$ which we call the representation space in general. So, the product group $G' \times G$ acts on $L^2(G, \mu)$ in the follow mean

$$G' \times G \times L^2(G, \mu) \rightarrow L^2(G, \mu)$$
$$[(g', g_0), \quad f] \rightarrow g' \circ f \circ g_0$$

where :

$$g' \circ f \circ g_0) : G \rightarrow R^n \text{ or } C^n$$
$$g \rightarrow g[f(gg_0)]$$

and μ is an invariant measure on G which is assumed locally compact. In such kind of group an invariant exists.

Harmonic analysis on G which consists of the definition of Fourier transform on a given group is suitable for separating the shape description and the transformation information. Effectively, when Fourier transform on G exists, the shift theorem remains available. It transforms the right action on a left operation. It is useful for extracting the transformation parameters from the object representation function. Therefore, both global object movement estimation and invariant features extraction under such action become simpler.

When objects are 2D or 3D grey level images, G' becomes the trivial group, and G is the group of all geometric transformations which the searched descriptors have to be invariant. However, when objects are represented by a parametrisation, G' is formed by the geometrical transformations and G represents all possible parametrisations. In this case, G and G′ are sub group of GL(n).

Definition 1: We say that two images f and h have the same shape if and only if it exists an element (g′, g) in G′ × G such that g′ 0 f 0 g = h (f and h will be assumed belonging to L²(G))

The relation *have the same shape* defines an equivalence relation in the representation space of object assumed to be equal $L^2(G)$, since $G' \times G$ forms a group.

I. 2. 2. Shape and the metric shape space :

A shape could be defined as an equivalence class of images in relation through the action of $G' \times G$.

Proposition 1: The shape set $\mathcal{S}$ *is the topologic quotient space* $L^2(G)/_{G' \times G}$.

This implies that the topology of the shape space is the quotient one. Such formulation explains mathematically the object proximity between the object in the mean of their shapes since Proposition 1 defined neighbourhood in the shape space.

Proposition 2: When $G' \times G$ *is compact group then the shape space becomes a metric space and its natural distance which we denote by D, is defined as :*

$$D(F, H) = \min_{f \in F} \min_{h \in H} \|f - h\|_{L^2(G)}$$

for all F and H in $\mathcal{S}$.

Proof: It is well known that when the quotient space can be a metric space its natural distance is D defined in proposition 2 [37]. As the space shape is formed by equivalence classes which form a partition of compact sets since any shape F can be parametrised by $G' \times G$ in the following mean : for any f belonging to $L^2(G)$ its shape is equal to the following set :

$$f = F = f(G' \times G) \left\{ g' 0 f 0 g / (g', g) \in G' \times G, f \in L^2(G) \right\}.$$

It is clear that F is a compact set. Therefore, when $D(F, H) = 0$ we have $F = H$. In the other hand, it is easy to see that D verifies the other distance axioms [37].

I. 3. Harmonic analysis formulation

The construction of a unitary operator or the Fourier transform on a group G, when it is possible, is the subject of the Harmonic analysis. Such construction needs the determination of the Haar measure on G and a set of unitary and irreducible group representations. We begin with recalling some elements of such theory.

I. 3. 1. Haar measure

Definition 2: Let G be a locally compact group, μ *is a left invariant measure on G if and only if* $\mu(gB) = \mu(B)$, *for every g in G and B a Borel set of G.*

Therefore, for every function f in $L^1(G)$, we have :

$$\int_G f(g_0 g) d\mu\,(g) = \int_G f(g) d\,(g)$$

Definition 3: μ *is called a left Haar measure on G if and only if it is positive and left invariant.*

Haar showed that in any locally compact group G, there exists a unique left or a right Haar measure up to a positive constant.

Definition 4: The group G is said to be unimodular if and only if the left Haar measure and the right one are similar.

All Abelian or compact groups are unimodular. It is well known that GA(n), GL(n), SL(n), SA(n), M(n), SO(n) and S^{n-1} are unimodular.

Examples :

– the normalized Haar measure in the real vector space R^n is the Lebesgue one

$$d\mu(x_1, x_2 \ldots, x_n) = dx_1, dx_2 \ldots dx_n,$$

– the Haar measure of the multiplicative group R* is

$$d\mu(x) = dx/x\,.$$

I. 3. 2. Group representation

Definition 5: Let H be an Hilbertien vector space. T is a representation of G if:

1. T(g) is an endomorphism of H for all g in G.

2. T(e) = Id_H *where* Id_H *is the identical operator of H and e is the identity element of G.*

3. T(gg′) = T(g) T(g′) for all g and g′ in G.

4. T is continuous on G.

Definition 6: T is irreducible if and only if H has not a proper subspace S invariant with respect to $T(T(S) \not\subset S)$.

Definition 7: T is unitary if and only if the matrix T(x) is unitary, for all x in G.

Definition 8: The set of all irreducible and unitary representations of G that we denote $\hat{G}$ is called the dual of G.

Proposition 2: If G is Abelian then $\hat{G}$ is a locally compact and Abelian group and any irreducible and unitary representation becomes a scalar. Then, there exists a unique normalized Haar measure $\mu_{\hat{G}}$ in $\hat{G}$.

I. 3. 3. Unitary operator and Fourier transform on a group

Definition 9: Let G is a unimodular group, m is his normalized Haar measure and assume that f belongs to $L^1(G, \mu)$. The following operator:

$$\hat{f}(\lambda) = \int_G f(g)\,[T_\lambda(g)]^{-1} d\mu(g) \qquad (1)$$

will be called the pseudo-Fourier transform if and only if $\{T(g)_\lambda\}$ forms a proper subset of all irreducible and unitary representations of G.

As we said in the introduction of this article, Shift theorem is the main idea for the extraction of invariant parameters under a given group of transformations and for the estimation of the geometrical transformation parameters. This theorem is conventionally known as that of Fourier Transform. So, it is easy to see that its extension to the Pseudo Fourier Transform remains valid.

Generalised shift theorem: *Let G be a unimodular group, μ is his normalized Haar measure and assume that f and h are two elements of $L^1(G, \mu)$. If $f(g) = h(g_0 g)$ for each g and g_0 in G, then we have the following relation between the corresponding pseudo Fourier transform:*

$$\hat{f}(\lambda) = T_\lambda(g_0)\hat{h}(\lambda)$$

Proof

When we substitute $f(g)$ in the expression of the pseudo Fourier transform by $h(g_0 g)$, we obtain:

$$\hat{f}(\lambda) = \int_G h(g_0 g)\,[T_\lambda(g)]^{-1} d\mu(g)$$

Let us consider the change of variable $u = g_0 g$:

$$\hat{f}(\lambda) = \int_G h(u)\,[T_\lambda(g_0^{-1} u)]^{-1} d\mu(g_0^{-1} u)$$

By using the fact that μ is a Haar measure of G and T_λ is a homomorphism, we obtain:

$$\hat{f}(\lambda) = [T_\lambda(g_0^{-1} u)]^{-1} \int_G h(u)\,[T_\lambda(u)]^{-1} d\mu(u)$$

Using again the fact that T_λ is a homomorphism, the theorem can be achieved.

Definition 10: When all unitary and irreducible representations of G can be determined then the pseudo Fourier transform becomes the Fourier transform on G if and only if, in the expression (1), the set $\{T_\lambda\}$ is formed by all these unitary and irreducible representations.

Examples

– If $G = R^n$ then $\hat{G} = R^n$, for all λ in R^n

$$\hat{f}(\lambda) = \int_{R^n} f(x) e^{-i < x, \lambda >} dx$$

is a multidimensional Fourier transform.

– If $G = R^*+$ then $\hat{G} = R$, for all λ in R

$$\hat{f}(\lambda) = \int_{R^*_+} f(x) x^{-i\lambda} \frac{dx}{x} = \int_0^{+\infty} f(x) x^{-i\lambda-1} dx$$

is the known Mellin transform.

When G is Abelian, the inverse Fourier transform exists and is defined by:

$$f(g) = \hat{\hat{f}}(g) = \int_{\hat{G}} \hat{f}(\lambda)\,[T_\lambda(g)]^{-1} d\mu_{\hat{G}}(\lambda)$$

for $\hat{f}$ belongs to $L^1(\hat{G}, \mu_{\hat{G}})$ where $\mu_{\hat{G}}$ is the normalized Haar measure of $\hat{G}$ and $T_\lambda(x)$ is an irreducible and unitary representation of G[7].

When $G = R^*+$, the dual group is the real line R and we have for every positive x

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\lambda)\, x^{i\lambda} d\lambda$$

This is the inverse Mellin transform.

In the Fourier transform space the action of $G' \times G$ is simpler and becomes only on the left

$$G' \times G \times L^2(\hat{G}, \hat{\mu}) \rightarrow L^2(\hat{G}, \hat{\mu})$$
$$[(g', g_0), \hat{f}] \rightarrow g\, T_\lambda(g_0)\hat{f}$$

I. 4. Invariance, completeness and stability

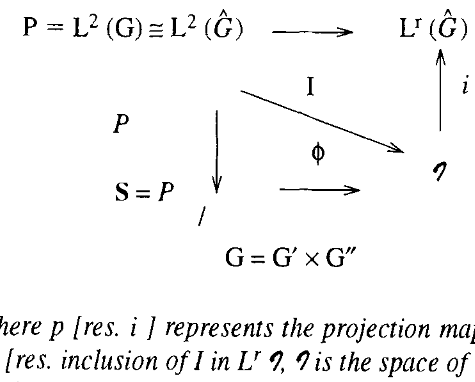
Definition 11: Let J is an index set, a map I from the space of all object representations P to the complex vector space C^J, i.e.

$$I: \mathcal{P} = L^2(G, \mu) \longrightarrow C^J$$
$$f \longrightarrow I(f) = \{I_n(f), n \in J\}$$

is called invariant under the action of the group G'xG on $\mathcal{P}$, if and only if I is constant on each shape (class of equivalence).

When Fourier transform exists, the index J could be computed and it is the dual set of G ($J = \hat{G}$).

Proposition 3: Assume that the Fourier transform exists on G. Let I be an invariant map, it is equivalent to say that the following diagram is commutative.



where p [res. i] represents the projection map of P on G [res. inclusion of I in L^r $\mathcal{I}$, $\mathcal{I}$ is the space of invariants and is a finite real number (r > 1) or infinite.

Completeness criterion introduced by Crimmins [34] could be formulated by the fact that the map ϕ is one to one and onto. It is also easy to see that completeness gives only an algebraic identification (one to one) between the shape space $\mathcal{S}$ and the space of invariant features $\mathcal{I}$. However, in practice, we often need to know whether the analysed objects have close shapes or not. So, the invariant descriptors have to be able to evaluate the degree of similarity between shapes. Such invariant descriptors are usually called stable [12] 14 [17]. With our topology approach, a rigorous definition to the stability becomes natural. So, we propose the following definition

Definition 13: A set of invariants is said to be stable if and only if the map f defined in proposition 3, is continuous on the shape space.

The continuity is the property which conserves the neighbourhood notion. Therefore, with such definition, Stability expresses the fact that if two invariant representations have a small difference, the objects they represent should have a small shape difference.

I. 5. Invariant distance on the shape space

A central problem in image analysis, in computer vision and in image coding is determining the extent to which one shape differs from another. The correlation and template matching can be viewed as techniques for determining the difference between shapes. In order to meet this goals, invariant descriptions could be used. The method consists on the computation of the Euclidean distance between these features. The obtained quantity is useful for the evaluation of the shape difference between images. In the following, we will give some responses to the two following questions:

– When does the distance exist in the mathematical mean?

– What distance do we have to choose?

Proposition 4: The existence of a complete and convergent set of invariant features {I(f)} implies that the shape space $\mathcal{S}$ is a metric space with one of the following metrics:

$$d_p(G, H) = \left(\int_{\hat{G}} |I(f)(\hat{g}) - I(h)(\hat{g})|^p \, d\mu_{\hat{G}}(\hat{g}) \right)^{1/p}$$

where p>1, f and h are two images having the shape G and H respectively

When the complete set of invariants is stable, the topology induced by the invariant map on the shape space coincides with its natural topology (topology of quotient space). However, when D, defined in proposition 2, is a metric, the distance d_p is equivalent to D if and only if the map of proposition 3 is uniformly continuous. In such situation the set of invariant features is said to be uniformly stable.

II. PLANAR CLOSED CONTOUR-EUCLIDEAN MOTION

We consider here the case of objects which could be represented with a planar and a closed contour and we search to extract a set of invariants under Euclidean motion group $M(2)$ by using the proposed approach. This construction will also be done by means of the properties cited above. Such restrictions allow us to illustrate the stability criterion and the existence of distances in the shape space. First, we show that the shape space is metric and that its natural metric is the Hausdorf one. Next, we present the distance induced by a complete and stable set of invariant features constructed with our approach.

II. 1. Object and parametrisation, group action

Using the notation of section I, it is easy to see that the group G′ corresponds to $M(2)$ which could be seen as the product of C by S^1. The suitable representation of planar and closed curve is a periodic parametrisation. In order to put all of these parametrisations in a space having known mathematical properties, we have to consider the same period for any parametrisation. This means that a normalisation procedure is required. Therefore, the group G (notation of section I) corresponds to the circle S^1 and the space of all normalised parametrisations can be assumed to be equal to $\mathcal{P} = L^2_C(S^1, d\theta)$

On the other hand, it is important to note that for a given curve there are several parametrisations. The difference between parametrisations for the same curve can be represented by the fact that the same arc is drawn with different speeds, when we want to extract invariants under Euclidean transformations. So, an arclenght reparametrisation is required. This fact implies that we go over the curve with a constant speed. Therefore, the following action summarises our formulation:

$$(C \times S^1) \times S^1 \times L^2_C(S^1, d\theta) \to L^2_C(S^1, d\theta)$$
$$[(b, \theta, s_0), f] \to e^{i\theta} f(s + s_0) + b$$

In order to justify rigorously the arguments abovementioned, we need to recall some definitions.

Definition 12: A parametrisation is a C^1-immersion f of an interval I into the complex plane C or into the real plane R^2

Definition 13: A periodic parametrisation is C^1-immersion f of the circle S^T (with T as radius) in the complex plane C.

Definition 14: Two parametrisations f and h are equivalent if and only if $h = f_0\psi$, where ψ is a change of variable (ψ' represents the speed).

Definition 15: An object is a closed arc.

Proposition 5: Any object O has a periodic parametrisation from which we can extract a periodic normalised arclength parametrisation (n.a.l.p.).

Proof of Proposition 5:

Let f be a T-periodic parametrisation of an object O

$$f: S^T \to C$$
$$t \to x(t) + iy(t)$$

The following parametrisation f^* defined by :

$$f^*: S^1 \xrightarrow{*} C$$
$$s \to f(t^{-1}(s))$$

where

$$s(t) = \frac{1}{L}\int_0^t \|f'(u)\|_2\, du \text{ and } L = \int_0^T \|f'(u)\|_2\, du$$

is a periodic n.a.l.p; of O. t^{-1} represents the inverse of the arclength function $s(t)$.

So, we say that O_1 et O_2 have the same shape if and only if for any two n.a.l.p. f and h of O_1 and O_2 respectively, we have:

$$(1) \qquad h(s) = e^{i\theta} f(s + s_0) + b$$

for all s in S^1, where θ and s_0 belong to S^1 and b is a complex number. S_0, θ and b represent respectively the difference between the arclength coordinates of the starting points on each curve, the rotation angle and the translation between the two objects. It is also important to remark that equation (1) defines an equivalence relation in the space $\mathcal{P}$.

Definition 16: A shape is an equivalence class of objects in relation. So, the shape space is the quotient space of $\mathcal{P}$ *by* $M(2) \times S^1$*.i.e.* $\mathcal{S} = L^2_C(S^1)/M(2) \times S^1$.

II. 2. Harmonic analysis formulation

As we have shown in section II. 1, G is equal to S^1 and the harmonic analysis has to be applied to this group. The Fourier transform exists and conform to the classical Fourier coefficients of a complex function defined on the unit circle. Shift theorem separates both shape and transformation information which can be formulated in the following lemma.

Lemma 1: Let f and h be the n.a.l.p. of two objects having the same shape, then:

$$a_n(f) = e^{i\theta} e^{2\pi i n l_0} a_n(h) + b\delta_n \text{ for all integer } n$$

where $a_n(f)$ *and* $a_n(h)$ *are respectively the sequence of Fourier coefficients of f and h.*

Therefore, the application of harmonic analysis transforms the previous action in the only left one:

$$(C \times S^1) \times S^1 \times l^2_C(Z) \to l^2_C(Z)$$

$$[(b, \theta, s_0), a_n f] \to e^{i\theta} e^{2\pi i n s_0} a_n(f) + b\delta_n$$

since Fourier transform is an isometric map. So, the shape space $\mathcal{S}$ is isomorphic to $l^2(Z)/M(2) \times S^1$. Unfortunately, M(2) is not a compact group, then the shape space does not appear explicitly metrizable.

II. 3. On the existence of distances on $\mathcal{S}$

In this paragraph, we intend to show the following important proposition in shape classification and motion analysis since it gives the expression of the natural distance between shapes and offers an efficient procedure which guarantees the uniqueness of rigid motion parameters.

Proposition 6: The shape space S of closed contours relative to the planar Euclidean transformations is equal to $l^2(Z^*)/ S^1 \times S^1$ *and is a metrizable space. Its natural metric is the Hausdorf distance which can be reduced to the following quantity:*

$$\Delta(F, H) = \inf_{(s,\theta) \in T_2} \left\| a_n(f) - e^{i(ns+\theta)} a_n(h) \right\|_{l^2(Z^*)}$$

$$= \inf_{(s,\theta) \in T_2} \left\{ \sum_{n \in Z^*} \left| a_n(f) - e^{i(ns+\theta)} a_n(h) \right|^2 \right\}^{1\backslash 2}$$

Proof of proposition 6:

From Lemma 1, it is easy to deduce that the sequence $\{a_n(f)$, for n in $Z^*\}$ represents completely the curve up to a translation. Therefore we have:

$$S^1 \times S^1 \times l^2_C(Z^*) \to l^2_C(Z^*)$$

$$[(\theta, s_0), a_n(f)] \to e^{i\theta} e^{i(ns+\theta)} a_n(f)$$

This implies that the shape space is the quotient space $l^2(Z^*)/ S^1 \times S^1$. The compactness of the group $S^1 \times S^1$. involves that the shape space is metrizable with the metric D defined in section I.

We will show now that D can be seen as the Hausdorf distance which is frequently used for the evaluation of distance between sets of points. Let F and H be two subsets of a metric space (X, d), then the following quantity

$$\Delta(F, H) = \max\big(\rho(F, H), \rho(H, F)\big)$$

where

$$\rho(F, H) = \max_{f \in F} \min_{h \in H} d(f, h)$$

is called the Hausdorff distance in the set of all closed and bounded subsets of X.

As we have seen above, the shapes are equivalence classes. So, all the shapes are disjoint sets. In the Euclidean case and after elimination of translations, they are closed and bounded. Since any planar and closed shape relative to the Euclidean motion could be represented by the following compact set:

$$F = \left\{ \left(e^{i\theta} e^{i(ns+\theta)} a_n(f)\right)_{n \in Z^*}, (\theta, s) \in [0, 2\pi] \times [0, 2\pi] \right\}$$

this one could be assimilated to the bidimensional Torus T_2. This fact proves that the shape space is metrizable [37]. So, the quantity ρ can be written as:

$$\rho(F, H) = \max_{(\theta, s) \in T^2} \min_{(\beta, v) \in T^2} \left\| e^{i\theta} e^{in\theta} a_n(f) - e^{i\beta} e^{inv} a_n(h) \right\|_{l^2(Z^*)}.$$

Then, we have:

$$\rho(F, H) = \max_{(\theta, s) \in T^2} \min_{(\beta, v) \in T^2} \left\| e^{i\theta} e^{in\theta} a_n(f) - e^{i(\beta-\theta)} e^{in(v-s)} a_n(h) \right\|_{l^2(Z^*)}$$

Using the following change of variable:

$$u = \beta - \theta \text{ and } w = v - s$$

the previous formula could be written as:

$$\rho(F, H) = \max_{(\theta, s) \in T^2} \min_{(u, w) \in T^2} \left\| a_n(f) - e^{iu} e^{inw} a_n(h) \right\|_{l^2(Z^*)}$$

The quantity inside the norm does not depend on θ and s. So, $\rho(F,H)$ Hausdorff distance can be reduced to :

$$\rho(F,H) = \min_{(u,w)\in T^2} \left\| a_n(f) - e^{iu} e^{inw} a_n(h) \right\|_{l^2(Z^*)}$$

In the same way, we can prove that :

$$\rho(F,H) = d(f,H) = d(F,H)$$

Finally, it is easy to see that :

$$\Delta(F,H) = D(F,H) = \max(h(F,H)\, h(F,H)) \\ = d(H) = d(F,H)$$

II. 4. A complete and stable set of invariants :

In this paragraph, we recall the complete and stable set of descriptors invariant with respect to the starting point on the curve and the planar Euclidean transformations introduced in [17].

For all p and q positive real numbers, the following sequence of complex numbers:

$$\begin{cases} I_n(f) = \dfrac{a_n^{n_0-n_1}(f)\, a_n^{n_0-n_1}(f)\, a_n^{n_0-n_1}(f)}{a_{n_0}(f)|^{n_1-n-p} |a_{n_1}(f)|^{n_1-n-q}} & \text{for } a_{n_0}(f) \neq 0 \text{ and } a_{n_1}(f) \neq 0 \\ I_n(f) = 0 & \text{for } a_{n_0}(f) = 0 \text{ or } a_{n_1}(f) = 0 \end{cases}$$

which belongs to $l^{2/(n_0-n_1)}(Z)$, forms a complete and stable set of invariant features with respect to planar Euclidean transformations and curve parametrisations.

The set $\{I_n\}$ is convergent because it belongs to $l^r(Z)$ for a given $r>1$ since:

$$I_n(f) = \left(\sum \left| I_n(f) \right|^{\frac{2}{n_0-n_1}} \right)^{n_0-n_1/2} \\ = \left| a_{n_0}(f) \right|^{p} \left| a_{n_1}(f) \right|^{q} \left\| \{a_n(f)\} \right\|_{2/n_0-n_1}$$

Therefore $\{I_n\}$ belongs to $l^{2}/(n_0-n_1)(Z)$. It is more convenient to compute Euclidean distance between features when, for example, some discriminant analysis algorithms are used for object classifications. So, n_0 and n_1 have to be chosen such that $n_0-n_1=1$ ($\{I_n(f)\}$ in $l^2(Z)$).

By using Lemma 1, the invariance can be easily obtained.

For $a_{n_0}(h) \neq 0$, $a_{n_1}(h) \neq 0$ and $a_n(h) \neq 0$ we have :

$$I_n(h) = \\ a^{n_0-n_1} e^{2\pi i n(n_0-n_1)l_0} a_n^{n_0-n_1}(f) a^{n_1-n} e^{2\pi i n(n_1-n_0)l_0} a_{n_0}^{n_1-n}(f) a^{n-n_0} \\ e^{2\pi i n_1(n-n_0)l_0} a_{n_1}^{n-n_1}(f) / \left|a\right|^{n_1-p-n_0-q} \left|a_{n_0}(f)\right|^{n_1-n-p} \left|a_{n_0}(f)\right|^{n-n_0-q}$$

By denoting: $\varphi_0 = \arg(a_{n_0}(f))$ and $\varphi_0 = \arg(a_{n_1}(f))$, the Fourier coefficients of a n.a.l.p. of any object can be expressed only according the set of invariant features $\{I_n\}$ up to an Euclidean motion and a starting point on the object

$$a_n(f) = I_n(f)^{\frac{1}{n_0-n_1}} e^{-i\left(\frac{n_1\varphi_0-n_0\varphi_1}{n_0-n_1}\right)} e^{in\frac{(\varphi_1-\varphi_0)}{n_0-n_1}} \\ \left[I_n(f)\right]^{\frac{p}{n_0-n}} \left[I_n(f)\right]^{\frac{q}{n_0-n}}.$$

This shows the completeness of this set of invariants. The stability can be obtained by showing that the map ϕ defined in Proposition 3 is continuous on the shape space. The proof is detailed in [14].

II. 5. The distance induced by invariants

It has been shown in [18], that completeness and stability properties induce the existence of a distance between shapes which have a right physical mean. Such a result is very important in object matching problems and object classification. This distance can be expressed as :

$$d_{2/n_0-n_1}(F,H) = (\{I_n(f)\}, \{I_n(h)\}) = \\ \left(\sum \left| I_n(f) - I_n(h) \right|^{2/(n_0-n_1)} \right)^{(n_0-n_1)/2}$$

Two fundamental remarks can be done:

– The set of invariants defined above is also invariant under the group of planar and positive similarities which is not a compact set. Moreover, this set of invariants makes the corresponding shape space metrizable.

– The distance D and the one induced by this set of invariants define an equivalent topology but they are not equivalent. In order to obtain this result, the uniform continuity of the map θ is required. With this condition, the set of invariants is called complete and uniformly stable. The existence of such a set is not known until now.

III. PLANAR CLOSED CONTOUR AND SPECIAL AFFINE MOTION

Here, we study planar closed contours with affine motions. We show how to represent dimensional space. This contour is projected on the image plane. The movement can be described with affine transformations if we assume that the contour is enough remoted from the camera. A complete set of invariants will be extracted under such transformations, afterwards, we give some precisions about possible shape distances.

III. 1. Object and parametrisation, group action

Following the notation of section I. 1, the group G' of the present case is the special affine motion group

$SA(2) = R^2 \times SL(2)$. The choice of the suitable parametrisation which we will develop later, implies that G is equal to S^1. The formulation of such problem can be described in the following operation:

$$(R^2 \times SL(2) \times S^1) \times L^2_{R^2}(S^1) \to L^2_{R^2}(S^1)$$
$$[(B, A, l_0), U_n(f)] \to Af(1 + 1_0) + B.$$

In order to validate this schema, we have to remind some definitions about affine geometry.

Propositions 7: Any object O has a periodic parametrisation from which we can extract a periodic normalised affine arclength parametrisation.

Proof of Proposition 7:

Let f be a T-periodic parametrisation of an object O. The quantity f_a^* defined by:

$$f_a^* \;:\; S^1 \to R^2$$
$$l \to f(t^{-1}(l))$$

where

$$f(t) = \frac{1}{L_a}\int_0^t \sqrt[3]{\left|\det\left(f'(u), f''(u)\right)\right|}\, du,$$

$$L_a = \int_0^T \sqrt[3]{\left|\det\left(f'(u), f''(u)\right)\right|}\, du$$

is a periodic normalised affine arclength of the object O.l^{-1} (t) represents the inverse function of the affine arclength function $l(t)$, and let denotes the determinant operator.

Definition 17: We say that two objects O_1 and O_2 have the same affine-shape if and only if

$$h(t) = Af(l + l_0) + B$$

where f and h are respectively a normalised affine arclength parametrisation of O_1 and O_2, A an element of SL(2, R) and B is a vector of R^2.

It is also important to that the last action defines in the space $\mathcal{P} = L^2_{R^2}(S^1)$ an equivalence relation.

Definition 18: An affine shape is an equivalence class of objects in relation. So, the affine shape space is the quotient space of $\mathcal{P}$ by the product group $SA(2, R) \times S^1 = \mathcal{S}_a$

Harmonic analysis formulation is available for the group S^1. The Fourier transform in the case is the Fourier coefficients of an affine arclength parametrisation of a given curve. The inverse Fourier exists, due to the fact that each element of $L^2_{R^2}(S^1)$ can be expanded into Fourier series. So, the following lemma gives the relation between Fourier coefficients of f and those of h.

Lemma 2 (Shift theorem): Let f and h are respectively the normalised affine arclength parametrisations of two objects having the same affine-shape then:

$$U_n(h) = e^{2\pi i n l_0} A\, U_n(h) + B\delta n$$

for all interger n, where $U_n(f)$ and $U_n(h)$ are respectively the bi-dimensional vectors formed by the Fourier coefficients of each component of f and h. B, A and l_0 are assimilated respectively to a translation vector, a special linear transformation and the difference between affine arclength coordinates of the starting point on each curve.

After normalisation with respect to the translations as we have done in the Euclidean case, the action on the Fourier space is reduced to the following operation:

$$(SL(2) \times S^1 \times l^2_{C^2}(Z^*)) \to l^2_{C^2}(Z^*)$$
$$[(A, l_0), f] \to e^{inl_0} A\, U_n(f)$$

Therefore the affine shape space $\mathcal{S}_a$ becomes the quotient $l^2_{C^2}(Z^*)/SL(2) \times S^1$. Unfortunately, it is important to note that the natural topology of the affine shape space is not metrizable since SL(2) is not a compact group because, for example, it contains the unbounded sub group G_a:

$$G_a = \left\{\begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix}, a\, R_+^*\right\}$$

To get rid of this handicap, we propose to construct, in the forthcoming paragraph, a complete and convergent set of invariants under special affine motion group and parametrisations which induce a distance on $\mathcal{S}_a$

III. 2. A complete set of invariants

In the following theorem, we intend to derive a complete set of invariant features under special affine motions and parametrisations. It is important to underline that the normalised affine parametrisation of an object O is a periodic fucntion f from n interval to the real plane. So, the Fourier coefficients of f which we denote $\{U_n(f)\}$ belongs to $l^2_{C^2}(Z)$

Theorem 1: The following two sequences of complex numbers:

$$J^1_{k_1}(f) = \left|\det\left(U_{k_1}(f), U_{k_0}(f)\right)\right|, \; J^1_{k_1}(f) = \left|\det\left(U_{k_2}(f), U_{k_0}(f)\right)\right|$$

$$J^1_{k_1}(f) = \left[\det\left(U_{k_1}(f), U_{k_0}(f)\right)\right]^{k_1-k_2}\left[\det\left(U_{k_1}(f), U_{k_0}(f)\right)\right]^{k_1-k}\left[\det\left(U_{k_2}(f), U_{k_0}(f)\right)\right]^{k-k_1}$$

for all $k \in N/\{0, k_1, k_2\}$

$$J^2_{k_1}(f) = \left|\det\left(U_{k_1}(f), U_{k_3}(f)\right)\right|, \; J^2_{k_2}(f) = \left|\det\left(U_{k_2}(f), U_{k_3}(f)\right)\right|$$

$$J^2_{k_1}(f) = \left[\det\left(U_{k}(f), U_{k_3}(f)\right)\right]^{k_1-k_2}\left[\det\left(U_{k_1}(f), U_{k_0}(f)\right)\right]^{k_2-k}\left[\det\left(U_{k_2}(f), U_{k_0}(f)\right)\right]^{k-k_1}$$

for all $k \in N/\{0, k_1, k_2\}$

form a complete set of invariant features under the group SA(2, R) and the curve parametrisations.

Proof:

Invariance: Let f and h be respectively two normalised affine arclength parametrisations of affine shapes F and H. As we have seen above, Shift theorem implies that:

$$_k(f) = e^{2\pi i k l_0} A\, U_k(h) \quad \textit{for all } k \in Z^*$$

Therefore :

$$J^1_{k_1}(h) = \left|\det\left(\mathbf{e}^{2i\pi k_1 l_0} A U_{k_1}(h),\, \mathbf{e}^{2i\pi k_0 l_0} A U_{k_0}(f)\right)\right|,$$
$$= \left|\mathbf{e}^{2i\pi(k_1+k_0)l_0} \det A \det\left(U_{k_1}(f), U_{k_0}(f)\right)\right|$$
$$= \left|\det\left(U_{k_1}(f), U_{k_0}(f)\right)\right| = J^1_{k_1}(f)$$

The same proof can be obtained for the feature $J^1_{k_2}$.
For all $k \in N/\{0, k_1, k_2\}$

$$J^1_{k_1}(h) = \det A \{\mathbf{e}^{2i\pi(k+k_0)(k_1-k_2)l_0}\left[\det\left(U_k(f), U_{k_0}(f)\right)\right]^{k_1-k_2}$$
$$\mathbf{e}^{2i\pi(k_1+k_0)(k_2-k)l_0}\left[\det\left(U_{k_1}(f), U_{k_0}(f)\right)\right]^{k_2-k}$$
$$\mathbf{e}^{2i\pi(k_2+k_0)(k-k_1)l_0}\left[\det\left(U_{k_2}(f), U_{k_0}(f)\right)\right]^{k-k_1}\}$$
$$= \mathbf{e}^{2i\pi(k+k_0)(k_1-k_2)+(k_1+k_0)(k_2-k)+(k_2+k_0)(k-k_1)l_0} J^1_k(f) = J^1_k(f)$$

By substituting the interger k_0 by k_3 in the formula of the second set defined in theorem 1, the proof of the invariance under *SL(2)* and parametrisation for this set becomes similar to those of the first set.

Completeness

The basic idea of the proof consists in showing that the knowledge of the invariant features allows the determination of a normalised affine arclength parametrisation up to a special affine transformation and a starting point on the curve.

Let us denote by $\theta_1 = \arg\{\det(U_{k_1}, U_{k_0})\}$,

$$\theta_2 = \arg\{\det(U_{k_2}, U_{k_0})\}$$

$\theta_3 = \arg\{\det(U_{k_1}, U_{k_3}\}$, $\theta_4 = arg\{\det(U_{k_3}, U_{k_0}\}$ and $\theta_5 = \arg\{\det(U_{k_0}, U_{k_3}\}$, where $\arg(z)$ represents the argument of the complex number z.

So, for all $k \in N/\{0, k_1, k_2\}$, we have :

$$\left[\det(U_k, U_{k_0})\right] = \left[J^1_k\right]^{1/k_1-k_2}\left[J^1_{k_1}\right]^{k_2-k/k_2-k_1}\left[J^1_{k_2}\right]^{k-k_1/k_2-k_1} \mathbf{e}^{\frac{i\theta_1(k_2 \mp k)+\theta_2(k-k_1)}{k_2 \mp k_1}} \quad \text{(a)}$$

We also have :

$$\left[\det(U_k, U_{k_3})\right] = \left[J^2_k\right]^{1/k_1-k_2}\left[J^2_{k_1}\right]^{k_2-k/k_2-k_1}\left[J^2_{k_2}\right]^{k-k_1/k_2-k_1} \mathbf{e}^{\frac{i\theta_3(k_2-k)+\theta_4(k-k_1)}{k_2 \pm k_1}} \quad \text{(b)}$$

For clarity, we propose to denote now the following quantities by :

$$U_k = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, U_{k_0} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, U_{k_0} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$\text{and} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \det(U_k, U_{k_0})_1 \\ \det(U_k, U_{k_0}) \end{bmatrix} \quad \text{(c)}$$

Using the above notation, equation (c) can be summarised by the following linear system :

$$\begin{cases} a_2x_1 - a_1x_2 = c_1 \\ b_2x_1 - b_1x_2 = c_2 \end{cases}$$

The unique solution of this linear system is :

$$U_k = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \dfrac{c_2a_1 - c_1b_1}{a_1b_2 - a_2b_1} \\ \dfrac{c_2a_2 - c_1b_2}{a_1b_2 - a_2b_1} \end{bmatrix} = \frac{\mathbf{e}^{-i\theta_5}}{J^2_{k_0}} \begin{bmatrix} c_2a_1 - c_1b_1 \\ c_2a_2 - c_1b_2 \end{bmatrix}$$

Using equations (a) and (b), it is easy to see that the Fourier coefficients U_k can be expressed only with the two invariant sets and the five argument values :

$$U_k = F(J^1_k, J^2_k, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$$

They correspond to the special affine transformation and the starting point on the curve since there is a relation between these five parameters. A representative curve of the considered affine shape can be reconstructed using the Fourier expansion. This achieves the proof of the completeness.

III. 3. An invariant distance:

The set of invariants defined in theorem 1, is not convergent. Therefore, it does not induce a distance between shapes. For this reason, we propose to define another complete set of invariants which is convergent in the following theorem:

Theorem 2: The following two sets of complex numbers :

$$I^1_{k_1}(f) = |\det(U_{k_1}, U_{k_0})|,\ I^1_{k_2} = |\det(U_{k_2}, U_{k_0})|$$

$$I^1_k(f) = [\det(U_k, U_{k_0})]^{k_1-k_2}[\det(U_{k_1}, U_{k_0})]^{k_2-k}$$
$$[\det(U_{k_2}, U_{k_0})]^{k-k_1}[\det(U_{k_1}, U_{k_0})]^{k-k_2+p}$$
$$[\det(U_{k_1}, U_{k_0})]^{k_1-k+q}$$

for all $k \in N/\{0, k_1, k_2\}$

$$I^2_{k_1}(f) = |\det(U_{k_1}, U_{k_3}|,\ I^2_{k_2}(f) = |\det(U_{k_2}, U_{k_3}|$$

$$I^2_k(f) = [\det(U_k, U_{k_3})]^{k_1-k_2}[\det(U_{k_1}, U_{k_3})]^{k_2-k}$$
$$[\det(U_{k_2}, U_{k_3})]^{k-k_1}[\det(U_{k_1}, U_{k_3})]^{k-k_2+p}$$
$$[\det(U_{k_1}, U_{k_3})]^{k_1-k+q}$$

for all $k \in N/\{0, k_1, k_2\}$

form a complete and convergent set of invariant features under special affine motion and parametrisations, when p and q are two strictly positive real numbers.

Proof of convergence:

Here, we will show that the invariant set belongs to a given $l^\alpha(Z)$, where $\alpha > 1$:

$$|I^1_k|^\alpha = |\det(U_k, U_{k_0})|^{|\alpha(k_1-k_2)|}|\det(U_{k_1}, U_{k_0})|^{\alpha p}$$
$$|\det(U_{k_1}, U_{k_0})|^q \text{ for all } k \in N/\{0, k_1, k_2\}$$

It is easy to see that :

$$\Sigma|I^1_k|^\alpha = [\det(U_{k_1}, U_{k_0})]^{\alpha p}|\det(U_{k_2}, U_{k_0})|^{\alpha q}$$
$$\Sigma|\det(U_k, U_{k_0}|^{\alpha(k_1-k_2)}$$

In the same way, we can obtain the same result for the second set

$$\Sigma\,|I_k^2|^\alpha = [\det(U_{k_1},U_{k_3})]^{\alpha p}\,|\det(U_{k_2},U_{k_3})\,|^{\alpha q}$$
$$\Sigma\,|\det(U_k,U_{k_3}\,|^{\alpha(k_1-k_2)}$$

By using the well known inequality for the determinant function, we obtain :

$$\|\{I_k\}\|_{l\alpha} \le \|\{I_k^1\}\|_{l\alpha} + \|\{I_k^2\}\|_{l\alpha} \le [\det(U_{k_1},U_{k_0})]^p$$
$$|\det(U_{k_2},U_{k_0})\,|^q \|U_{k_0}\|_2^{(k_1-k_2)}\left(\Sigma\|U_k\|^{\alpha(k_1-k_2)}\right)^{1/\alpha}$$
$$+ [\det(U_{k_1},U_{k_3})]^p\,|\det(U_{k_2},U_{k_3})\,|^q$$
$$\|U_{k_3}\|_2^{(k_1-k_2)}\left(\Sigma\|U_k\|^{\alpha(k_1-k_2)}\right)^{1/\alpha}$$

Which implies :

$$\|\{I_k\}\|_{l\alpha} \le \{[\det(U_{k_1},U_{k_0})]^p\,|\det(U_{k_2},U_{k_0})\,|^q\|U_{k_0}\|_2^{(k_1-k_2)}$$
$$+ [\det(U_{k_1},U_{k_3})]^p\,|\det(U_{k_2},U_{k_3})\,|^q\|U_{k_3}\|_2^{(k_1-k_2)}\}$$
$$\|U_k\|_{l\alpha}^{(k_1-k_2)}.$$

Therefore, we have shown that the set of invariants is convergent and belongs to $l^\alpha(Z)$ when the Fourier sequence also belongs to $l^{\alpha(k_1-k_2)}(Z)$. The Euclidean distance between invariants could be used for the special case of: $\alpha = 2$ and $(k_1 - k_2) = 1$.

Such result is very important in object matching problems and shape classification. So, as we have seen in the general case (section I), the shape space $\mathcal{S}$ becomes a metric space with this following set of metrics :

$$d_\alpha(F, H) = \|\{I_n(f)\} - \{I_n(h)\}\|_{l\alpha} = (\Sigma\,|\,I_n(f) - I_n(h)\,|^\alpha)^{1/\alpha}$$

for any real number α such that $\alpha > 1$ where f and h are two n.a.l.p. of objects having respectively affine shape F and H.

This is another way to obtain a rigorous distance on the affine shape space. But, in this paper we do not prove the stability which guarantees that the distance defined above induces an equivalent topology onto the natural quotient space. However, stability can be shown easily because in practice we compute the discrete Fourier transform (DFT). This involves that both the invariant and the parametrisation spaces become finite dimension vector spaces.

Therefore, a projection theorem for continuous functions is available [37].

IV. PLANAR GRAY LEVEL OBJECTS AND SIMILARITIES

Let us consider now one of the crucial problems in image representation; it concerns the invariant description of planar gray level images under planar similarities.

IV. 1. Group actions and shape space

Following the notations of section I, in the case of planar gray level images with respect to similarity transformations, the group G′ is reduced to the trivial group $\{Id_R\}$, G contains all positive similarities centred on the origin of axes, forms a commutative group isomorphic to the space of polar coordinates :

$G = \{ (r, \theta), r > 0$, and θ belongs to $[0, 2\pi[\ \} = R_+^* \times S^1$

with the well known multiplication

$$(r, \theta), (r', \theta') = (r, r', \theta + \theta')$$

G is a locally compact group. The normalized Haar measure is

$$d\mu(r, \theta) = dr/_r d\theta.$$

The suitable object space is $L_R^2(R_+^* \times S^1, dr/_r d\theta)$, therefore the formulation of the invariance problem can be summarised with the following action :

$$(R_+^* \times S^1) \times L_R^2(R_+^* \times S^1, dr/_r d\theta) \to L_R^2(R_+^* \times S^1, dr/_r d\theta)$$
$$[(\alpha, \beta), \qquad f \qquad] \to f(\alpha r, \theta + \beta).$$

Harmonic analysis in such commutative group has been studied in [7]. The dual of G is the commutative group $R \times Z$. So it is immediate to see that the Fourier transform on G is defined as :

$$(3) \quad \hat{f}(k, v) = M_f(k, v) = \int_0^{+\infty}\int_0^{2\pi} f(r, \theta) e^{-ik\theta} r^{-iv} \frac{dr}{r} d\theta,$$

for $k \in Z$ and $v \in R$.

It represents the Fourier Mellin transform of a two dimensional image $f(r, \theta)$ expressed in polar coordinates. The origin of the polar coordinates can be taken in the image centre of the gravity in order to obtain the invariance under translations. However it is important to note that the integral defined by formula (3) diverges in general. The convergence can be obtained by considering the analytical Fourier Mellin transform [7]

$$M_f(k, s = -\sigma_0 + iv) = \int_0^{+\infty}\int_0^{2\pi} f(r, \theta)\, e^{-ik\theta} r^{-s} \frac{dr}{r} d\theta,$$

for all $\sigma_0 > 0$, $k \in Z$ and $v \in R$.

This integral converges for the positive values of the real part of the complex number s.

From the commutative harmonic analysis, the inverse of the FMT (IFMT) exists and can be written as :

$$f(r, \theta) = \sum_k \int_R M_f(k, \sigma_0 - iv)\, e^{ik\theta} r^{-\sigma_0 + iv} dv.$$

This is important for the definition of a complete set of invariant features and then for the reconstruction of a distance between gray level shapes. The relation between the analytical Fourier Mellin transform (AFMT) of two images f and g having the same shape ($g(r, \theta) = f(\alpha r, \theta + \beta)$), where α is assimilated to a scale factor and β is a rotation parameter, is given by the following system of equations (shift theorem applied to AFMT) :

$M_g(k, s = -\sigma_0 - iv) = \alpha^s e^{ik\beta} M_f(k, s)$

for every integer k, for all real number v and for a fixed positive value σ_0. So, in the following diagram we can see the action on the analytic Fourier space :

$(R^*_+ \times S^1) \times L^2_C(R \times Z, dv \otimes d\delta) \to L^2_C(R \times Z, dv \otimes d\delta)$
$[(\alpha, \beta), M_f] \to \alpha^s e^{ik\beta} M_f(k, s)$
where δ is the discrete measure on Z. This implies that the planar gray level shape space is the quotient space $L^2_C(R \times Z, dv \otimes d\delta)/_{R^*_+ \times S^1}$
Unfortunately, the distance D defined in section I, does not exist because G is not a compact set. In the following paragraph, we construct a complete and convergent set which allows us to define a metric between shapes.

IV. 2. A complete and convergent set of invariant primitives

Theorem 3: When we assume that $\{I_f(k, s)\}$ is a convergent set then
$I_f(k, s) = M_f(k, s)\,[M_f(1, 1)]^{-k}\,|M_f(1, 1)|^k\,[M_f(0, 1)]^{-s}$
is a complete and convergent set of invariant features under planar similarities.

Proof of convergence: Invariance and completeness have been shown in [7]. For such a set, we only recall here the proof of convergence of this set. The p-norm of the sequence of invariants can be expressed as:
$$N_p(I_f) = \Big(\sum_{k \in Z} \int_{-\infty}^{+\infty} |I_f(k, -\sigma_0 + iv|^p dv\Big)^{1/p} =$$
$$[M_f(0, 1)]^{-\sigma_0}\Big(\sum_{k \in Z} \int_{-\infty}^{+8} |M_f(k, -\sigma_0 + iv)|^p dv\Big)^{1/p}$$
since:
$$|I_f(k, \sigma_0 - iv| = |M_f(k, s)|[M_f(0, 1)]^{-\sigma_0}$$
This implies that:
$N_p(I_f) = [M_f(0, 1)]^{-\sigma_0} N_p(I_f)$ when $M_f(0, 1)$ is not zero.
The set of invariant features is convergent since σ_0 is a fixed number.

IV. 3. An invariant metric

The set of invariants $\{I_f(k, s)\}$ is convergent since it can be assumed to belong to $L^2(Z \times R, \mu)$ (finite energy of the image signal). So;
$$N_2(f) = \Big(\int_{-\infty}^{+8} \sum_{k \in Z} |I_f(k, s = -\sigma_0 + iv|^2 dv)\Big)^{1/2} < +\infty$$
It implies that the shape space $\mathcal{S}$ is a metric space with the following distance:
$$d_2(G, H) = \Big(\int_{-\infty}^{+8} \sum_{k \in Z} |I_g(k, -\sigma_0 + iv) - I_h(k, -\sigma_0 + iv)|^2 dv)\Big)^{1/2}$$
where g and h are two gray level images having respectively the shape G and H.

V. THREE DIMENSIONAL OBJECT AND EUCLIDEAN MOTIONS

Recently, there has been renewed interest in closer studies of 3D-geometric shapes. Perhaps the reason is that there is more 3D available, either as range data (obtained by stereo measurement) or as a complete 3D lattice data (obtained by computerised tomography). There are two aspects of shape due to the type of the image data: the surface aspect and the volume aspect. In this section, we shall describe how our unitary approach introduced in section I, allows the extraction of a relevant invariant 3D descriptors in the two aspects since the formulation adjusts itself to the type of the considered 3D object.

V. 1. Descriptors for three-dimensional surfaces

Three-dimensional feature extraction of 3D surfaces is a complex problem, particularly if all Euclidean three dimensional motion invariance has to be required. One of the main reasons is that there does not exist a general three dimensional surface parametrisation. So, the classification of surface parametrisations is not well known.

However, most of the 3D surfaces usually studied in pattern recognition are closed and connected, which could be represented by a spherical or a periodic bidimensional parametrisation. harmonic analysis gives an efficient solution for these categories of objects when suitable reparametrisation is computed. So in this paragraph, we intend to show how to construct a global set of invariant descriptors under a general three dimensional Euclidean motion in these two kinds of surfaces.

V. 1. 1. Spherical shaped form

Definition 19: A spherical parametrisation of a three dimensional surface is an immersion from S^2 to R^3.

Definition 20: A closed 3D surface is said to be a spherical shaped form if and only if it has a spherical parametrisation (see Fig. 1).

Following the notation of section I, G′ and G represent respectively the 3D Euclidean motion group M(3) and the unit sphere S^2. This means that the invariant representations of spherical shaped forms need the application of harmonic analysis on S^2 which is a topological compact group. Its well-known unique Haar measure can be written as:
$$d\mu(\theta, \phi) = {}_{1/4\pi}\sin\theta\, d\theta\, d\phi.$$

Compactness implies that the Fourier transform exists and it is discrete. It corresponds to the Fourier coefficients with respect to a given basis of $L^2(S^2)$. For computational simplicity, we propose to consider the classical basis which could be expressed according only to the Legendre associated functions (see Appendix I). Let us denote here such basis by:
$$\{e_{n,m}(\theta, \phi), -m < n < m, m \text{ belongs to } Z\}$$

Spherical shaped objects can be represented by an orthogonal spherical arclength parametrisation obtained after a reparametrisation procedure. when it is possible.

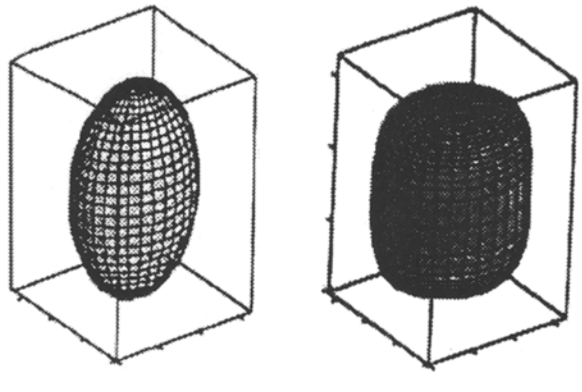


Fig. 1. Examples of spherical shaped forms (superquadrics)

Exemples de formes sphériques

Such representation can be assumed to belong to $L^2{}_R 3(S^2)$. So, the invariance formulation of this case can be summarised by the following action :

$$(R^3 \times SO(3) \times S^2) \times L^2{}_R 3(S^2) \to L^2_R 3(S^2)$$
$$[(B, A, (\varphi_0, \psi_0)), f] \to A f(\varphi + \varphi_0, \psi + \psi_0) + B.$$

By computing Fourier coefficients of an orthogonal arclength parametrisation, the Shift theorem reduces the last action to only the following left operation :

$$(R^3 \times SO(3) \times S^2) \times l^2_C 3(Z \times Z) \to l^2_C 3(Z \times Z)$$
$$[(B, A, (\varphi_0, \psi_0)), \{a_{n,m}(f)\}] \to$$
$$e_{n,m}(\varphi_0, \psi_0) A\, a_{n,m}(f) + B\delta_{n,m}$$

Using this result, relevant invariant descriptors under 3D Euclideean motions can be constructed. We propose the following set :

$$I_{n,m}(f) = \sum\sum |a_{n,m}(f)|$$

V. 1. 2. Torus shaped form

Now we intend to consider another kind of closed surfaces which we call a *Torus shaped form.*

Definition 21: A bidimensional periodic parametrisation is an immersion from the bidimensional Torus T_2 into R^3.

$$f: T_2 \sim S^1 \times S^1 \to R^3$$
$$(\theta, \phi) \to t\,(x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))$$

Such parametrisation represents the Torus shaped form. The formulation in this case can also be realised by the determination of the diagrams which summarise the group actions. G′ represents in this case the 3D Euclidean motion group M(3) and G is T_2. As we have seen above an orthogonal and normalised arclength parametrisation is required in order to transform the right action into a translation one. All of these parametrisations belong to the same space $L^2{}_R 3(S^2)$. So, we have the following operation :

$$(R^3 \times SO(3) \times T_2) \times L^2{}_R 3(T_2) \to L^2_R 3(T_2)$$
$$[(B, A, (\theta_0, \phi_0)), f] \to A f(\theta + \theta_0, \phi + \phi_0) + B.$$

As T_2 is a compact set, which can be seen as the product group $S^1 \times S^1$, its normalized Haar Measure is the product of that of S^1. It can be expressed as :

$$d\mu(\theta, \phi) = d\theta\, d\phi$$

The compactness implies that the dual group is discrete and is isomorphic to Z × Z. All their irreducible and unitary representations are :

$T_{nm}(\theta, \phi) = exp -[\mathbf{i}(n\theta + m\phi)]$ *for n and m any integers*

Then, the Fourier transform on this group is the bidimensional Fourier series. All periodic bidimensional parametrisations can be expanded in Fourier series. The Shift theorem applied on an orthogonal and normalised arclength parametrisation of such surfaces implies that the operation, in the Fourier transform space, is reduced to :

$$(R^3 \times SO(3) \times T_2) \times l^2_C 3(Z \times Z) \to l^2_C 3(Z \times Z)$$
$$[(B, A, (\theta_0, \phi_0)), \{a_{n,m}(f)\}] \to$$
$$e^{i(n\theta_0 + m\phi_0)} A\, a_{n,m}(f) + B\delta_{n,m}$$

then, the extraction of invariant features becomes easier. The immediate solution which verifies stability but does not fulfil completeness consists of the following set :

$$I_{n,m}(f) = \sum_{(n,m)\, Z \times Z / \{(0,0)\}} \|a_{n,m}(f)\|^2$$

A set of invariants under three dimensional affine transformations can be defined by using the determinant operator, a case which we do not study here.

V. 2. Descriptors of three dimensional gray level objects

Such a description is very useful for medical application like the description of ulna or radius bones anatomical structures [19]. This application requires three dimensional invariant descriptors under 3D Euclidean motion transformations. The group representation theory is suitable. The formulation allows us to define a relevant set of invariant primitives. So, the computation of the irreductible and unitary representations of M(3) is the first step.

V. 2. 1. Harmonic analysis on M(3)

A 3D gray level image is considered as a function f defined in R^3 which is assumed to have its support confined in a bounded domain of M(3). This group acts on the representation space $L^2(M(3))$ as follow :

$$M(3) = R^3 \times SO(3) \times L^2_R(M(3)) \to L^2_R 3\,(M(3))$$
$$g = (B, A, f] \to f \circ g$$

The motion of the three-dimensional Euclidean space is a cross product of the three-dimensional translation space T(3) and the group of the three dimensional Euclidean rotations SO(3). A general rotation h can be determined uniquely by the three Euler angles (ϕ, θ, ψ). h is the product of a rotation by the angle φ around the axis Ox, a rotation by the angle θ around the axis Oy, and a rotation by the angle ψ around the axis Oz. The unique normalised Haar measure of M(3) is given by :

$$d\mu(h, x, y, z) = dh\, dx\, dy\, dz$$

where dh is the normalized Haar measure of SO(3)

$$dh(\phi, \theta, \psi) = \frac{1}{16\pi^2}\sin\theta\, d\theta\, d\varphi\, d\psi.$$

Here, we do not consider all irreductible and unitary representations of this group, but only a class of representations, since the other irreductible and unitary representations are not known.

However, the subset of all representations defines an operator which allows the construction of a relevant set of invariant features since the shift theorem is still valid.

Definition 22: A representation T(g) is called a representation of class 1 relative to H if in its space there are non zero vectors invariant with respect to H

For every element g of M(3) corresponds a representation $T_\lambda(g)$ acting on $L^2(S^2)$, which turns $\rho(\xi)$ into the function

$$T_\lambda(g)\,\rho(\xi) = e^{i\lambda <a,\,\xi>}\,\rho(h^{-1}\xi)\ \forall\ \lambda \in \mathbb{R}_+^* \text{ and } \xi \in S^2,$$

where : $<a, \xi>$ is the inner product of R^3. These irreductible representations of class 1, allow to define an operator which verifies the shift theorem as it was defined in (I. 3. 3 : definition 9).

Following the notation of section I, for every f belonging to $L^2(M(3)$, the pseudo-Fourier transform on M(3) can be expressed as

$$[\hat{f}(\lambda),\rho(\xi) = \int_{M_3} f(x, y, z)\, e^{-i\lambda <(x,y,z)^t,\, h(\xi)>}\,\rho(h(\xi))\, d\mu(h, x, y, z)$$

In the case of three-dimensional gray level images, this transform can be expanded into the Fourier series transform with respect to the basis of $L^2(S^2)$, which is written according to the associated Legendre functions, and its components are reduced to the following formula :

$$[\tilde{f}_{1,m}(\lambda),\rho]\,(\xi) = c(1, m)\int_0^{2\pi}\int_0^{\pi}\hat{f}(\lambda\sin\theta\sin\varphi, \lambda\sin\theta\cos\varphi)\sin\theta\, P_1^m(\cos\theta)e^{+im\varphi}d\varphi\, d\theta$$

where $\hat{f}$ represents the Fourier transform on the additive group R^3 (the classical three-dimensional Fourier transform) and $c\,(1, m)$ is a constant.

The proof is given in Appendix I.

Remark: When f is a 3D gray level image, then it can be considered as a function in M(3) independent with respect to (ϕ, θ, ψ).

Proposition 8: The function I_f defined on R_+^* by:

$$I_f(\lambda) = \sum_{1=0}^{+\infty}\sum_{m=-1}^{m=+1}\tilde{f}_{1,m}(\lambda)\,\overline{\tilde{f}}_{1,m}(\lambda)$$

forms a set of invariants for 3D *gray level object under a general* 3D *Euclidean motion.*

$\overline{\tilde{f}}_{1,m}(\lambda)$ denotes the conjugate of the Fourier component $\tilde{f}_{1,m}(\lambda)$ of the function f.

It is easy to show that if f is in $L^2(M(3))$, then $I_f(\lambda)$ converges for all λ. These descriptors can be used for 3D gray level images indextion, for example in medical volumes images data bases.

VI. APPLICATION TO OBJECT ORIENTED IMAGE CODING

The ITU-T has standardised several block-based hybrid coder for coding of moving images where each image of a sequence is subdivided into independently moving blocks. Each block is coded by 2D-motion compensed prediction and transform coding (DCT). This corresponds to a source model of 2D square blocks moving with a translation in the image plane, which fails at boundaries of natural moving objects. At low data rates, source model causes coding artefacts known as blocking and mosquito effects.

In order to avoid these coding distortions, the concept of object based coding has been introduced [47,48]. A coder based in this approach divides an image sequence into moving objects. We propose here to describe objects by the global rigid movement of exterior contour (a translation followed by a rotation and a scale factor), by its shape (invariant descriptors under similarities) and by its texture. We intend to apply the formulation developed in section II to a videophone sequence.

The object motion analysis is recognised as a key point for many applications in robotic vision and machine intelligence or, recently, in object oriented coding. It consists of the estimation of the geometric transformation parameters existing between a given object extracted from two consecutive images a sequence. It is usually based on primitives that may be, according to the context, segments of straight lines or curves, characteristic points, pixels blocks or regions. Matching the primitives is carried out by minimising a cost function, or maximising a correlation function, both based on these attributes. It is important to note that that efficiency of an estimation method depends on the primitives that are used criterion which could be the generalised correlation function or a given distance between functions. A numerical optimisation method is generally required. In this section, we intend to use the exterior profile of the objects as features. We also intend to consider the Euclidean distance between respective parametrisation of each planar and closed boundary, assuming that these objects move with a planar and rigid motion. Other hypothesis can be done that will not be developed in this article as :

– planar objects submitted to three dimensional motions which can be described by affine transformations,
– planar grey level objects submitted to rigid movements modelled by planar Euclidean motions
– three-dimensional objects (surfaces or volumes) which move in the 3D space.

The joint topology and harmonic analysis proposed in this paper seem very suitable to give answers to object oriented coding since such a model can be adapted to the different cases which we classify above.

Later on, we show how the whole theoretical results in the contour-rigid case serve efficiently and are essential in oriented coding application.

For these experiments, the test sequence Miss Claire has been used. It has been selected by international expert groups. Figure 2 shows some pictures of this test sequence. The first pre-processing consists in extracting the contours as shown in Figure 3.

FIG. 2. Original Claire sequence

Séquence originale de Claire

FIG. 3. Contour extraction

Extraction de contour

The main idea consists of the compensation and the coding of homogenous regions assumed to be animated with rigid movements between two consecutive images. However, we admit of small non linear shape deformations

The formulation described in Section II seems suitable for the coding problems. Effectively, data reduction for very low bit rate image coding can be obtained for the following reasons :

1. All pixels of a region are assumed to have the same motion which is related directly to the parameters of the geometrical transformation. For planar similarities, we only have four real parameters.
2. The whole exterior boundary of a given region can be coded with a limited number of descriptors (about 20 features for a perimeter of nearly 400 points).
3. Texture of regions is also described with a limited number of features (this point will not be studied in this paper).

A prediction procedure will be applied to all these primitives which generally change slowly along image sequences.

The first type of features which is the motion parameters requires the follow properties :

1. *Robustness*
2. *Uniqueness*: Hausdorf distance which is a shape space metric as we have seen in section I, guarantees this fact.
3. *Real time* computation : only a limited number of the normalised points will be useful for the estimation of such primitives.

The second kind of features which are the shape descriptors have to verify the following criteria :

1. *Invariance*: The descriptors must be independent of a rigid motion (a planar similarity).
2. *Completeness* allows the reconstruction of the object up to a transformation in the receiver.
3. *Stability* gives robustness under small distortions caused by failures in transmission, quantization and non linear deformation of the object between two consecutive images.
4. *Simplicity* and *real time* computation (based on a monodimensional fast Fourier transform algorithm).
5. Generally, a given scene contains several objects. Therefore, we usually need to establish the correspondence between similar objects in the two consecutive images before the movement estimation step. This fact requires the computation of a *shape distance*. Such a distance presents algorithm simplicity, allowing the matching between objects in real time.

VI. I. I Discrete arclength reparametrisation

A curve is usually represented by a parametrisation. It is well known that there are different curve parametrisations to represent a given curve. Therefore, the normalised arclength parametrisation has to be used when the invariance under similarities is required and when the displacement of the object is assumed to be rigid and planar, such a parametrisation is also needed for the estimation of the motion (see Fig. 4). An arclength parametrisation is obtained by a reparametrisation procedure which does not have to depend on the location. We propose to use the truncated Fourier expansion of the original curve points.

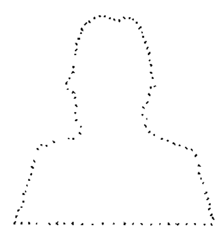


FIG. 4. A discrete arclength normalisation with 128 points.

Normalisation de longueur d'arc discrète à 128 points.

Such curve algorithm can be expressed as :

$$f(t) = \frac{1}{n}\sum_{m=0}^{n-1} r_m B_m(t, m)$$

where

$$B_m(t, m) = \frac{1}{M}\sum_{k=0}^{M-1} e^{-\frac{2i\pi mk(m-t)}{n}}, \quad n << M$$

where M is the number of the original contour points denoted by r_m and n is the truncated number of the Fourier harmonics. It is easy to verify that this curve algorithm is periodic and independent of the orientation of axes in the mean of [19, 20].

So, a discrete normalisation can be achieved by computing the length function defined by:

$$l(t) = \int_0^t \|f'(u)\|_2 \, du$$

$$= \frac{2\pi}{Mn^2} \int_0^t \Big\| \sum_{m=0}^{n-1} \sum_{k=0}^{M-1} k r_m \, e^{-\frac{2i\pi mk(m-t)}{n}} \Big\| du$$

Then, a uniform sampling of the inverse function of the arclength one defined in the last formula can be derived numerically. In Figure 4, we illustrate the type of experimental result which we have obtained in a Claire contour with 128 normalised points.

It is important to note that this step of normalisation is obtained by using twice the FFT algorithm.

VI. 1. 2. Invariant features extraction

The computation of invariants is obtained in real time, since the DFT is computed on the obtained normalised points of the contour.

Figures 5.a, 5.b illustrate the relevance of these descriptors. Clearly, they converge to zero when indexes tend to infinity. Therefore, the shape information is located near the origin and it is contained in just a few invariants. A contour reconstruction from a limited number of Fourier coefficients was proved in [4, 6] to be without a lot of shape deformation.

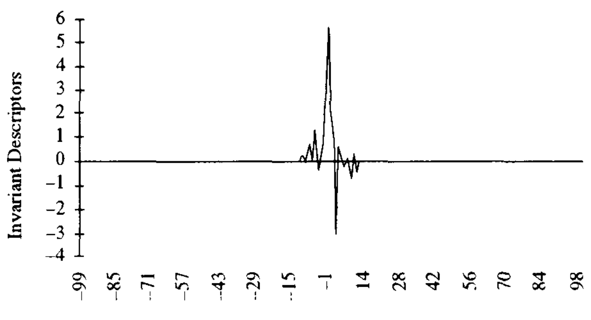


FIG. 5.a - Real part of complete and stable invariant descriptors

Partie réelle de descripteurs invariants stables et complets

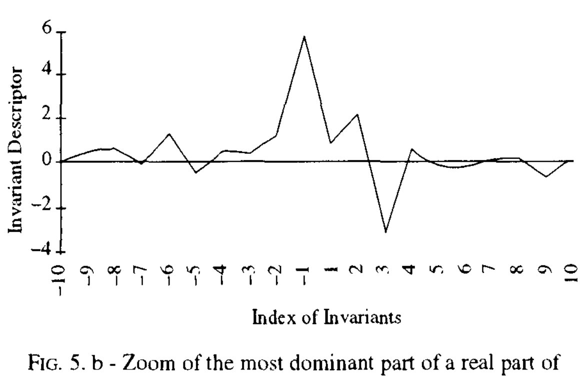


FIG. 5. b - Zoom of the most dominant part of a real part of invariant descriptors

Zoom de la partie la plus dominante d'une partie réelle de descripteurs invariants

VI. 2 Global motion estimation of objects

VI. 2. 1. Hausdorf distance and the uniqueness of the motion

The following step is the movement estimation between two objects having the same shape and extracted from two consecutive images.

Let f and h be respestively two normalised arclength parametrisations of two object having respectively the shapes F and H. As we have shown in Section II that the following quantity is a metric between shapes:

$$d(F, H) = \inf_{(l_0, \theta) \in T_2} \| f(1) - e^{i\theta} h(l + l_0) \|,$$

In the Fourier domain, and by using Shift theorem, the computation of such distance is reduced to the minimisation of the following expression:

$$\sum_{n=-\infty}^{n=+\infty} | a_n(f) - e^{j(nl_0+\theta)}, a_n(h)|^2$$

Persoon and Fu showed that there is a simple numerical solution to this problem, since it consists of the extraction of all the zeros of the function F given by:

$$F(l_0) = \sum_n \rho_n \sin(\psi_n + nl_0) \sum_n \rho_n \cos(\psi_n + nl_0) -$$

$$\sum_n \rho_n \cos(\psi_n + nl_0) \sum_n \rho_n \sin(\psi_n + nl_0)$$

$$\text{where } \rho_n e^{l\psi_n} = a_n(f) * a_n(h)|^2$$

The optimal value of θ that represents the rotation parameter is obtained by the next formula:

$$\tan\theta = - \frac{\sum \rho_n \sin(\psi_n + nl_0)}{\sum \rho_n \cos(\psi_n + nl_0)}$$

where l_0 is the solution to the equation $F(l_0) = 0$ which minimises the square error among all the set of solutions.

VI. 2. 2. Motion estimation

In practice, the scale factor α can be obtained by minimising the following quantity:

$$\sum_{n=-v}^{n-v} | a_n(f) - \alpha e^{j(nl_0+\theta)}, a_n(h)|^2$$

This is possible, because we can assume that the scale factor values representing the zoom between two curves of the same object belonging to two consecutive images, is bounded.

After the numerical resolution of the equation $F(l_0) = 0$ and the computation of θ, the scale factor α can be reduced by using the following formula:

$$\alpha = \frac{\sum_n \rho_n \cos(\psi_n + nl_0 + \theta)}{\sum_n a_n(f)\, a_n(h)}$$

Example: In order to illustrate the behaviour of this motion estimation algorithm, we consider the contour represented in Figure 6. This contour was rotated artificially by 10 degrees in the inverse clockwise direction (Fig. 6.).

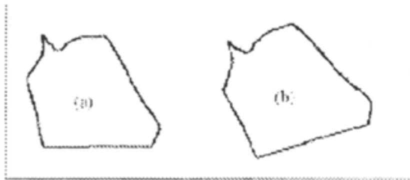


FIG. 6. Original contour 1. (b) Original contour 2.

(a) contour original 1, (b) contour original 2

$F(l_0)$ only contains a limited number of harmonics. So, it is possible to find all zeros of $F(l_0)$ by using numerical techniques. Here, we apply a bracketing method to solve this equation. The function $F(l_0)$ has more than one root in the interval $[0,2\pi]$ (for the two contours represented in Figure 6, the function $F(l_0)$ presents several roots, see Fig. 7).

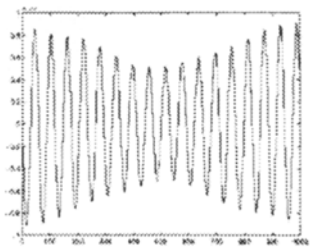


FIG. 7. The function $F(l_0)$ computed with 64 values of normalised points

Fonction $F(l_0)$ calculée avec 64 valeurs de points normalisés

As $F(l_0)$ exhibits several zeros, we compute the Hausdorf distance for each solution. Thus, the zero which presents the smallest value of Hausdorf is kept. The adjustment of the parameter ν, which represents the number of normalised points, can be reduced significantly, as we can see in Figure 8.a and Figure 8.b. This implies that a reduction in the computing times can be obtained by choosing the minimum number of the normalised points.

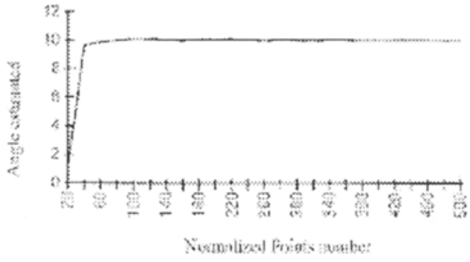


FIG. 8.a. The rotation angle estimation with different number of normals points.

Estimation de l'angle de rotation avec des nombres points normalisés differents

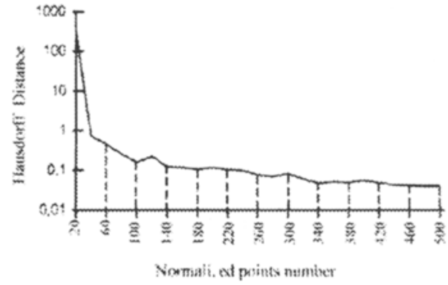


FIG. 8.b. Hausdorff distance according to the number of normalised points.

Distance de Hausdorff en fonction du nombre de points normalisés

Figure 9 shows another Claire sequence (n° 2) having a modification in the shape of the profile object for the two last contours.

FIG. 9. Claire contour sequences n° 2.

Séquences de contour de Claire n° 2

The estimated value of the Hausdorff distance between contours (i and $i + 1$) of the Claire contour sequence represented in Figure 10, shows that this quantity follows the contour shape evolution. In fact, we can observe that the value of this distance increases considerably between the contours 10 and 11. So, it corresponds to our interpretation of shapes.

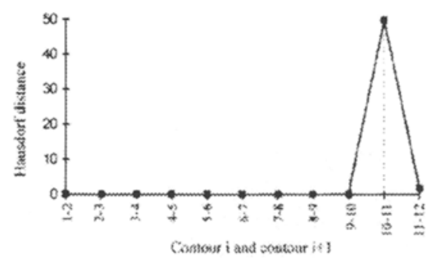


FIG. 10. Hausdorff distance.

Distance de Hausdorff

In order to reconstruct the regions in the receiver, we have to transmit the prediction error of each invariant and each motion parameter. After motion estimation, about 20 invariants descriptors are used for this videophone application for each object. Figure 11.a and Figure 11.b represent the variation of such features along the sequences. As we can observe, these values do not change much for Claire sequence. However, there is an important variation between contour 10 and contour 11. Such experiment result proves that transmission of the prediction error on invariants could be done with a very low bit rate without losing a lot of information.

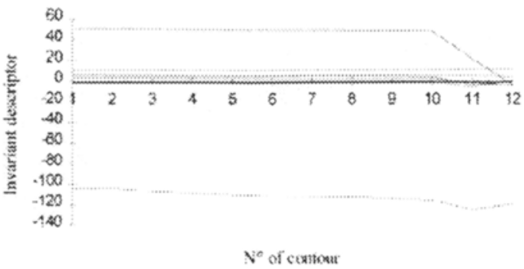


FIG. 11.a. Real part of invariant descriptors.

Partie réelle de descripteurs invariants

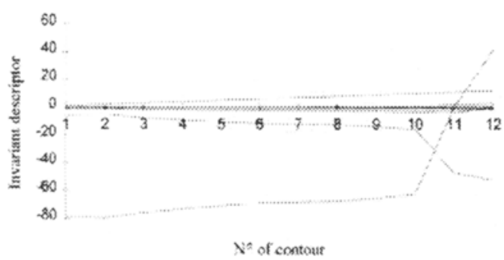


FIG. 11.b. Imaginary part of Invariant descriptors.

Partie imaginaire de descripteurs invariants.

CONCLUSION

The mathematical formulation presented in this paper has been applied to coding, in the planar closed contour case, with a rigid motion (Euclidean motion). The generalisation to the affine case would be very useful for such an application and it would be more efficient than Euclidean case. It can also model the 3D dimensional planar contour displacement. The disadvantage comes from the fact that the parameter estimation is not unique since SL*(2,*R*)* is not compact. The grey level case can be solved, too. Our future works concern the planar grey level object compensation with a rigid movement assumption and the planar affine animated with a general rigid 3D motion. Three-dimensional grey level object movement estimation can also be treated with such an approach. However, completeness is not verified for the moment, which is important for the coding applications since object reconstruction is not possible. The 3D surface object case is a more complex problem because of the non existing of a parametrisation as we have shown in section IV.

The more complex movement estimation problem remains projective 3D grey level object one. In this case, the Harmonic analysis seems suitable to formulate the problem but the theoritical solution needs the determination of the unitary and irreducible representation of the group SL*(2,*R) which is not to be done.

Acknowledgements

The author thanks Mr. Olivier Avaro for his valuable idea about the application of the geometric invariance approach to the coding field, also thanks Mr. Gérard Eude, for his encouragements. The author also thanks the CNET-DSE *and the* CCETT-Rennes *for having financially supported this work.*

Manuscrit reçu le 22 décembre 1997

Appendix I : A basis of $L^2(S_2)$

The well known basis of the $L^2(S^2)$ is defined as

$\{e_{m,l}(\theta,\phi); -1 \leq m \leq 1, 1 \in \mathrm{IN} \text{ and } \xi = (\theta,\phi) \in S^2\}$

where $e_{m,l}(\xi)$ are functions which are defined in spherical S^2 by :

$$e_{m,l}(\xi=\theta,\phi) = \sqrt{\frac{(2l+l)(l+m)!}{(l+m)}}\, P_l^m(\cos\theta)\, e^{-im\phi},$$

where $P_l^m(\cos\theta)$ are associated Legendre functions :

$$P_l^m(z) = \frac{(-1)^{m+1}(1-z^2)^{m/2} d^{m+1}}{2^l l!} \frac{d^{m+1}}{dz^{m+1}}(1-z^2)l$$

where l and m are integers ($m \leq 0$). In case where is :

$$P_l^m(z) = (-1)^m \frac{(1-m)!}{(1+m)!} P_l^m(z).$$

Appendix II

Let f be a function in $L^2_{\mathbb{C}}(M(3))$ which represents a 3D grey level object. Then we can assume that it is independent of the parameters of the three rotation parameters ϕ, θ, ψ.

By making the charge of variables in \$ which we denote by :

$$\xi' = h(\xi)$$

the proposed operator can be written as follow

$$[\tilde{f}(\lambda)\rho)](\xi) = \int_{M(3)} f(x,y,z) e^{-i\lambda <(x,y,z)^t,\xi'>} \rho(\xi') d\xi' dx dy dz.$$

Then, the function $\tilde{f}(\lambda)$ associated to ρ is a scalar one. This function which we call the pseudo-Fourier transform of f comes from the following intergral.

$$\forall \xi \in S^2 \text{ and } \rho \in L^2(S^2)$$

$$[\tilde{f}(\lambda)\rho)](\xi) = \int_{M(3)} f(x,y,z)[T_\lambda^{-1}(h,(x,y,z)^t)\rho](\xi) dh dx dy dz$$

The linear forms $\{\tilde{f}(\lambda), \lambda > 0\}$ defined in space $L^2(S^2)$ can be written according to its compotents in the basis B (defined in Appendix I) in the following mean :

$$\tilde{f}_{(l,m)}(\lambda) = K_{lm} \int_0^{2\pi} \int_0^{\pi} \int_{IR^3} f(x,y,z) e^{-i\lambda < x \sin\theta \sin\varphi + y \sin\varphi \cos\varphi + z \cos\varphi >} P_l^m(\cos\theta) e^{+im\varphi} \sin\theta \, d\varphi \, d\theta \, dx dy dz,$$

$$= K_{l,m} \int_0^{2\pi} \int_0^{\pi} \hat{f}(\lambda \sin\theta \sin\varphi, \lambda \sin\theta \cos\varphi, \lambda \cos\theta) \sin\theta \, P_l^m(\cos\theta) e^{+im\varphi} d\varphi \, d\theta$$

where $\hat{f}$ represent the usual Fourier transform and

$$K_{l,m} = \frac{1}{2\pi} \frac{(2l+1)(l-m)!}{(l+m)}$$

REFERENCES

[1] MING-KUEL (H.), Visual pattern recognition by moment invariants, *IEEE Trans. IT.*, pp 179-187, (Feb 1962).

[2] SADJADI (F. A.) et al , Three dimensional moment invariants, *IEEE Trans PAMI, 2*, pp. 127-136. (March 1980).

[3] CHEONG-HUAH et al, 3-D moment forms : their construction and application to object identification and position, *IEEE Trans. PAMI* **11**, no. 10. (Oct. 1989).

[4] ZHAN (C. T.) ROSKIES (R. Z.), Fournier Description for plane closed curves, in *IEEE Trans. C,* **21**, n° 3, pp. 261-81, (1972).

[5] GRUNLUND (G. H.), Fourrier pre processing for hand print character recgnition, *IEEE. Trans. C,* **21**, pp. 195-201. (1972).

[6] PERSOON (E.) FU (K. S.), Shape discrimination using Fourrier descriptors, *IEEE Trans. SMC,7*, n° 3, pp. 170-179. (1977).

[7] GHORBEL (F.), A complete invariant description for gray-level images by the harmonic analysis approach, *Pattern Recognition Letters*, pp. 1043-1051. (15 April 1994).

[8] BOUGRENET (J. L. de), GHORBEL (F.), Scale invariant pattern recognition applied to image data compresion, *Pattern recognition letters*, **8**, pp. 55-58. (1988).

[9] FONGA (H.), Analyse harmonique en reconnaissanse de formes Doctorate thesis, *Institut National Polytechnique de Grenoble* (France). (1992).

[10] TIMOTY (P.). Wallace,Three dimensional shape analysis using local shape description, *IEEE Trans. PAMI,* **3**, (May, 1981) pp. 310-322.

[11] BURDIN (V), Modélisation et analyse de structure tridimensionnelles ; application à la morphométrie des os longs, *Doctorate thesis*, Université Rennes I. (1992).

[12] WEISS (I), Noise resistant invariant of curves. *Geometric in Computer Vision, edited by* Joseph L. Mundly and Andrew Zisserman, Artificial Intelligence series, *MIT Press*, (1992).

[13] GOURD (F.). GAUTIER (J. P.), Une méthode d'analyse harmonique en reconnaissance de formes, *Traitement du signal,* **6**, no3, pp. 161-178. (1989).

[14] GHORBEL (F.), Vers une approche mathématique unifiée des aspects geométriques et reconnaissance de formes planes, *Doctorate thesis*, Université Rennes I. (1990).

[15] GAUTIER (J. P.), BORNARD (G), SILBERMANN (M). Motion and Pattern analysis : harmonic analysis on motion groups and their homogenoeus spaces, *IEEE Trans*. SMC **21**, no 1 (Jan/Feb 1991).

[16] Lenz-Groups representations and their homogeneous spaces, *Lecture Notes in Computer Science, Springer*. (1991).

[17] GHORBEL (F.), Stability of invariants Fourrier descriptors and its interference on the shape classification, *11th International Conference on Pattern Recognition, (ICPR)*, 30 August-3 September, The Hague (NL). (1992).

[18] GHORBEL (F.) (1992), Distance et invariance pour la clasification d'objets dans les images, *International meeting on distance analysis DISTANCIA* ' 92, June 22-26, 1992, Rennes.

[19] GHORBEL (F.). BUDIN (V.), Invariant curve approximation vith radial representation, Wavelets, Images, and Surface fitting, Pierre-Jean Laurent/Alain Le Méhauté/ Larry L. Schumaker (eds). 269-276. *AK Peters Ltd*, Wellesley USA(1994).

[20] BEZ (H. E.), A unified of invariance for the curve and surface defining algorithms, *IMAJ. Appl Maths*, **30**, pp. 325-341. (1983).

[21] BEZ (H. E.), On invariant curve forms, *Computer Geometrical Design*, **3**, pp. 193-204. (1986).

[22] PAVLIDIS (T.). Structural recognition, *Springer Verlag*, Berlin. (1977).

[23] DIEUDONNÉ (J.). *Élément d'analyse moderne, Gautier-Villars*. (1974).

[24] ARBTER (K.), SNYDER (W. E.), BURKHARDT (H.). HIRZINGER (G.), Application of affine-invaraint Fourrier Descriptiors to Recognition of 3-D objects, *IEEE Trans., PAMI*-**12**, no. 7, pp. 640-647. (July 1990).

[25] GHORBEL (F.), BOUGRENET DE LA TOCNAYE (J. L. de), Automatic control of lamellibranch larva growth using contour invariant feature extraction, *Pattern Recognition*, **23**, no. 2. (March 1990).

[26] SEGMAN (Jos), RUBINSTEIN (J), ZEEVI (Y. Y.), The canonical coordinates method for pattern deformation : theoretical and computational considerations, *IEEE Trans PAMI,* **14**, 12, pp. 1171-1183. (1992).

[27] SEGMAN (Jos) ZEEVI (Y. Y.), Image analysis by wavelet-type transforms : group theoric approach, *Mathematical Imaging and vision*, (1993).

[28] MEYER (F.) et al, Morphological segmentation, *J. Visual Communication and Image Representation*, **1** no. 1, pp. 21-46, (Sept. 1990).

[29] SALEMBIER (P.) et al, Morphological multiscale image segmentation, SPIE vol. 18, *Visual Communications and Image Processing*, pp. 620-627 (1992).

[30] SERRA (J.), Image analysis and mathetical morphology, Vol II, *Theoretical Advances, Academic Press.* (1988).

[31] FREEMAN (H.), Shape description via use of critical points, *Pattern recognition,* 10, pp. 159-156, (1978).

[32] VILLENKIN, Fonctions spéciales et théorie de la représentation des groupes, *Dunod*, Paris. (1969).

[33] DIEUDONNÉ (J.) CARRELL J., Invariant theory old and new, *Academic press*, New York. (1971).

[34] CRIMMINS (T. R.), A complete set of Fourrier descriptors for two-dimensional shapes, *IEEE Trans., SMC*-**12**, no. 6, pp. 848-55, (Nov/Dec. 1982).

[35] YIP (R. K. K.), TAM (P. K. S.), LEUNG (D. N. K.), Application of elliptic Fourier descriptors to symmetry detection under parallel projection, *IEEE, Trans. PAMI,* **16** no. 3. (March 1994).

[36] SEKITA (I.), KURITA (T.), OSTSU (N.), Complex autoregressive model for shape recognition, *IEEE Trans., PAMI,* **14**, pp. 489-496, (April. 1992).

[37] DIEUDONNÉ (J.), Eléments d'analyse *Gauthier-Villars* Tome II, (Collection cahier scientifique), Paris I. (1974).

[38] DIEUDONNÉ (J.), *Eléments d'analyse, Gauthier-Villars* Tome VI, (Collection cahier scientifique), Paris I. (1974).

[39] SHENG (Y), Processeur optique en temps réel d'image video : application optique des moments bidimentionneles, *Doctorate thesis, Faculté des sciences et Techniques*, Université de Franche-comté. (1986).

[40] CHEN (S.), A new vision system and the Fourier descriptors method by group representations theory, *CDC Conference*, Las Vegas, USA. (1985).

[41] SADJADI (F. A.) et al. Three dimensional moment invariants, *IEEE Trans PAMI.*, **2**, pp. 127-136. (March 1980).

[42] CHONG-HUAH et al, 3-D moment forms : their construction and application to object identification and positioning, *IEEE Trans PAMI*.**11**, no 10. (Oct. 1989).

[43] YOU (Y. C.) et al (1979), A syntactic approach to shape recognition using attributed grammars, *IEEE Trans PAMI*. SMC, **9**, pp. 334-345, (June1973).

[44] BARR, (A. H.) Superquadrics and angle-preserving transformations IEEE Trans *CGA.*, **1**, pp. 11-23. (1986).

[45] SOLINA (F) BAJCSY (R), Recovery of parametric models from range images : the case for superquadricks with global deformations, *IEEE trans PAMI.*, **12**, no 2, pp 131-47, (1990).

[46] ZRIBI (M.), FONGA(H.) GHORBEL (F), Invariant description for 3D gray level volumes under a general 3D Euclidean motion, *Internationat Conference on Pattern Recognition ICPR*, Austria. (1996).

[47] (J) OSTERMAN, Based analysis-synthesis coding (OBASE) based on source model of moving flexible 3D Objects, *IEEE Trans. IP*, **3**. no 5. (Sept 1994).

[48] BANHAM (M. R.) et al, Low bit rate video coding using robust motion vector regeneration in the decoder, *IEEE Trans. IP*, **3**, no 5 (Sept 1994.).

[49] GHORBEL (F), DAOUDI (M), MOKADEM (A.) AVARO (O), Global planar rigid motion estimation, applied to object-oriented coding, *International Conference on Pattern Recognition ICPR*, Austria (Aug 1996).

[50] MOKADEM (A.), DAOUDI (M), GHORBEL (F.) : A shape distance by complete and stable invarient description for contour tracking, *International Conference on Pattern ICPR*, Austria, August (1996).