Randomness and Geometric Features in Computer Vision Xavier PENNEC, Nicholas AYACHE
Presented at the Int. Conf. on Computer Vision & Pattern Recognition (CVPR'96)
San Francisco, Cal, USA – June 16-20, 1996

Randomness and Geometric Features in Computer Vision

Xavier PENNEC Nicholas AYACHE
EPIDAURE Project*
INRIA, B.P. 93, 2004 route des Lucioles
06902 SOPHIA ANTIPOLIS Cedex, FRANCE
Email: Xavier.Pennec@sophia.inria.fr

Abstract

It is often necessary to handle randomness and geometry in computer vision, for instance to match and fuse together noisy geometric features such as points, lines or 3D frames, or to estimate a geometric transformation from a set of matched features. However, the proper handling of these geometric features is far more difficult than for points, and a number of paradoxes can arise. We analyse in this article three basic problems: (1) what is a uniform random distribution of features, (2) how to define a distance between features, and (3) what is the "mean feature" of a number of feature measurements, and we propose generic methods to solve them.

1 Introduction

Many algorithms in computer vision and object recognition deal with simple geometric features like points, for example the Iterative Closest Point ([4, 21]), the geometric hashing ([13, 19]), and the alignment ([3, 11]) algorithms. On the other hand, models of the real world often lead to the consideration of more complex features: lines ([8]), planes, oriented points, frames ([16, 18]), etc...

The handling of these features raises some problems and can lead to paradoxes such as Bertrand's paradox concerning geometric probabilities. We have previously shown ([18]) that additive noise is not suited for describing the uncertainty of frames and should be replaced by a compositive model of noise. Other examples are presented in this article and demonstrate the need for a particular attention when dealing with geometric features.

We determined that the key point to analyse geometric feature is their behavior under a transformation group (see 1).

The article is organized in four parts. Section 2 focuses on the nature of geometric features, namely

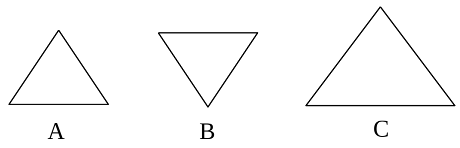

Figure 1: Comparison of geometric objects: we can say that all three triangles are similar (relative to similarity transformations); or that only **A** and **B** are congruent (relative to rigid transformations); or that they are all different (relative to translations).

points on a manifold, and we investigate transformation groups (rigid, affine,...) that operate on this manifold. In Section 3, we investigate the standard geometric probabilities and in particular need to define an invariant measure on random features (under the considered transformation group) in order to obtain a meaningful result. This leads to the computation of the prior probability of a false match in recognition algorithms. Section 4 is devoted to metric properties and section 5 to their use in providing a definition of the expected and average features which is invariant of the representation and stable under the transformation group. Due to the lack of space, we refer to ([17]) for the experiements.

2 Features and transformations

2.1 Transformations: Lie groups

There are many familiar transformations: translations, rotations, similarities, affine transformations... More generally, a transformation of a set X is a one-to-one map of X onto X. If g is a transformation, we will denote by $g \star x = g(x)$ the application of the transformation to an element $x \in X$, and by $g^{(-1)}$ the inverse map. If g_1 and g_2 are two transformations, the map $g(x) = g_2(g_1(x))$ is also a transformation: the composition of g_1 and g_2 ($g = g_2 \circ g_1$). A set of transformations $\mathcal{G}$ which is stable under these two maps is a transformation group of X.

The important class of Lie groups is obtained if $\mathcal{G}$ has a separate topological structure (a Hausdorff space) and the composition and inversion maps are

*http://www.inria.fr/Equipes/EPIDAURE-eng.html differentiable ($\mathcal{G}$ is then a differentiable manifold). In fact, most usual transformation groups are Lie ones, as soon as the group is continuous and has reasonable operations. In this article, we will use the 3D rigid motion group as an example application. An element of this group can be defined as the composition of a rotation with a translation. It can be represented by $F = (R, t)$, where the translation t belongs to $\mathbb{R}^3$ and where R is a rotation matrix (a 3x3 matrix satisfying $R.R^\top = R^\top.R = Id$ and $\det(R) = +1$), and hence belongs to the special orthogonal group SO_3. The inverse and compose maps are easily written ("." is the matrix multiplication):

$$F^{(-1)} = (R^\top \,,\; R^\top.t)$$
$$F = F_2 \circ F_1 = (R_2.R_1 \,,\; R_2.t_1 + t_2)$$

2.2 Geometric features: manifolds

Geometric features are generally defined as sets of points in the plane or 3D space, and the set of all geometric features of a given type can be described by a parameter p and a function $\beta(p, x)$ which associates the parameter p to the geometric feature (the set $\{x \in \mathbb{R}^n \ / \ \beta(p, x) = 0\}$). The function β describe a particular type of geometric feature (lines, planes, curves, triangles, ...) with a specific *representation* p. Usual sets of geometric features, such lines, curves, surfaces... are regular and constitute differential manifolds. This means that they are not traditional vector spaces, but locally may be treated as if they were (the set is locally diffeomorphic to a vector space $\mathbb{R}^m$). Spheres or smooth surfaces are such manifolds, as is the set of rotation matrices which is equivalent to $\mathcal{P}^3$ (the projective space of $\mathbb{R}^4$) by means of unit quaternions (see [18, 2]). Points trivially constitute a manifold since they already are a vector space.

There are often numerous ways to represent a given manifold, with different properties. We can for instance define a manifold as an embedding in $\mathbb{R}^k$ with differentiable constraints (e.g. rotation matrices) For other purposes, in particular differentiation, it is necessary to have a *minimal representation* where the dimension of the parameter is the dimension of the manifold (e.g. the rotation vector), or more generally a set of charts forming an atlas of the manifold, exactly the same way we need several charts to represent the earth surface in a continuous way everywhere. Each chart is defined by a one-to-one differentiable map $\varphi_i(p)$ from the representation into the manifold and an open definition domain $\mathcal{D}_i$. The set of charts must cover the manifold and must overlap each other so that it is possible to move from one chart to another.

In the case of geometric objects, the transformation usually applies to the plane or the 3D space (or more generally $\mathbb{R}^n$), but we want to work directly on features and thus must take particular care that their nature is preserved during transformations. Consider, for instance, that two orthonormal axes are no longer orthonormal after a general affine transformation. The first constraint is then for the manifold $\mathcal{M}$ to be globally invariant under the considered transformation group $\mathcal{G}$. We can then define the image of the feature p, satisfying $\beta(p, x) = 0$, by a transformation $g \in \mathcal{G}$ as being the feature $p' \in \mathcal{M}$ realizing $\beta(p', g \star x) = 0$. We will write $p' = g \star p$. With this definition, the group $\mathcal{G}$ is also a transformation group of the manifold $\mathcal{M}$.

2.3 Representation of rotations and frames

It is well known that a rotation matrix can be characterized by an angle θ around an axis n (unit vector), but since the axis is constrained, this couple is not minimal (the dimension of the representation is 4 instead of 3) and the axis is moreover not defined for the identity transformation. The rotation vector $r = \theta.n$ is always defined (multiply, since θ is modulo 2π) and differentiable. In order to define an atlas of rotations, we need in theory four charts, but we use in practice only one

- Non reflection rotations are represented by rotations vectors r from the open ball $\mathcal{B}^3(0, \pi)$.

and we remembered that, at the boundary of the domain, $r = \pi.n$ and $r' = -\pi.n$ are identified. Let $\mathtt{R}(r)$ and $\mathtt{r}(R)$ denote the mappings between rotation vectors and matrices, we can now write directly the composition and inversion laws on the representation:

$$r^{(-1)} = \mathtt{r}(\mathtt{R}(r)^\top) \qquad \text{and} \qquad r_2 \circ r_1 = \mathtt{r}(\mathtt{R}(r_2).\mathtt{R}(r_1))$$

In this article, we used frame features under the rigid motion group. A frame is defined by a point with an associated orthonormal trihedron. A particularity of frames is that they are equivalent to rigid transformations. Indeed, any frame defines a basis for 3D space and so we can represent each frame by the rigid transformation which map the canonical basis to itself. It is then easy to verify that composition and application are equivalent. Frames and motions are represented by a rotation vector and a translation: for convenient notation, we will write $f = (r, t)$ and consider it as a column vector. In this framework, a point of the Euclidean space is denoted by x, and the standard operators can then be written:

- Composition: $f = f_2 \circ f_1 = (r_2 \circ r_1 \,,\; r_2 \star t_1 + t_2)$
- Inversion: $f^{(-1)} = \left(r^{(-1)} \,,\; r^{(-1)} \star (-t)\right)$
- Application to a point: $x' = f \star x = r \star x + t$
- Application to a frame: $f = f_2 \star f_1 = f_2 \circ f_1$

2.4 Homogeneous features

A special kind of relation between the manifold and the group will be very important: let $\mathcal{O} \in \mathcal{M}$ be an element called the origin. The manifold $\mathcal{M}$ is *transitive* or *homogeneous* for the group $\mathcal{G}$ if any other element of the manifold can be obtained by a transformation of $\mathcal{G}$, i.e. if $\mathcal{G} \star \mathcal{O} = \{g \star \mathcal{O} \ / \ g \in \mathcal{G}\} = \mathcal{M}$. This means that the features we consider have no invariants. In fact, we assume that we can split the features into an invariant part (which we do not consider here) and a variable part under the group of transformation.

In the case of homogeneous features, we identify the manifold with equivalence classes of group elements in the following way. Let $\mathcal{H}$ be the subset of transforms that leave $\mathcal{O}$ invariant:

$$\mathcal{H} = \{h \in \mathcal{G} \ / \ h \star \mathcal{O} = \mathcal{O}\} \tag{1}$$

$\mathcal{H}$ is a group and is called the *isotropy* or *stability* group of $\mathcal{G}$ at $\mathcal{O}$. The left cosets $g \circ \mathcal{H}$ can be identified with elements of $\mathcal{M}$. Indeed, if $g \star \mathcal{O} = x$, then $g \circ \mathcal{H}$ is the set of transformations which map $\mathcal{O}$ to x. We will write

$$\hat{x} = \{g \in \mathcal{G} \ / \ g \star \mathcal{O} = x\} \tag{2}$$

If we consider for instance point features with origin $\mathcal{O} = 0$ and rigid transformation, we have $\mathcal{H} = \{(R, 0) \ / \ R \in SO_3\}$ and $\hat{x} = \{(R, x) \ / \ R \in SO_3\}$ where R is any rotation. For frame, taking as the origin the canonical basis ($\mathcal{O} = (\ I_d, 0)$), then $\mathcal{H}$ and all its cosets are reduced to a single point: $\hat{f} = \{f\}$. This special case where the manifold is equivalent to the group leads to important simplifications in the theory. As an intermediate example, we consider oriented points: we set the origin to $\mathcal{O} = (0, e_3)$, which means that all rotations R_z around the third axis e_3 will let it invariant: $\mathcal{H} = \{R_z \in SO_3 \ / \ R.e_3 = e_3\}$.

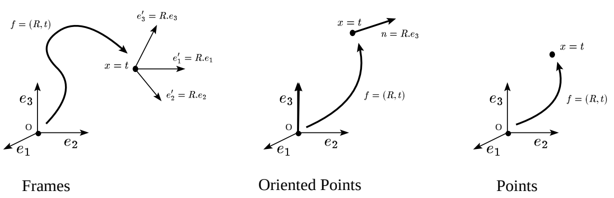

Figure 2: Chosen origins for frames, oriented points and points, and transformation of it by a rigid motion.

3 Classical Geometric probabilities

The first class of problems in geometric probability are questions of how to measure the probability of occurrence of an event when some geometric elements are randomly distributed. Bertrand's paradox illustrate the need to consider invariance by a transformation group in order to obtain a single and well defined result. In fact, the problem lie in the notion of a *uniform* distribution (or measure). Some more recent results in Lie group theory provide a means of computing the left invariant measure on the group $\mathcal{G}$, which induces the invariant measure on homogeneous manifolds. An application is presented with the generalization of the false positives analysis.

3.1 Bertrand's paradox

The problem raised by J. Bertrand in 1907 is to find the probability that a "random" chord of a circle has a length greater than the side of an inscribed equilateral triangle. Without loss of generality, we can fix the radius to 1 and the side length of the triangle is then $\sqrt{3}$. This problem can be tackled by at least three methods, which are illustrated in figure 3.

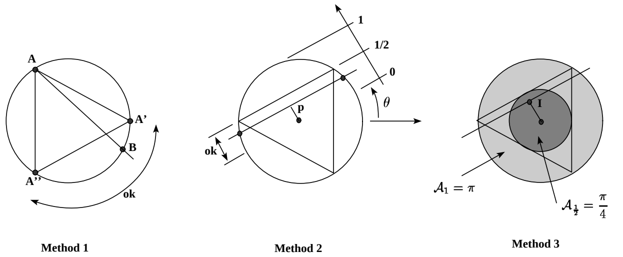

Figure 3: Three methods to compute the probability that a "random" chord of a circle has a length greater than the side of an inscribed equilateral triangle. From left to right, the methods give a probability of $\frac{1}{3}$, $\frac{1}{2}$ and $\frac{1}{4}$.

Method 1: By definition, a chord intersects the circle in two points, and we may assume that these two points are equally and independently distributed on the circle. Assume that one of the points is A in figure 3. Then the second point has to lie between A' and A'' in the circle for the chord to be greater than the triangle side. This is just $\frac{1}{3}$ of the circumference and the searched probability is then $\frac{1}{3}$.

Method 2: A chord is characterized by its distance p to the center (between 0 and 1) and its orientation θ w.r.t. a fixed line (between 0 and 2π). If we draw the equilateral triangle with a side parallel to the chord, we can see that the distance d has to be less than $\frac{1}{2}$ in order to have a chord length greater than $\sqrt{3}$. Assuming a uniform orientation and distance to the origin, we find then a normalized probability of $\frac{1}{2}$ (since p varies between 0 and 1).

Method 3: A chord is uniquely defined by the orthogonal projection I of the circle center onto it. It has to lie inside the disc of radius $\frac{1}{2}$ in order to have a sufficient length. So, assuming I is uniformly distributed over the interior of the circle, the normalized probability is $\frac{1}{4}$.

The above three solutions are correct but they do not refer to the same notion of uniformity in the way we choose the chord. Using the (p, θ) representation (described in the second method), we can compute (see [12]) that the probability measures are respectively

$$d\sigma_1 = \frac{dp.d\theta}{2\pi\sqrt{1-p^2}} \qquad d\sigma_2 = \frac{dp.d\theta}{2\pi} \qquad d\sigma_3 = p.\frac{dp.d\theta}{\pi}$$

and only the second one is invariant under the action of rotations, translations and reflections.

The solution to this problem is to impose an invariance constraint, or more precisely to *define* the notion of uniformity: for instance, uniform on $\mathbb{R}$ means that the probability for a point to lie on an interval $]x; x+d[$ is the same for all x. This is basically an invariance by translation. In the same way, and since we can only compare geometric objects with a transformation group, we define the uniform (or invariant) measure (the infinitesimal volume element) as the measure being invariant by the action of any fixed element f of the group. Let $d\mathcal{M}(x)$ be such a measure ($\mathcal{M}$ stands for manifold); this means that $d\mathcal{M}(f \star x) = d\mathcal{M}(x)$ for any x. The invariant measure can be sought directly on a given representation (see [12]), but a general formalism is developed in [20] to extract it from the left-invariant measure $d_L\mathcal{G}$ of the group $\mathcal{G}$.

3.2 Invariant measures on a Lie group

We can require left invariance ($d_L\mathcal{G}(g \circ f) = d_L\mathcal{G}(f)$) for any fixed $g \in \mathcal{G}$, or right invariance ($d_R\mathcal{G}(f \circ g) = d_R\mathcal{G}(f)$). Since the group acts on the left (the application of transformation f to feature x is $f \star x$), we are mainly interested in left-invariance and we will mean by invariant measure the left-invariant one.

To be mathematically correct, we require that for any continuous real function α on $\mathcal{G}$ with compact support, we have:

$$\forall g \in \mathcal{G} \quad \int_{f\in\mathcal{G}} \alpha(g \circ f).d_L\mathcal{G}(g \circ f) = \int_{f\in\mathcal{G}} \alpha(f).d_L\mathcal{G}(f)$$

If the group is locally compact, then it can be proven ([10]) that there exists only one left-invariant measure (up to a scale factor) that verifies the above properties. This measure is called the (left) Haar measure of the group. In a symmetric way, there is also a unique right Haar measure.

It is interesting to note that the left and right invariant measures are generally different: their difference is quantified by the *module* $\Delta\mathcal{G}$ defined by: $d_L\mathcal{G}(f) = \Delta\mathcal{G}(f).d_R\mathcal{G}(f)$. The group is called unimodular if the module is identity everywhere, which means that left and right measures are equal. A compact group is always unimodular, but only locally compact groups can have different left and right Haar measures. For instance, left and right Haar measures are identical on SO_3, since the 3D rotation group is compact, and the 3D rigid motion group is unimodular although the group is only locally compact due to the introduction of translations (see below).

The left and right invariant measures can be generally computed from the Maurer-Cartan equations ([20]), but a very interesting theorem allows easy computation in the case of a *minimal* representation: assume that the definition domain of a chart almost covers the group[1] and that the Jacobian of the left translate of the identity $J_L(f)$ exists and is continuous almost everywhere. Then the invariant measure can be written:

$$d_L\mathcal{G}(f) = \frac{df}{|J_L(f)|} \quad \text{with} \quad J_L(f) = \left.\frac{\partial(f \circ e)}{\partial e}\right|_{e=I_d} \tag{3}$$

where $|J| = |\det(J)|$. The right-invariant measure can be derived in the same way using the Jacobian J_R of the right translate of the identity, and the module is thus $\Delta\mathcal{G}(f) = |J_R(f)| \,/\, |J_L(f)|$. Using this scheme, we can show that the uniform measure for rotations using the rotation vector representation is

$$d_L\mathcal{G}(r) = d_R\mathcal{G}(r) = \frac{4\sin^2(\theta/2)}{\theta^2}\, dr \tag{4}$$

where $\theta = \|r\|$. For rigid transformations with our representation $f = (r, t)$, we have just to multiply by the translation measure dt. Rotations and rigid transformations are thus unimodular.

3.3 Invariant measure on manifolds

If the manifold is not homogeneous, we factorize it into $\mathcal{M} \times \mathcal{I}$ where $\mathcal{I}$ is the invariant part on which we can put any measure. For the homogeneous part $\mathcal{M}$, we saw in section 2.4 how to identify it with the quotient space $\mathcal{G}/\mathcal{H}$. We can find, thanks to the above section, the (left) invariant measures $d_L\mathcal{G}$ on $\mathcal{G}$ and $d_L\mathcal{H}$ on $\mathcal{H}$, and write $d_L\mathcal{G} = d\mathcal{M}.d_L\mathcal{H}$ where $d\mathcal{M}$ is a measure on the manifold $\mathcal{M}$ (or the quotient space $\mathcal{G}/\mathcal{H}$). Santalo gives in [20] several forms of a necessary and sufficient condition for $d\mathcal{M}$ to be an invariant measure. One of them can be stated as follows (e is an element of the manifold $\mathcal{M}$):

$$\forall h \in \mathcal{H}, \qquad \left| \left.\frac{\partial(h \star e)}{\partial e}\right|_{e=\mathcal{O}} \right| = 1 \tag{5}$$

This means that the measure of the infinitesimal volume element at the origin remains unchanged under any transformation that keep the origin unchanged. If this condition is not verified, then there is no invariant measure on the manifold, otherwise we can compute it (with a minimal representation of the manifold) in a way similar to what we did for the group (f_x is now any transformation of $\hat{x}$):

$$d\mathcal{M}(x) = \frac{dx}{|J(f_x)|} \quad \text{with} \quad J(f_x) = \left.\frac{\partial(f_x \star e)}{\partial e}\right|_{e=\mathcal{O}} \tag{6}$$

[1] Since we integrate function (and not distributions), we can "forget" a subset of the group that have a null measure.

3.4 Prior probability of a false match

Assume that x is a uniform random feature in the first image (characterized by a set of possible features I_1). What is the probability that this feature be accepted as a match with feature y in the second image, under a given global transformation f ?

If we characterize the possible matches for y by an "error volume" $Z(y)$ around y, we can write this probability as the conditional probability

$$P(f \star x \leftrightarrow y) = P((f \star x) \in Z(y) | x \in I_1) = \frac{\mathcal{V}((f \star I_1) \cap Z(y))}{\mathcal{V}(I_1)}$$

where $\mathcal{V}(X)$ is the "volume" of the set X. With the assumption that the volume $\mathcal{V}(Z(y))$ is sufficiently small with respect to the volume of the image, we can consider that the transformed image I_1 either contains the whole set $Z(y)$ or does not intersect it at all. This allows us to approximate the above probability by

$$P(f \star x \leftrightarrow y) = \varepsilon \frac{\mathcal{V}(Z(y))}{\mathcal{V}(I_1)}$$

with ε beeing 1 if $f^{(-1)} \star y \in I_1$ and 0 otherwise. A desirable property for our "error volume" $Z(y)$ is that it be comparable at every point (as we usually fix the same bound for error on all the points): this means that, for any pair of points y and y' on the manifold, there exists a transformation f such that $f \star y = y'$ and $Z(y') = f \star Z(y)$ (the error volume is said to be homogeneous). A stronger hypothesis is that *for every transformation* f, the error volume on the transformed point is the transformation of the error volume: $Z(f \star y) = f \star Z(y)$. The volume is said to be isotropic in this case, and is completely determined by its shape around the origin. In both cases (homogeneity and isotropy), the volume of the error volume is invariant: $\mathcal{V}(Z(y)) = \mathcal{V}(Z(\mathcal{O})) = \mathcal{V}_0$. The basic probability of a false match we obtain can now be applied as usual in an analysis of the frequency of false positives (see [9, 14, 15]).

As a practical example, we considered in [16] that two frames are matched if the distance between their point is less than a threshold d_0 and if the rotation needed to adjust their trihedra has an angle less than a threshold θ_0. This is in fact a bound on an invariant distance (see section 4.5), the volume is thus invariant and we can compute it at the origin. If $f = (r, t)$ is a frame, it is in the error volume $Z(Id)$ if $\theta = \|r\| < \theta_0$ and $\|t\| < d_0$. Using the invariant measure of equation 4, we can compute the volume of the error zone:

$$\mathcal{V}_0 = \int_{\theta < \theta_0} \int_{\|t\| < d_0} d\mathcal{M}(r, t) = 2\pi(\theta_0 - \sin\theta_0).\left(\frac{4\pi}{3} d_0^3\right)$$

If we assume a cubic image of side l (256 for instance) this gives a Euclidean volume $V_I = l^3$ for points in which trihedrons are not constrained: the rotation volume is $2\pi^2$. We obtain finally the basic probability of false match:

$$P(f \star x \leftrightarrow y) = \left(\frac{\theta_0 - \sin\theta_0}{\pi}\right) \frac{4}{3} \left(\frac{d_0}{l}\right)^3 \varepsilon$$

We have isolated in the first term the probability of false match due to trihedra only, which reflects the gain in selectivity when using frames instead of just points. This function is plotted in figure 4 and shows very interesting results: even for a bound of $\theta_0 = \pi/2 = 90\,\text{deg}$, more than 80% of the random matches are rejected. For a bound of $\theta_0 = \pi/10 \simeq 20\,\text{deg}$, the probability of a false match drops to 0.0016: we would have to divide the bound on the position by 10 to obtain an equivalent selectivity using points only.

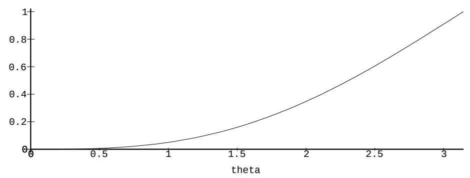

Figure 4: Basic probability of a false match for trihedrons with a bound on the angle for the adjustment rotation of theta.

4 Metric properties

A distance between points is often used to classify, quantify and minimize. This is even the core of some algorithms (Iterative Closest Point for instance). In order to be able to compare features (and generalize a number of methods on points), we also need a distance function on features. However, the distance has to be carefully chosen: a highly desirable property for this distance is the invariance under the transformation group and the invariance under the chosen representation (see example below).

A general method is given to generate an invariant distance on the manifold from a metric on the transformation space, and this allows the generalization of several algorithms based on distance to homogeneous features. The invariance property automatically guarantees the stability of the result with respect to the considered group of transformations.

4.1 The paradox of the closest line

We saw with the paradox of Bertrand different representations of 2D line. We use in this section another minimal representation based on the line equation ([2]): the equation of a 2D line is $a.x + b.y + c = 0$. In order to obtain a minimal representation, we need to eliminate one parameter:

- **Chart 1**: lines that are not parallel to the X axis are represented by $d = (a, p) \in \mathbb{R}^2$ and have equation: $a.x + y + p = 0$.

- **Chart 2**: lines that are not parallel to the Y axis are represented by $d' = (a', p') \in \mathbb{R}^2$ and have equation: $x + a'.y + p' = 0$.

In the first chart, the line $d = (a, p)$ cuts the Y axes at the point $(0, -p)$ and has a director vector $(1, -a)$. This is symmetric in the second chart: the line $d' = (a', p')$ cuts the X axes at the point $(-p', 0)$ and has a director vector $(-a', 1)$.

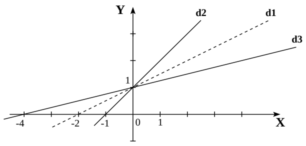

Figure 5: Three lines in the plane.

We draw in figure 5 three lines. The problem is to choose which line d_2 or d_3 is the closest one to d_1. A definition of the distance that seems to be reasonable is $\text{dist}(d_1, d_2) = \sqrt{(a_1 - a_2)^2 + (p_1 - p_2)^2}$. The coordinates of the three lines in the first chart are $d_1 = (-1/2\,;\,-1)$, $d_2 = (-1\,;\,-1)$, $d_3 = (-1/4\,;\,-1)$, and the distance between line d_1 and d_2 turns out to be equal to the distance between line d_1 and d_3:

$$\text{dist}(d_1, d_2) = \text{dist}(d_1, d_3) = 1/2$$

The line d_1 is thus the mid-line between d_2 and d_3. If we now consider the lines in the second chart, their coordinates are $d_1 = (-2\,;\,2)$, $d_2 = (-1\,;\,1)$, $d_3 = (-4\,;\,4)$ and the distances are now

$$\text{dist}(d_1, d_2) = \sqrt{2} \qquad \text{and} \qquad \text{dist}(d_1, d_3) = 2\sqrt{2}$$

We have then the following paradox: d_1 is the mid-lines of d_2 and d_3 in one chart, and d_2 is the closest line to d_1 in the second chart. Which chart gives the most reasonable result ? In fact, visually, the line d_3 is the closest one: there is an angle of 12.5^o between d_1 and d_3, versus an angle of 18.5^o between d_1 and d_2 (these angles are obsiously invariant by rigid transformation).

4.2 Invariant distance on a manifold

Let $x, y \in \mathcal{M}$ and $g \in \mathcal{G}$, then we require that $\text{dist}(x, y) = \text{dist}(g \star x, g \star y)$. This means in particular that this distance is completely defined by the distance $N(x) = \text{dist}(x, \mathcal{O})$ of a feature x with the origin: if we use transformation $f_y^{(-1)}$ or $f_x^{(-1)}$, we have

$$\text{dist}(x, y) = N(f_y^{(-1)} \star x) = N(f_x^{(-1)} \star y) \qquad (7)$$

The symmetry, positivity, definite character and triangular inequality for the distance are translated, under the invariance assumption, into the three following properties:

- $N(x) \geq 0$ and $(N(x) = 0) \Leftrightarrow (x = \mathcal{O})$.
- $N(f_x^{(-1)} \star \mathcal{O}) = N(x)$ for $f_x \in \hat{x}$ and thus $N(h \star x) = N(x)$ for any $h \in \mathcal{H}$.
- $N(x) + N(y) \geq N(f_y^{(-1)} \star x) = N(f_x^{(-1)} \star y)$ for any $f_x \in \hat{x}$ and $f_y \in \hat{y}$

These properties are very close to those required in order to define a norm on a vector space (without the positive linearity). To distinguish the function N from the associated invariant distance, we call N the "norm" of the manifold. We note that we have so far defined the metric on the manifold and not in a particular chart. In practice, we use a "principal chart", centered around the origin and covering almost the manifold. The norm N is defined in this chart, and when we have to use the distance $\text{dist}(x, y)$, we compute $N(f_y^{(-1)} \star x)$.

4.3 Invariant distances on a Lie group

Assume that we work now on the transformation group $\mathcal{G}$. We can require the distance to be either left or right invariant. We are only interested here in the left-invariant distance since it can induce an invariant distance on an homogeneous manifold. As above, a left invariant distance is determined by a metric N_g on the group, satisfying the following properties:

- $N_g(f^{(-1)}) = N_g(f)$.
- $N_g(f) \geq 0$ and $(N_g(f) = 0) \Leftrightarrow (f = I_d)$.
- $N_g(f) + N_g(g) \geq N_g(g^{(-1)} \circ f) = N_g(f^{(-1)} \circ g)$.

and the corresponding left invariant distance is

$$\text{dist}_L(f, g) = N_g(g^{(-1)} \circ f) = N_g(f^{(-1)} \circ g)$$

4.4 Distance induced by the group

Let N_g be a norm on the group $\mathcal{G}$. We define the induced semi-norm on the homogeneous manifold $\mathcal{M}$ as

$$N(x) = \inf_{(h_1, h_2) \in \mathcal{H}^2} \left(N_g(h_1 \circ f_x \circ h_2)\right) \qquad (8)$$

The corresponding invariant distance satisfies

$$\text{d}(x, y) = \inf_{\{f_x \in \hat{x}\,;\, f_y \in \hat{y}\}} \left(\text{dist}_L(f_x, f_y)\right) \qquad (9)$$

If the infimum of is reached for every point x, then the semi-metric is separable and is thus a metric. This property is always true if the isoptropy group $\mathcal{H}$ is compact but is not automatically verified otherwise (for instance, there is no metric induced on points by the similarities or affine transformations).

4.5 Practical use on rigid transformations

The Euclidean distance on $\mathbb{R}^3$ is induced by the L_2 norm: $d_t(x, y) = \|x - y\|$. On the other hand, it can be shown ([1]) that the angle θ of a rotation is a metric that induce a left and right invariant distance on SO_3, the rotation group. With the rotation vector representation, we have $d_\theta(r_1, r_2) = \|r_2^{(-1)} \circ r_1\| = \|r_1 \circ r_2^{(-1)}\|$ and the norm on the rigid motion group is the combination of these two metrics:

$$N_\lambda(f) = N_\lambda((r, t)) = \|f\| = \sqrt{\lambda^2\,\|r\|^2 + \|t\|^2}$$

where λ is a fixed parameter that allow to tune the importance of the trihedron (rotation part) with respect to the position (or translation part). We usually scale each of the two terms by the inverse of their variation domain (π for θ and the diameter l_0 of the image or the interest object for the translation: $\lambda = l_0/\pi$). When we have a an information about the noise level (i.e. standard deviations σ_θ and σ_t), we can also use $\lambda = \sigma_t/\sigma_\theta$.

We can check that the distance induced by this metric on the original space is the Euclidean distance. The left-invariant distance is thus

$$\text{dist}_L(f_1, f_2)^2 = \|f_2^{(-1)} \circ f_1\|^2 = \lambda^2 \|r_2^{(-1)} \circ r_1\|^2 + \|t_1 - t_2\|^2$$

We note that this distance is *not* right invariant, although the rigid motion group is unimodular (and thus left and right invariant measures are identical). This reflects the fact that $N_\lambda(f_1 \circ f_2) \neq N_\lambda(f_2 \circ f_1)$.

4.6 Utility of invariant distances

The distance between two points is often used in geometric algorithms: the Iterative Closest Point algorithm, developed in [4, 21] is the best example. Another example is the classic method of computing the transformation that maps a set of points x_i in one image, to a set of points y_i in another image: it is the transformation which the minimizes the sum of squared distances $C(f) = \sum_i \text{dist}^2(f \star x_i, y_i)$.

All these algorithms can be extended to homogeneous features in a straightforward way using the invariant distance previously defined. The nice property that we have in using this distance is the "conservation" of the result under a change in the reference frame (the application of a transformation). Let F be the transformation minimizing the above criterion $C(f)$. Assuming that the features x_i are transformed by a transformation g ($x_i' = g \star x_i$), the new result is $F' = F \circ g$, but if the features of the second image are globally transformed by g ($y_i' = g \star y_i$), then we need here the invariance property of the distance to conclude that the new minimum is $F'' = g^{(-1)} \circ F$, which gives the expected result $F = g \circ F''$. The same experiment can be done if both images are transformed by the *same* transformation g (this means a global change of the reference frame), and the invariance of the distance is required to prove that the transformation found is $F''' = g^{(-1)} \circ F \circ g$, i.e. only the change of the reference frame.

5 Expectation of random features

Uncertainty on geometric features (and more generally on measurements) is usually characterized by a probability density function whose expected value corresponds to the exact value. In order to deal efficiently with it from a computational point of view, we need to keep only a few number of parameters characterizing this pdf. The usual way ([5, 2, 22]) is to consider the representation of the random feature as a random vector and, assuming that the pdf is quasi-Gaussian, approximate it up to the second order by its expectation and covariance matrix. We focus in the sequel on the expectation $\bar{x}$ and its statistical measurement (or its discrete version): the empirical mean. The classical definition is, for a pdf p_x (in the parameter space) and a set of measured features $\{x_i\}$:

$$\bar{x} = \mathbf{E}(x) = \int_{\mathcal{D}} y.p_x(y).dy \quad \text{and} \quad \mathbf{E}(\{x_i\}) = \frac{1}{n} \sum_i x_i$$

We claim that these operators are not properly defined. In particular, the result of the integral or the sum is *not* ensured to be in the definition domain and defines not necessarily a feature: for instance, the arithmetic mean of several rotation matrices is generally not an orthogonal matrix and is therefore not a rotation itself (particularly for large deviations). The second reason, is that the expectation does not commute in general with the application of a fixed transformation (see example below). This means that the mean value of a pdf depends on the chosen reference frame, which is unacceptable.

5.1 Standard expectation of a 2D random line

We consider for this example 2D oriented vector lines, which can be represented by a point on the unit circle, and therefore an angle θ with a given axis. We fix the domain of θ to be $\mathcal{D} =]-\pi, \pi]$. The action of a rotation of angle λ is simply the addition (modulo 2π). We can define an uncertain line by its probability density function $\rho(\theta)$, and the classical way to obtain the expected value is to integrate in the parameter space:

$$\bar{\theta} = E(\theta) = \int_{\mathcal{D}} \theta.\rho(\theta) \frac{d\theta}{2\pi}$$

where the term 2π is a normalization factor. We note that $d\theta$ is the uniform measure for lines under rotation. Let $\rho_0(\theta) = 2\cos(\theta/2)^2$ be such a pdf (drawn in figure 6). It is normalized and its expectation is $\bar{\theta}_0 = 0$. If we now change the reference frame, i.e. apply a rotation of angle λ, we obtain the new pdf $\rho_\lambda(\theta) = \rho_0(\theta - \lambda)$, and the expected value is now

$$\bar{\theta}(\lambda) = \int_{\mathcal{D}} \theta.\rho_0(\theta - \lambda) \frac{d\theta}{2\pi} = 2\sin(\lambda/2)\cos(\lambda/2)$$

which is different from the rotated expectation $\underline{\theta}(\lambda) = \lambda$. In particular, for a rotation of $\lambda = \pi$, we find that $\underline{\theta}_\pi = \pi$ and $\bar{\theta}_\pi = 0$.

The first idea to avoid this type of problems is to "center" the definition domain of the chart around the

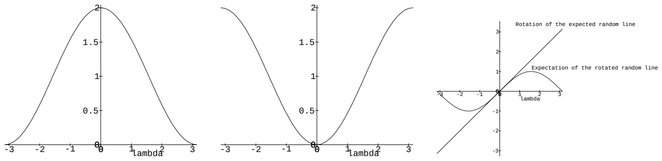

Figure 6: Left: Original pdf $\rho_0(\theta) = 2\cos(\theta/2)^2$. Expectation is $\bar{\theta}(0) = 0$. Middle: pdf after a rotation of angle $\lambda = \pi$. The expectation $\bar{\theta}(\pi)$ is also 0, whereas the rotated expectation is $\underline{\theta}_\pi = \pi$. Right: comparison of the expectation of the rotated line and the rotation of the expected line.

expected feature. If this can be easy in the case of a circle, as above, it may be more problematical for some other features such as 3D lines, or frames, where the manifold is far more complex. The question is similar for the mean value, especially with scattered measurements. Last, but not least, if the chart is centered around the expected (or the mean) feature, then the problem is already solved. The Frechet expectation is a well-posed formalism to implement this idea: the "centrality" of a feature is based on its distance with other measurements, and the mean or expected values are the features that optimize the "centrality".

5.2 Frechet expectation of a random feature

Let $\mathbf{v}$ be a random vector in $\mathbb{R}^n$. Frechet ([7]) observed that the variance $\sigma^2_{\mathbf{v}}(x) = E(\operatorname{dist}(\mathbf{v}, x)^2)$ is minimized at the expected value $\bar{v}$. The second point is that if the expectancy of the representation of a feature (a vectorial integral) is not well defined (because features are not vectors), the expectation of a real mapping (and in particular the distance to a fixed feature) is properly defined.

Let $dist$ be an invariant distance on the manifold $\mathcal{M}$ under the group $\mathcal{G}$. This distance is independant of the representation. In order to get rid of the representation problem and be sure to integrate correctly on the manifold, we use from now on the "invariant" pdf previously introduced to describe random features. Let then $\mathbf{x}$ be a random feature of pdf $p_{\mathbf{x}}$. The expected square distance of a deterministic feature y with the random feature $\mathbf{x}$ is defined by

$$\sigma^2_{\mathbf{x}}(y) = E_{\mathbf{x}}(\operatorname{dist}(y, \mathbf{x})^2) = \int_{\mathcal{M}} \operatorname{dist}(y, z)^2 . p_{\mathbf{x}}(z) . d\mathcal{M}(z) \tag{10}$$

If $\sigma^2_{\mathbf{x}}(y)$ is finite for all y, we call every feature $\bar{x}$ minimizing $\sigma^2_{\mathbf{x}}$ an expected feature, and we denote by $\mathbb{E}(\mathbf{x})$ the set of all expected features of the random feature $\mathbf{x}$. We have thus

$$\mathbb{E}(\mathbf{x}) = \arg\min_{y \in \mathcal{M}} \left(\mathbf{E}_{\mathbf{x}}(\operatorname{dist}(y, \mathbf{x})^2) \right) \tag{11}$$

If there is at least one expected feature $\bar{x}$, then $\sigma_{\mathbf{x}} = \sigma_{\mathbf{x}}(\bar{x})$ is called the standard deviation of $\mathbf{x}$ and $\sigma^2_{\mathbf{x}}$ the variance. In general, there can be several expected values. However, Emery and Mokobodzki show in [6], under some strong conditions, the existence and uniqueness of the expected value.

In a very similar way, we can define the set of empirical means of features $x_1, \ldots x_n$ by

$$\mathbb{E}(\{x_i\}) = \arg\min_{y \in \mathcal{M}} \left(\sum_i \operatorname{dist}(y, x_i)^2 \right) \tag{12}$$

and if there is at least one mean $\bar{y}$, we call $s = \sqrt{\frac{1}{n} \sum_i \operatorname{dist}(y, x_i)^2}$ the empirical standard deviation and s^2 the empirical variance.

Other types of central values can be defined in this framework: we define more generally the "mean deviation" at order α by

$$\sigma_{\mathbf{x},\alpha}(y) = \left(E_{\mathbf{x}}(\operatorname{dist}(y, \mathbf{x})^\alpha)\right)^{1/\alpha}$$

If this function is finite over $\mathcal{M}$, the features $\bar{x}_\alpha$ minimizing it are called central features of order α. To be more practical, we obtain the modes of the pdf for $\alpha = 0$, the median features for $\alpha = 1$, of course the mean or expected features for $\alpha = 2$, and the "barycenter" of the support of the pdf (which is a compact set) if the mean deviation is finite for $\alpha = \infty$. This can also be applied to define the empirical central features at any order.

All the above definitions can be applied to the transformation group with our previous definitions of the left invariant distance and measure.

5.3 Invariance properties

The nice properties of the Frechet expectation and mean features are, in our case, due to the invariant distance; these sets are stable under the transformation group:

$$\mathbb{E}(g \star \mathbf{x}) = g \star \mathbb{E}(\mathbf{x}) \quad \text{and} \quad \mathbb{E}(\{g \star x_i\}) = g \star \mathbb{E}(\{x_i\})$$

Since the distance we use does not depends on the representation, the results of all minimization are ensured to be also independent of the representation. We have thus obtained a stable definition (and a mean of computation via optimization) for the expected features of a pdf and for the empirical mean features.

5.4 Application to data fusion: the mean frames

Assume that we have a set of frame measurements $f_i = (r_i, x_i)$. We are looking for the mean frame $f = (r, x)$ in the Frechet sense. From our definition of the distance from f to f_i, the least squares criterion reduces to

$$C_\lambda(f) = \lambda^2 \sum_i \|r_i^{(-1)} \circ r\|^2 + \sum_i \|x - x_i\|^2$$

We can thus minimize independently for the position and the orientation. Moreover, the solution is independent of the parameter λ.

The position is given by the barycenter of the frame positions, and the orientation is obtained by an iterative gradient descent (equations for the derivation can be found in [18]). The gradient descent can be repeated for several starting points to verify that the global minimum is obtained an to test its uniqueness.

Another method following the same principles but incorporating second order informations (covariance matrices) was proposed in [18].

6 Experiments: a data fusion problem

This section is reachable in the corresponding INRIA research report ([17]).

7 Conclusion

We showed in this paper that a wide range of paradoxes arise when we try to generalize to geometric features the classical algorithms used for points, and we demonstrate, in the case of homogeneous features, that they can be avoided by the careful definition of basic operators which respect the following rules:

- Invariance with respect to the representation.
- Invariance or commutativity with respect to the action of the associated transformation group.

We are currently working on a theory of uncertainty on geometric features continuing the formalism introduced in this paper.

Acknowledgments

Part of this work was supported by a fellowship from D.R.E.T./D.C.N. (France).

References

[1] S. L. Altmann. *Rotations, Quaternions, and Double Groups.* Clarendon Press - Oxford, 1986.

[2] N. Ayache. *Artificial Vision for Mobile robots - Stereo-vision and Multisensor Perception.* MIT-Press, 1991.

[3] N. Ayache and O.D. Faugeras. Hyper : A new approach for the recognition and positioning of two-dimensional objects. *IEEE Trans. on Patern Analysis and Machine Intelligence*, 8(1):44–54, 1986.

[4] Paul Besl and Neil McKay. A method for registration of 3−D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[5] H.F. Durrant-Whyte. *Integration, Coordination and Control of Multi-Sensor Robot Systems.* Kluwer Academic Publishers, 1988.

[6] M. Emery and G. Mokobodzki. Sur le barycentre d'une probabilité dans une variété. In M. Yor J. Azema, P.A. Meyer, editor, *Séminaire de probabilités XXV*, number 1485 in Lect. Notes in Math., pages 220–233. Springer-Verlag, 1991.

[7] M. Frechet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincarré*, 10:215–310, 1948.

[8] W.E.L. Grimson. *Object Recognition by Computer - The role of Geometric Constraints.* MIT Press, 1992.

[9] W.E.L. Grimson and D.P. Huttenlocher. On the sensitivity of geometric hashing. In *Proc. third ICCV*, pages 334–338, 1990.

[10] G. Hochschild. *The structure of Lie groups.* Holden-Day, 1965.

[11] D.P. Huttenlocher and S. Ullman. Object recognition using alignment. In *Proc. of ICCV*, pages 72–78, 1987.

[12] M.G. Kendall and P.A.P. Moran. *Geometrical probability.* Number 10 in Griffin's statistical monographs and courses. Charlkes Griffin & Co. Ltd., 1963.

[13] Y. Lamdan and H.J. Wolfson. Geometric hashing : A general and efficient model-based recognition scheme. In *Proc. of Second ICCV*, pages 238–289, 1988.

[14] Y Lamdan and H.J. Wolfson. On the error analysis of geometric hashing. In *IEEE Int. Conf. on Comput. Vis. and Patt. Recog.*, pages 22–27, 1991.

[15] X. Pennec. Correctness and robustness of 3D rigid matching with bounded sensor error. Research report 2111, INRIA, November 1993.

[16] X. Pennec and N. Ayache. An $\mathcal{O}(n^2)$ algorithm for 3D substructure matching of proteins. In a. Califano, I. Rigoutsos, and H.J. Wolson, editors, *Shape and Pattern Matching in Computational Biology.* Plenum Publishing, 1994.

[17] X. Pennec and N. Ayache. Randomness and geometric features in computer vision. Research Report ??, INRIA, March 1996.

[18] X. Pennec and J.P. Thirion. Validation of 3D registration methods based on points and frames. In *Proceedings of the 5th Int. Conf on Comp. Vision (ICCV'95)*, pages 557–562, Cambridge, Ma, June 1995.

[19] I. Rigoutsos and R. Hummel. Distributed bayesian object recognition. In *Proceedings of Int. Conf on Comput. Vis. and Pat. Recog*, pages 180–186. IEEE Computer Society Press, June 1993.

[20] L.A. Santalo. *Integral Geometry and Geometric Probability*, volume 1 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Company, 1976.

[21] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994. Also Research Report No.1658, INRIA Sophia-Antipolis, 1992.

[22] Z. Zhang and O. Faugeras. *3D Dynamic Scene Analysis: a stereo based approach*, volume 27 of *Springer series in information science*, chapter 2: Uncertainty Manipulation and Parameter Estimation, pages 9–27. Springer Verlag, 1992.