# Evidential Reconstruction of a Sportive Gesture from a Multisensor Acquisition

David Helbert, Bertrand Augereau

# *Evidential Reconstruction of a Sportive Gesture from a Multisensor Acquisition*

David Helbert* and Bertrand Augereau*
*XLIM, CNRS, University of Poitiers, France*

## Abstract

We describe a reconstruction method of three dimensional marker trajectories from several cameras in the applicative framework of sportive gesture analysis. The principle is to apply a tracking algorithm on extracted markers from each sequence of images, then to fuse data to reconstruct three dimensional space movements of an athlete. The measurement-to-track association algorithm is based on the formalism of Dempster-Shafer's Theory of Evidence in the frame of the extended open-world, a powerful tool to represent uncertainty. Finally, the fusion enables to obtain a system which becomes efficient in its capacity to detect either the appearance or the disappearance of markers.

## 1 Introduction

Movement dynamical analysis is essential to realize a gestural analysis. It proves to be important at the time of studies on human movements. The analysis of sportive gestures aims at improving performances of athletes and to teach an "ideal" sportive practice in the point of view of the mechanical science. Kinematic analyses rely either on points on the skeleton or on the joint centers. These points are not apparent, it is thus necessary to materialize them to track the movements. For this kind of applications, markers are often put on the skin or on a suit and positions are captured via sensors.

**Corresponding author. Email: david.helbert@univ-poitiers.fr

The tracking system can be divided into the three following modules: the object detection, the tracking filter and the measurement-to-track data association. Our work consists in associating a measurement of markers with a set of tracks on each sensor without *a priori* knowledge and in reconstructing their trajectories in the three dimensional (3D) space. Some probabilistic methods do exist [Hall and Llinas(2001)] ; the most famous one are the nearest-neighbour method, the maximum of likelihood method, the probabilistic data association filter and the multiple hypothesis filter. Most of the time, only the imprecision is really processed and the notion of reliability bound to information sources is often omitted. The Dempster-Shafer theory of Evidence offers a privileged mathematical frame for the measurement-to-track association with conflict and ambiguity resolution. It avoids problems met by the classic association methods and it enables to initialize a new trajectory, as well as to handle object appearances and disappearances [Royère et al.(2000), Gruyer et al.(2000), Mourllion et al.(2005), Lemeret et al.(2008)]. In [Rombaut(1998)], M. Rombaut has suggested a measurement-to-track data association algorithm in the mathematical frame of the theory of Evidence to process both qualitative and quantitative information. D. Gruyer has resumed this method and he has developed a new measurement-to-track data association algorithm [Gruyer et al.(2002)].

Markers cannot sometimes be detected; this phenomenon is very common in this type of application: it can be due to an eclipse of a marker by a sportsman limb or to an extraction error. The global system of 3D track reconstruction thus consists in the

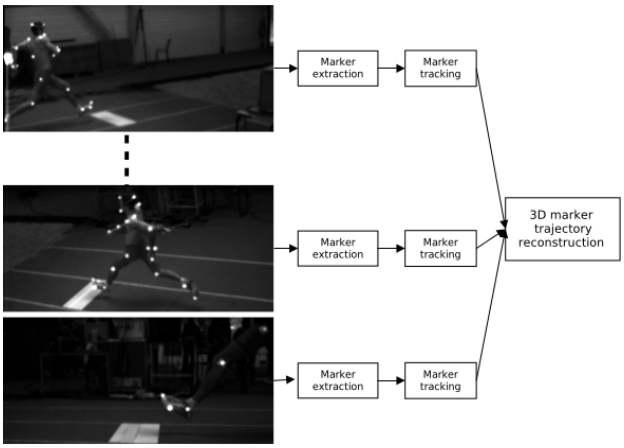



Figure 1: The global principle of 3D gesture reconstruction.

marker tracking for each sensor in the framework of Dempster-Shafer theory of Evidence, the 3D reconstruction by stereoscopic approach and the fusion of 3D tracks (cf. figure 1).

The organization of the paper is as follow. In section 2, the basis of the Dempster-Shafer theory of Evidence and an evidential method of measurement-to-track data association are explained. In section 3, we present a new marker tracking algorithm and we illustrate the efficiency. In section 4 we propose a 3D marker reconstruction and fusion to correct 3D trajectories. Finally, in section 5 we draw the conclusions and give future work to be done.

# 2 Evidential Association in Extended Open-World

The theory of Evidence, also called the belief function theory, is based on the A. Dempster's works [Dempster(1967)] and it has been resumed by G. Shafer under a more accomplished mathematical formalism [Shafer(1976)].

## 2.1 Definitions

The principle of the theory of Evidence is to manipulate beliefs on a finite set of possible hypotheses $\Theta = \{H_1, H_2, ..., H_n\}$ named frame of discernment. This set is composed of $n$ hypotheses which must be exhaustive and exclusive and it is referred to its powerset, a reference set, denoted by $2^\Theta = \{A/A \subseteq \Theta\}$.

The basic probability assignment (bpa) $m : 2^\Theta \rightarrow [0, 1]$ represents the distribution of a mass unit among $2^\Theta$ elements:

$$m(\oslash) = 0 \quad \text{and} \quad \sum_{A \subseteq \Theta} m(A) = 1, \tag{1}$$

where $\oslash$ is the empty set and $m(A)$ is the measure of belief that supports $A$.

The belief function $Bel(A)$, associated with the bpa $m(.)$, is a function that assigns a value in $[0, 1]$ to every nonempty subset $A$ of $\Theta$. The degree of belief in $A$ is defined by :

$$Bel(A) = \sum_{B/B \subseteq A} m(B). \tag{2}$$

P. Smets has also developed the belief transferable model which is an extension of the mass functions [Smets(1990)]. This approach disregards the probabilistic interpretation of the knowledge and it is based on two levels of information perception:

- the credal level where beliefs are entertained and quantified by belief functions,
- the pignistic level which is dedicated to the decision.

In the frame of this model, the definition of the pignistic probability, $BetP : 2^\Theta \rightarrow [0, 1]$, is based on the Laplace's principle of insufficient reason: in the absence of any reason for privileging a particular hypothesis, we suppose that all hypotheses are equiprobable.

$$BetP(A) = \sum_{B \subseteq \Theta} \frac{|A \cap B|}{|B|} m(B) \qquad \forall A \in \Theta, \tag{3}$$

where $|B|$ represents the cardinal of subset $B$ of $\Theta$.

The combination proposed by A. Demspter in consists in obtaining an unique mass distribution by combining multiple measures of belief through their bpas defined on the same frame of discernment [Dempster(1967)] . The combination of two bpas $m_1(.)$ and $m_2(.)$, noted $m_{1,2}(.)$, is given by the orthogonal sum formula, a conjunctive and associative operation.

## 2.2 The Extended Open-World

When G. Shafer has formalized the theory of Evidence in [Shafer(1976)], it took place within the framework of the closed-world supposing that all hypotheses are exhaustive and exclusive (eq. 1). According to P. Smets [Smets(1990)], the empty set $\oslash$ can be considered as a possible solution and can represent all the unknown hypotheses. This world, called open-world, does not correspond to the definition of Shafer which stipulates that the mass on the empty set has to be null. From this point of view, the frame of discernment consists of all known hypotheses but not all possible hypotheses. Furthermore, these hypotheses are exclusive but not necessarily exhaustive. In that case, we combine masses using a conjunctive operator.

To distinguish the real conflict and the appearance of a new hypothesis, an exhaustive study of the problem has to be realized or sources have to provide reliable information. These two solutions are difficult to obtain with complex systems. C. Royère has also proposed in [Royère et al.(2000)] a new frame of discernment: the extended open-world. An hypothesis singleton is introduced within the frame of discernment including all hypotheses which are not modelled. This hypothesis, noted $*$, is exclusive with regard to the other hypotheses of the frame of discernment which becomes exhaustive. The empty set $\oslash$ then represents the class of the conflict due to one or several non-reliable sources.

No source is associated with the hypothesis $*$. If this one appears during the fusion, the complementary of the hypothesis can be introduced:

$$\bar{H}_i = \{H_1, H_2, .., H_{i-1}, H_{i+1}, .., H_n\}. \quad (4)$$

In an application of object tracking, the reference set can be limited by putting the following constraint: *"a detected object can be put in relation only with one and a known object"* [Gruyer et al.(2002)]. The mass distribution is calculated to manipulate relations between an $i^{\text{th}}$ perceived object $\mathcal{X}_i$ and known objects $\mathcal{Y}_j$:

- the mass $m_{i,j}(\mathbf{Y}_j)$ is associated with the proposition *"the perceived object $\mathcal{X}_i$ is in relation with the known object $\mathcal{Y}_j$"*,
- the mass $m_{i,j}(\bar{\mathbf{Y}}_j)$ is associated with the proposition *"the perceived object $\mathcal{X}_i$ is not in relation with the known object $\mathcal{Y}_j$"*,
- the mass $m_{i,j}(\Theta)$ is associated with the total ignorance,
- the mass $m_{i,j}(*)$ is associated with the proposition *"the perceived object $\mathcal{X}_i$ is in relation with nothing known"*.

For example, a perceived object $\mathcal{X}_1$ can be associated with two known objects $\mathcal{Y}_1$ and $\mathcal{Y}_2$. In figure 2, we illustrate the following frame of discernment:

$$\Theta = \{\mathbf{Y}_1, \mathbf{Y}_2, *\}. \quad (5)$$

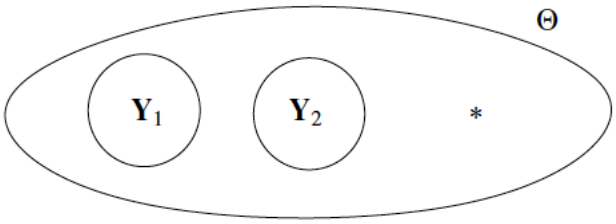



Figure 2: The extended open-world.

The hypothesis $*$ means that $\mathcal{X}_1$ is not locally associated with any known object. We can so build the following reference set:

$$2^{\Theta} = \{\oslash, *, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_1 \cup \mathbf{Y}_2, \bar{\mathbf{Y}}_2, \bar{\mathbf{Y}}_1, \Theta\} \quad (6)$$

We can remark that $\bar{\mathbf{Y}}_2 = \mathbf{Y}_1 \cup *$ and $\bar{\mathbf{Y}}_1 = \mathbf{Y}_2 \cup *$.

In case we have two hypotheses $\mathbf{Y}_1$ and $\mathbf{Y}_2$, we can calculate two distributions of the masses $m_{1,1}(.)$ and $m_{1,2}(.)$ which represent the relation between a perceived object $\mathcal{X}_1$ and two known objects $\mathcal{Y}_1$ and $\mathcal{Y}_2$. The combination is noted as $m_{1,12} = m_{1,1} \wedge m_{1,2}$.

The property of the operator associativity enables to combine the third source. The mass $m_{1,12}(*)$ is then affected in $m_{1,12}(* \cup \mathbf{Y}_3)$. In the case of $n$ sources, we can generalize the final mass distribution by using the properties of operator associativity and commutativity to represent the relation between the perceived objects $\mathcal{X}_i$ and the known objects $\mathcal{Y}_j$. The frame of discernment is thus:

$$\Theta = \{\mathbf{Y}_1, \mathbf{Y}_2, .., \mathbf{Y}_n, *\}. \tag{7}$$

The general form of mass distributions resulting in the combination is written as:

$$m_{i,.}(\mathbf{Y}_j) = m_{i,j}(\mathbf{Y}_j) \prod_{k=1,k\neq j}^{n} (1 - m_{i,k}(\mathbf{Y}_k)), \tag{8}$$

$$m_{i,.}(\bar{\mathbf{Y}}_j) = m_{i,j}(\bar{\mathbf{Y}}_j) \prod_{k=1,k\neq j}^{n} m_{i,k}(\Theta), \tag{9}$$

$$m_{i,.}(\bigcup_{j=p}^{q} \mathbf{Y}_j \cup *) = \prod_{k=p}^{q} m_{i,k}(\Theta) \prod_{k=1,k\neq[p,..,q]}^{n} m_{i,k}(\bar{\mathbf{Y}}_k), \tag{10}$$

$$m_{i,.}(*) = \prod_{j=1}^{n} m_{i,j}(\bar{\mathbf{Y}}_j), \tag{11}$$

$$m_{i,.}(\Theta) = \prod_{j=1}^{n} m_{i,j}(\Theta). \tag{12}$$

The product of masses on the hypotheses $\bar{\mathbf{Y}}_j$ corresponds to the non-modelled hypotheses. In the open-world, these hypotheses are allocated to the empty set $\oslash$ whereas in the extended open-world, the mass on hypothesis $*$ includes the non-modelled hypotheses. The mass associated with the empty set $\oslash$ thus corresponds to impossible cases and it is used during the decision-taking.

## 2.3 Generation of a mass distribution

### 2.3.1 Classical Models

Mass distributions allow to take into account all imperfections such as the incompletude, the imprecision or the uncertainty. These mass distributions have been developed to satisfy three properties: separability of the hypotheses evaluation, coherence with the Bayesian approach and coherence with the probability association of sources. L.M. Zouhal et al. developed in [Zouhal and Denœux(1998)] a generalization of the theory of beliefs of fuzzy sets to resolve classification problems. They used a set of rules which enables the construction of initial mass structures adapted to this problem. In [Rombaut and Berge-Cherfaoui(1997)], M. Rombaut has suggested a mass distribution which is inspired by the Denœux 's works [Denœux(1995)] constructed with help of a distance function $d_{i,j}$ between the perceived objects $\mathcal{X}_i$ and the known objects $\mathcal{Y}_j$. The usually used distances are the Euclidian distance or Mahalanobis distance. In case of information is fuzzy numbers, the proposed distance $d_{i,j}$ is based on the complement in '1' of a concordance degree.

### 2.3.2 A Mass Distribution Depending on Uncertainty

In our application, the extracted marker can be composed of some pixels. During a fuzzy modelling, the perceived objects $\mathcal{X}_i$ are much more precised than the predicted objects $\mathcal{Y}_j$ what implies that the support of $\mathcal{Y}_j$ can be much bigger than the one of $\mathcal{X}_i$. The possibilist or the Jacquard similarity measurement are not thus adapted to our application. Indeed, the first one can weaken the confidence on the perception which agrees the prediction, whereas the second one gives too much weight to the predicted objects and it discriminates the measurement. The idea is thus to build a mass distribution which takes into account only the prediction error and the hypothesis that a perceived object is maybe not associated and non-associated with a known object at the same time. The data association processing has to determine the optimal relation between the observation and the prediction to assign the measurement to an appropriate

existing track or to initialize a new track.

If the distance between objects is small or if the prediction error is high, the validation region can then overlap and several observations can be in this region. A perceived object can thus be a candidate with several tracks what generates conflict (cf. figure 3).

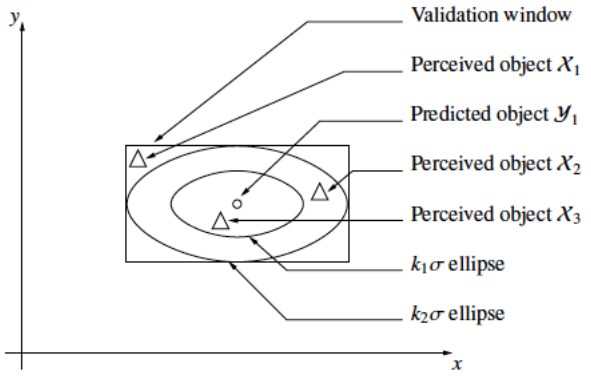



Figure 3: Three perceived objects are candidate with a predicted one.

The locus of points belonging to a confidence interval describes an ellipse (resp. ellipsoid) in the case of a Gaussian distribution with two (resp. three) random variables. The aim is to look for the relation between a predicted object and a perceived one belonging to the validation region. The computation of masses is based on the Rombaut's approach and it is made according to the position of perceived objects and the uncertainty ellipses. We define $k_1\sigma$ and $k_2\sigma$ uncertainty ellipses, $k_1 < k_2$, and we suggest the following mass distribution:

- case 1: the perceived object is inside of $k_1\sigma$ ellipse (object $\mathcal{X}_3$ in figure 3), *it is in relation with the predicted object*;
- case 2: the perceived object is outside of $k_2\sigma$ ellipse (object $\mathcal{X}_1$ in figure 3), *it is not in relation with the predicted object*;
- case 3: the perceived object is between $k_1\sigma$ ellipse and $k_2\sigma$ ellipse (object $\mathcal{X}_2$ in figure 3). *We do not know if there is a relation between both objects.*

To simplify writings, we note $X_i$ the coordinates of the i$^{\text{th}}$ perceived object $\mathcal{X}_i$, $Y_j$ the coordinates of the j$^{\text{th}}$ predicted object $\mathcal{Y}_j$ and $\Sigma_{X_iY_j}$ the nonempty prediction error matrix.

We note $d_{i,j,k\Sigma}$ the cartesian distance between the position of a predicted object and the uncertainty ellipse in the direction of the perceived object and $d_{i,j}$ the cartesian distance between the perceived object and the predicted one.

We propose to define the mass distribution, resulted in the association between a perceived object $\mathcal{X}_i$ and a predicted one $\mathcal{Y}_j$ as following:

$$m_{i,j}(\mathbf{Y}_j) = \begin{cases} \frac{\alpha_0}{2}\left(1+\sin\left(\frac{\pi(1+2d_{i,j})}{2d_{i,j,k_1\Sigma}}\right)\right) & \text{if case 1,} \\ 0 & \text{if cases 2, 3,} \end{cases} \quad (13)$$

$$m_{i,j}(\bar{\mathbf{Y}}_j) = \begin{cases} 0 & \text{if case 1,} \\ \frac{\alpha_0}{2}\left(\sin\left(\frac{\pi}{2}\frac{d_{i,j}-\frac{\Delta d}{2}}{\frac{\Delta d}{2}}\right)-1\right) & \text{if case 2,} \\ 1 & \text{if case 3,} \end{cases} \quad (14)$$

$$m_{i,j}(\Theta) = \begin{cases} 1-m_{i,j}(\mathbf{Y}_j) & \text{if case 1,} \\ 1-m_{i,j}(\bar{\mathbf{Y}}_j) & \text{if cases 2, 3,} \end{cases} \quad (15)$$

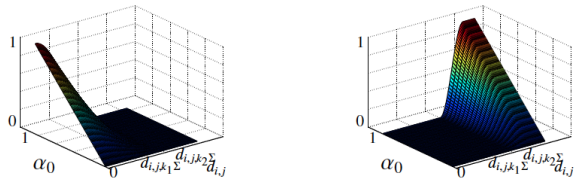
where $\alpha_0 \in [0,1]$ quantifies the sensor reliability, $k_1 < k_2$ and $\Delta d = d_{i,j,k_2\Sigma} - d_{i,j,k_1\Sigma}$.

Figure 4 illustrates the distribution of these masses according to the reliability factor and the distance between the perceived object and the predicted one. If the perceived object is inside of the first uncertainty ellipse and it tends toward the predicted object, the mass $m_{i,j}(\mathbf{Y}_j)$ tends toward the reliability factor. In the same way, if $\mathcal{X}_i$ is outside the second ellipse, the mass $m_{i,j}(\bar{\mathbf{Y}}_j)$ is equal to the reliability factor and $m_{i,j}(\mathbf{Y}_j)$ is null. Furthermore, if the reliability factor is lower than 0.5, the mass on the hypothesis $\Theta$, *"we do not know"*, is superior to the other masses.

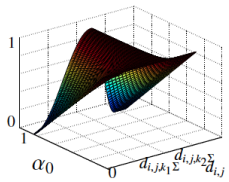
# 3 Multi-object Tracking

The tracking filter has two main roles: estimating the state of an entity from data provided by sensors and tracking an entity.

Predictions and estimations of object states are calculated with help of an adaptative Kalman filter.



(a) $m_{i,j}(\mathbf{Y}_j)$. (b) $m_{i,j}(\bar{\mathbf{Y}}_j)$.



(c) $m_{i,j}(\Theta)$.

Figure 4: A mass distribution in function of the distance between a perceived object and a predicted one and the sensor reliability factor $\alpha_0$.

The proposed object tracking method consists of five stages (cf. figure 5): prediction of object states, mass distribution, mass combination, decision and estimation of object states.

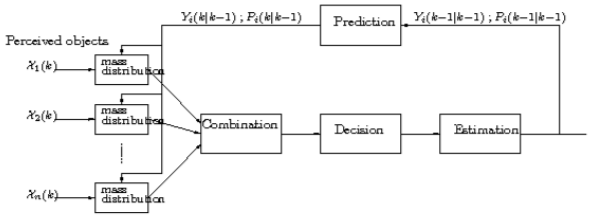



Figure 5: The principle of object tracking.

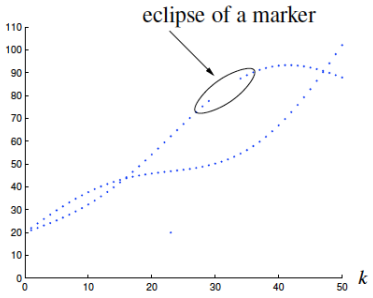
## 3.1 Illustration of Object Pursuits

We synthetically generate two tracks (cf. figure 6(a)). At time $k$, data measurements are associated with tracks. Figure 6(b) represents the temporal evolution of the mass distribution in the set $\Theta_{1,1}$ and figure 6(c) in the set $\Theta_{1,2}$. The dynamic behaviour of mass distributions is illustrated with help of three cases:

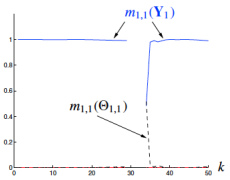
- object disappearance and reappearance: when the observation $\mathcal{X}_1$ is missing ($30 < k < 33$), the mass distribution in the set $\Theta_{1,1}$ cannot be calculated. The prediction $\mathcal{Y}_i$ at the following horizons is therefore calculated so that to possibly re-associate the track with the reappearing object $\mathcal{X}_1$. If this prediction is in relation with the observation, the mass $m_{1,1}(\mathbf{Y}_1)$ becomes no null.

- crossing of trajectories: at time $k = 16$, the perceived object $\mathcal{X}_2$ cannot be candidate with the association of the track $\mathcal{Y}_1$. The masses $m_{1,2}(\mathbf{Y}_2)$ or $m_{1,2}(\Theta_{1,2})$ are thus no null.

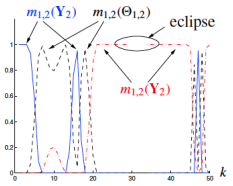
- appearance of a third observation: at time $k = 23$, a new object $\mathcal{X}_3$ appears in the scene, the masses $m_{1,3}(\bar{\mathbf{Y}}_3)$ and $m_{2,3}(\bar{\mathbf{Y}}_3)$ are naturally equal to $\alpha_0$.




(a) Perceived objects in function of the instants $k$.




(b) Relation with the perceived object $X_1$

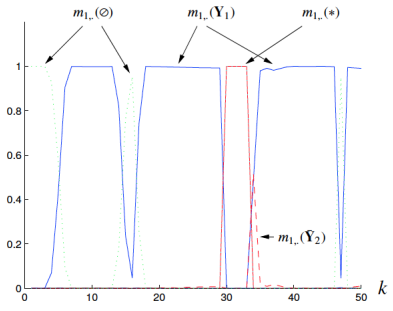



(c) Relation with the perceived object $X_2$

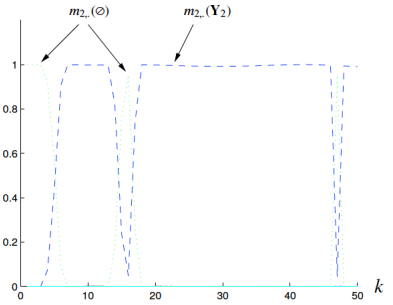
Figure 6: The mass distribution associated with the relation between the predicted object $Y_1$ and both perceived objects.

## 3.2 Dynamic Evolution of the Mass Distribution in the Extended Open-World

We combine mass distributions in the extended open world to look for the relation between the perceived objects and the predicted ones. Figures 7(a) and 7(b) illustrate the mass distributions which are the result of the combination.




(a) The relation with the predicted object $Y_1$.




(b) The relation with the predicted object $Y_2$.

Figure 7: Combination results: association with both predicted objects.

At the beginning of an acquisition or during a crossing of two objects, we have a conflict although the mass $m_{1,\cdot}(\mathbf{Y}_1)$ is superior to $m_{1,\cdot}(\mathbf{Y}_2)$. This conflicting mass $m_{i,\cdot}(\oslash) \neq 1$ can be used to normalize the mass distribution during the decision-taking:

$$k_i = \frac{1}{1 - m_{i,\cdot}(\oslash)}. \tag{16}$$

When the perceived object $\mathcal{X}_1$ disappears, the mass on the hypothesis $*$ is close to one; it means that no perceived object can be associated with the predicted object $\mathcal{Y}_1$. There are masses on the both hypotheses *"$\mathcal{X}_1$ is in relation with $\mathcal{Y}_1$"* and *"$\mathcal{X}_2$ is not in relation with $\mathcal{Y}_1$"*. It is thus necessary to define a method to take decision.

## 3.3 Choice of Decision

The combination in the extended open-world generates masses on the union of hypotheses. We have to take into account all information to take decision.

When the perceived object $\mathcal{X}_1$ appears, the piece of information restrained in masses on unions of hypotheses is primordial, for example, it is necessary to take into account the mass $m_{1,\cdot}(\bar{\mathbf{Y}}_2)$ in the degree of belief on the hypothesis $\mathbf{Y}_1$. Consequently, we maximize the pignistic probability based on the equidistribution of hypotheses consisting to distribute the mass of ignorance on the other singleton hypotheses.

## 3.4 Propagation of Tracks

### 3.4.1 Initialisation of a track

Let be a marker appears in the scene and it is not associated with existing tracks. During the first instants, all perceived markers can stand as candidate for the association. A combinatorial tree is so constructed and each branch corresponds to the association between a perceived marker and a predicted one. These propagation lasts some instants before to take a decision.

### 3.4.2 Disappearance of a Track

A marker can disappear because of a marker eclipse by a sportsman limb or because of an information extraction error. It can be considered as disappeared if the predicted marker is not associated with a perceived one. Consequently, the final decision stands as the hypothesis $*$: it is not thus modelled in the frame of discernment.

A disappeared marker can later reappear, information on the track is so propagated during a limited

number of samples to be associated with new markers.

### 3.5 Illustration of Tracking Performances

Markers are extracted from images with histogram analyses and connected component analyses. Their trackings are illustrated at figure 8.

Moreover, the error between the perception and the estimation enables to justify the choice of a Kalman filter. The mean error of the position abscissas is 0.0084 the standard 0.1663 and the mean error of the position ordinates is 0.0118, the standard 0.2371. These errors are satifactory to explore results in models for sportive gesture analysis.

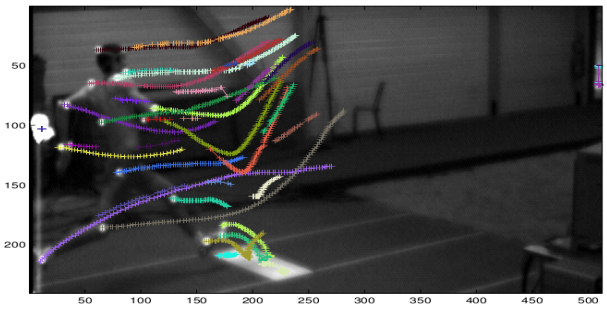



Figure 8: Markers tracking on a sensor.

## 4 Reconstruction of Three Dimensional Markers

### 4.1 Stereoscopic Reconstruction Method

The video acquisition system used for the sportive gesture analysis is composed of five CCD sensors with a resolution 512(H)×256(V) pixels, 250 frames per second and 256 grey scales. These cameras are placed to obtain all sportive gestures of a long-jump athlete. It is theorically possible to make ten different stereoscopic reconstructions with five sensors ($C_{n_s}^2$, with $n_s$ sensors).

The stereoscopic reconstruction requires to know the position of both cameras, the focal length and the transformation between the scene and the image [Faugeras(1993)]. Certains of these parameters are bound by the camera. So, camera internal parameters, called intrinsic parameters, depend on the focal length and the optical center position, whereas the extrinsic parameters characterize the transformation between the scene and the camera. To characterize the camera, we have chosen a simplified model, called pinhole model. This model consists of an optical center, an image center and all light rays propagate along a line and they linearly project onto the image plane.

In the first time, we compute the geometric model of each camera with an unique calibration pattern in using the classical geometric tools in computer vision. In help of the parameters of the right and left cameras, we determine the transformation matrix between the left camera space and the right camera space. Finally, the epipolar geometry enables us to characterize relationships between two stereoscopic images by the standardization knowledge on the one hand and to reduce the searching zone of the corresponding point on the other hand. The principle of the epipolar geometry consists in etablishing the homogeneous of a left (resp. right) image point which is on a known line in the right (resp. left) image named epipolar line.

### 4.2 Principle of the Marker Stereoscopic Reconstruction

At time $k$, a 3D trajectory of a marker is reconstructed in using the label of its projections at time $k-1$. However, it is necessary to valid the epipolar constraint. We have to initialize the 3D tracks $\mathcal{A}_i^{l,r}$ from $l^{th}$ and $r^{th}$ sensors, noted $S_l$ and $S_r$, with $l \neq r$, from 2D non-associated markers.

This method is easy, but it is necessary to fuse all reports issued from stereoscopic reconstructions to correct tracks and to obtain those are perceived by other sensor couples.

## 4.3 Correction and Fusion of Three Dimensional Reconstructed Markers

### 4.3.1 Covariance Intersection Algorithm

It concerns to combine $n$ estimates with the mean value of a random variable when the correlation between estimates is unknown. We note the estimated statistical mean of $a_i$ , $\hat{a}_i$ and the deviation $\tilde{a_i} \triangleq a_i - \hat{a}_i$. We can also note that correlations $\tilde{P}_{ii} = E\{\tilde{a_i}\tilde{a_i}^T\}$ can be unknown but the estimates are consistant and cross-correlations $\tilde{P}_{ij} = E\{\tilde{a_i}\tilde{a_j}^T\}$, $\{(i,j) \in |[1,n]|^2, i \neq j\}$, can so be unknown and they are not equal to 0.

The objective is to fuse the $n$ consistant estimates of $\mathcal{A}_i$ to build a new consistant estimate $\mathcal{C}\{c, P_{cc}\}$ which is linear and no biaised. The CI results in the geometric interpretation of the Bar-Shalom equation [Bar-Shalom and Campo(1986)] which consists in combining linearly the statistical means and in determinating analyticaly the covariance of the result. The Covariance Intersection algorithm (or CI) is a data fusion algorithm with a convex combination of means and covariances [Julier and Uhlmann(1997)]. If $P_{ii}$ are fixed, the covariance ellipse of $P_{cc}$ always stays in the intersection of covariance ellipses of $P_{ii}$ for different choices of $P_{ij}$ [Uhlmann(1996)].

The combination of the $n$ reports $\mathcal{A}_i\{a_i, P_{ii}\}$ creates a new estimates $\mathcal{C}\{c, C\}$ defined by:

$$P_{cc}^{-1} = \left(\sum_{i=1}^{n} \omega_i P_{ii}^{-1}\right) \text{ and } P_{cc}^{-1} c = \left(\sum_{i=1}^{n} \omega_i P_{ii}^{-1} a_i\right), \tag{17}$$

where $a_i$ is the statistical mean, $P_{ii}$ the correlation of the $i^{th}$ estimates and $\sum_{i=1}^{n} \omega_i = 1$. In [Julier and Uhlmann(1997)], it is demonstrated that the update equation is consistant for all values of $P_{ij}$ and $\omega$, $\{(i,j) \in |[1,n]|^2, i \neq j\}$.

### 4.3.2 Principle of Three dimensional Stereoscopic Report Correction

*Notation :* Only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state. In what follows, the $a(t_1|t_2)$ represents the estimate of $a$ at time $t_1$ given observations up to, and including time $t_2$. The main aims are to fuse, at time $k$, the $n$ 3D marker reports $\mathcal{A}_i(k|k)$ of statistical mean $a_i(k|k)$ and covariance $P_{ii}(k|k)$, noted $\mathcal{A}_i(k|k)\{a_i(k|k), P_{ii}(k|k)\}$, $i \in |[1,n]|$, and so, to correct their tracks for each stereoscopic reconstruction and to produce final tracks $\mathcal{C}_i(k|k)\{c(k|k), P_{cc}(k|k)\}$.

The i$^{\text{th}}$ tracked marker $\mathcal{Z}_i^{l,r}$ has so corresponding to the fusion of 3D tracked markers $\mathcal{A}_i^{l,r}$ issued of stereoscopic reconstructions with sensors $S_l$ and $S_r$. The principle thus consists in correcting the prediction report of a marker in fusing it with the partial estimates of other markers:

- Prediction of the 3D marker, noted $\mathcal{A}_i^{l,r}(k|k-1)$;
- Partial estimations of the other markers, noted $\mathcal{A}_i^{g,d;p}(k|k)$, $\forall d \neq r$ and $\forall g \neq l$, in using the reconstructed positions of $\mathcal{Z}_i^{l,r}(k|k-1)$;
- Fusion of the prediction $\mathcal{A}_i^{l,r}(k|k-1)$ with partial estimations $\mathcal{A}_i^{g,d;p}(k|k)$ using CI algorithm; the result is noted $\mathcal{A}_i^{l,r;*}(k|k-1)$;
- Update of $\mathcal{A}_i^{l,r,*}(k|k-1)$ with the position of $\mathcal{Z}_i^{l,r}(k|k-1)$; the final estimation is naturally noted $\mathcal{A}_i^{l,r}(k|k)$.

In the CI algorithm, the weights of each track are computed from the degree of belief on the hypothesis *"the reconstructed object is in relation with the predicted object"*.

### 4.3.3 Fusion of 3D stereoscopic tracks

We use the computation method of mass distribution in section 2.3.2 to associate measurements with tracks:

- computation of mass distributions $m_{i,j}(\mathbf{Y}^{l,r})$ to associate 3D markers $\mathcal{A}_i^{l,r}(k|k-1)$ with those from reconstructions $\mathcal{Z}_i^{l,r}(k|k-1)$,
- mass distribution in the extended open-world.

Reliability coefficients interviening in the computation of mass distributions are fixed for each stereoscopic reconstruction in empiric way: they correspond to probabilities that a marker from the calibration pattern is projected onto the corresponding extracted interval in the image.

At time $k$, the computation of weights $\omega_i^{l,r}$ associated with the prediction of the final track $\mathcal{Z}_i(k|k-1)$ is given by:

$$\omega_i^{l,r} = \frac{m_{i,i}(\mathbf{Y}_i^{l,r})}{\sum_{g=1}^{n}\sum_{d=1}^{n} m_{i,i}(\mathbf{Y}_i^{g,d})}, \qquad (18)$$

with $\Theta = \{\mathbf{Y}_i^{l,r}, \bar{\mathbf{Y}}_i^{l,r}\}$.

The degree of belief can be view as the total belief allocating at the hypothesis considering available information. If the mass $m_{i,j}(\mathbf{Y}^{l,r})$ is null, $\omega_i^{l,r}$ is then null. The partial estimation of $\mathcal{A}_i^{l,r;p}(k|k)$ does not consequently influence the final estimation of other markers in relation with $\mathcal{Z}_j(k|k-1)$, $j \neq i$.

We have opted for a convex combination with the CI algorithm to fuse data, each track is so corrected in taking account of the marker tracking report of the other stereoscopic reconstruction. At time $k$, these tracks can be fused to obtain only an unique track for each marker. The CI of the estimates $\mathcal{A}_i^{l,r}(k|k)$ enables to obtain the final estimate, noted $\mathcal{Z}_i(k|k)$ [Helbert et al.(2004)].

Non fused markers are combined to initialize new tracks in testing the uncertainty ellipsoid intersection.

#### 4.3.4 Illustrations of Fusion Performances

We have applied our algorithm to reconstruct marker trajectories with data coming from five cameras (cf. figure 9). Figure 10(a) represents an extract of a track corresponding to the movements of a marker from three stereoscopic reconstructions and their fusion.

Figure 10(b) illustrates the interest of the redundancy. At the beginning of the tracking, magenta track is close to the real marker position in the scene (blue points) contrary to the track resulting in the fusion (red line). One of sensors provides a position error in the 2D tracking which affects on the magenta reconstructions and yellow ones. The error on the fusion reduced by the belief influency of the yellow track and of the green one.

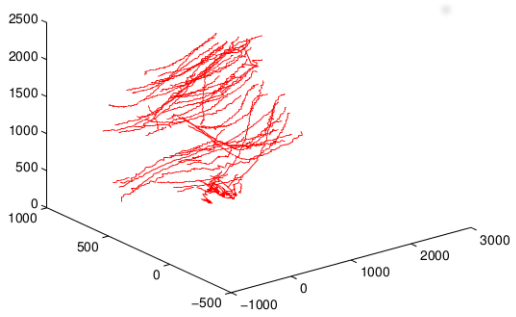
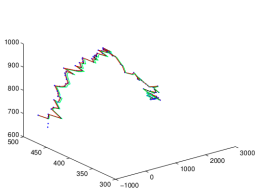
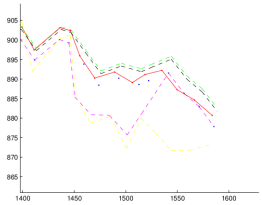



Figure 9: Fusion of stereoscopic reconstructions.



(a) Tracks in the scene



(b) The influency of the fusion

Figure 10: Results of the fusion with the CI algorithm. In figure 10(b), the blue points characterize the real marker position in the scene, the dash tracks the stereoscopic reconstructions and those in red the result of the fusion)

## 5 Conclusion and Perspectives

The proposed method enables to reconstruct the 3D marker trajectories without *a priori* knowledge such as the number of markers or their dynamic behaviour. The Dempster-Shafer theory of Evidence is a particularly interesting tool for information fusion, it is then easy to handle the marker appearances and the disappearances.

The 3D reconstruction takes into account the

tracking results in different sequences of images to reconstruct the 3D marker trajectories. These are then corrected and fused with the help of the covariance intersection algorithm. This last one has been chosen to fuse different update reports with different measurements, for example position or velocity. Moreover, the combination is adapted to the fusion of estimates when the correlation is unknown.

At last, this method privileges the redundancy of information to correct markers trajectories and to track markers which are not perceived by certain sensors.

At the present time, we are developing a real time system of 3D movement reconstruction in collaboration with mechanics. The aim is to propose a new tool for the 3D animation and the gesture analysis.

# References


[Bar-Shalom and Campo(1986)] Bar-Shalom, Y., and Campo, L. (1986), "The effect of the common process noise on the two-sensor fused track covariance," *IEEE trans. on Aerosp. Electron. Syst.*, AES(22), 803–825.

[Dempster(1967)] Dempster, A. (1967), "Upper and Lower Probabilities Induced by a Multi-Valued Mapping," *Annals of Mathematical Statistics*, 38.

[Denœux(1995)] Denœux, T. (1995), "A k-nearest Neightbor Classification Rule Based on Dempster-Shafer Theory," *IEEE Transactions on Systems, Man and Cybernetics*, 25(5), 804–813.

[Faugeras(1993)] Faugeras, O., *Three-Dimensional Computer Vision: a geometric view point*, Cambridge, Massachusetts, USA: MIT Press (1993).

[Gruyer et al.(2002)] Gruyer, D., Mangeas, M., and Royère, C. (2002), "A new approach for credibilitic multi-sensor association," in *International Conference on Information Fusion*.

[Gruyer et al.(2000)] Gruyer, D., Royère, C., and Cherfaoui, V. (2000), "Credibilistic multi-sensors fusion for the mapping of dynamic environment," in *3rd International Conference on Information Fusion*, Paris.

[Hall and Llinas(2001)] Hall, D., and Llinas, J., *Handbook on Multisensor Data Fusion*, crc press inc ed., Collection : Electrical Engineering and Applied Signal Processing (2001).

[Helbert et al.(2004)] Helbert, D., Augereau, B., and Fernandez-Maloigne, C. (2004), "Information Fusion using Covariance Intersection in the frame of the Evidence Theory," in *Advanced Concepts for Intelligence Vision Systems*, september, Brussels, Belgium.

[Julier and Uhlmann(1997)] Julier, S., and Uhlmann, J. (1997), "Non-divergence estimation algorithm in the presence of unknown correlations," in *The American Control conference*, Vol. 4, Piscataway, USA: IEEE, pp. 2369–2373.

[Lemeret et al.(2008)] Lemeret, Y., Lefevre, E., and Jolly, D. (2008), "Improvement of an association algorithm for obstacle tracking," *Information fusion*, 9, 234–245.

[Mourllion et al.(2005)] Mourllion, B., Gruyer, D., Royère, C., and Theroude., S. (2005), "Multi-Hypotheses Tracking Algorithm Based on the Belief Theory," in *IEEE ISIF*, July, Philadelphie, Pennsylvanie, USA: IEEE.

[Rombaut(1998)] Rombaut, M. (1998), "Decision in Multi-Obstacle Matching Process Using Theory of Belief," in *AVCS'98, Amiens, France*, 1-3 juillet.

[Rombaut and Berge-Cherfaoui(1997)] Rombaut, M., and Berge-Cherfaoui, V. (1997), "Decision Making in Data Fusion using Dempster-Shafer's Theory," in *3th IFAC Symposium on Intelligent Components and Instrumentation for Control Applications, Annecy, France*, 9-11 juin.

[Royère et al.(2000)] Royère, C., Gruyer, D., and Cherfaoui, V. (2000), "Data association with believe theory," in *IEEE FUSION*, Paris, France.

[Shafer(1976)] Shafer, G., *A Mathematical Theory of Evidence*, Princetown University Press (1976).

[Smets(1990)] Smets, P. (1990), "The combination of evidence in the transferable belief model," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 12(5), 447–458.

[Uhlmann(1996)] Uhlmann, J. (1996), "General data fusion for estimates with unknown cross covariances," in *SPIE Aerosense Conference*, April.

[Zouhal and Denœux(1998)] Zouhal, L.M., and Denœux, T. (1998), "An Evidence-Theoretic k-NN rule with Parameter Optimization," *IEEE Transactions on Systems, Man and Cybernetics - Part C*, 28(2), 263–271.