

A Support Method for the Contextual Interpretation of Biomechanical Data

Emmanuel Roux, Anne-Pascale Godillon-Maquinghen, Patrice Caulier, Stéphane Bouilland, and Denis Bouttens

Abstract—In the clinical field, biomechanical data provided by advanced technical devices are still underexploited. Data analysis usually consists of extracting attributes or computing synthetic values from temporal data and exploiting them by means of a monovariate statistical method. This article proposes a method to support clinicians, especially those in orthopedics, in the contextual interpretation of biomechanical data. We propose to characterize temporal biomechanical data by means of fuzzy space–time windows and to induce fuzzy decision trees to map the biomechanical and clinical data related to patients. Then, we present a method for objectively explaining a given clinical characteristic of a particular patient; this method is derived using the fuzzy rule base generated from the trees and a satisfiability measure. We have applied our method to real data in order to provide an objective explanation of the subjective self-evaluation of the functional status of patients with a shoulder prosthesis, and evaluate it by means of the stratified tenfold cross validation method. The mean explanation rate—which corresponds to the mean proportion of the patients belonging to test sets whose functional state is explained by the proposed method—exceeds 80% for more than half of the decision trees, and exceeds 70% for 94% of the trees. By supporting clinicians during the biomechanical data interpretation process, our method helps them take the objective biomechanical measurements in the medical practice into account, particularly in orthopedics. It can also make subjective evaluations more objective by mapping subjective and objective data.

Index Terms—Functional evaluation, fuzzy decision trees, fuzzy rules, objective explanation, stratified tenfold cross validations, three-dimensional (3-D)-movement measurement.

I. INTRODUCTION

BIOMECHANICAL measurement techniques have developed considerably over the past twenty years. These techniques, which permit easy, reliable, three-dimensional (3-D) measurements of movements, accelerations and forces, have principally been applied to lower limb and gait analysis and have become quite effective in these application domains over the last two decades. Medical practitioners, particularly in orthopedics, immediately sensed the benefits that could be derived from using these tools to objectively evaluate treatment results or to improve diagnosis methods, for example.

Manuscript received December 21, 2004; revised May 3, 2005. This work is sponsored by the Centre National de la Recherche Scientifique (C.N.R.S.), the Nord-Pas de Calais Region and the Institut Régional de Recherche sur le Handicap (I.R.R.H.).

E. Roux is with the LTSI, INSERM U642, Rennes, France (e-mail: emmanuel.roux@univ-rennes1.fr).

A.-P. Godillon-Maquinghen and P. Caulier are with the LAMIH, UMR CNRS 8530, Valenciennes, France (e-mail: anne-pascale.maquinghen@univ-valenciennes.fr; patrice.caulier@univ-valenciennes.fr).

S. Bouilland and D. Bouttens are with the Groupe Hopale, Berck/Mer, France (e-mail: sbouilland@hopale.com; dbouttens@hopale.com).

Digital Object Identifier 10.1109/TITB.2005.856861

However, biomechanical data are still underexploited in the clinical field. In fact, the temporal data sets provided by biomechanical measurements are relatively large and consequently pertinent information is often difficult to extract from the data. In addition, most clinicians are unfamiliar with the use of these sets and do not make the most out of the information available. Consequently, clinical reports take a lot of time to be generated and provide too many and unspecific information. Thus, gait analysis laboratories are often viewed as inefficient, unproductive and uneconomical [1]. Moreover, in the clinical field, data analysis usually consists of extracting attributes or computing synthetic values from the temporal data and exploiting them by means of a monovariate statistical method [2], [3]. Data relative to several trials of a given subject are usually averaged, preventing the representation of multimodal phenomena, resulting, for example, from different movement strategies of the subject. All of these data analysis methods, intended to give clinicians easily intelligible, interpretable and exploitable values, are well known, but provide limited information.

However, Artificial Intelligence offers a variety of techniques liable to help physicians interpret and exploit biomechanical data in a clinical context. Chau [4], [5] presents a review of advanced data analysis methods within gait analysis. Neural Networks take a large part in this review. According to Chau [4], [5], other methods: fuzzy clustering, multiple correspondence analysis, fractal methods, wavelets and time-frequency analysis, are marginal and have to be further evaluated within the applications discussed in his paper. Within movement analysis, neural networks are used for movement simulation, notably in order to obtain simulated data of unobservable variables [5], for prognosis, by predicting postoperative gait characteristics according to preoperative data [6], or classification, for a diagnosis support for example. However, due to the “black-box” properties of neural networks, it is difficult to provide a biomechanical interpretation of the input–output relationships. Dr. Gait III and its predecessor QUAWDS (Qualitative Analysis of Walking DisorderS) are knowledge-based systems that exploit pathological gait data [1], [7]. These support systems combine a qualitative model of the “normal” gait and an association model defined by means of a rule base. These two models are derived from the background knowledge of the domain and from more specific experts’ knowledge. In our case, i.e., within upper limb analysis, it seems difficult to describe a “normal” movement given the wide variety of functions the upper limb fulfills, the high redundancy of its kinematic chain and consequently the numerous movement strategies it offers. Kuchar *et al.* [8] propose a case-based reasoning system within

gait analysis. They use data provided by an accelerometer and data relative to age, sex, etc. A rule base is used to provide a brief interpretation of the data, and then similar case(s) are retrieved from the case base. Besides methodological difficulties associated with case-based reasoning, i.e., the case base completeness, the case model, the definition of the similarity measure and the adaptation of the similar case, case-based reasoning does not automatically extract general knowledge from data.

In this paper, we first propose a method to automatically extract general knowledge from biomechanical data in a way that the data characterizing and coding, the data analysis method and the results can be easily understood and interpreted by the user. We propose to use fuzzy decision trees. To our knowledge, it is the first time fuzzy decision trees are applied within movement analysis. This general knowledge is then used to assist clinicians in the interpretation and the exploitation of biomechanical data. The aim of our procedure is to explain, for a given patient, the presence or the absence of a given clinical characteristic, using the available biomechanical data on this patient and an abductive reasoning, exploiting a fuzzy rule base.

The paper is organized as follows. First, we briefly describe the biomechanical data and explain our statistical data characterizing and coding methods, using fuzzy space–time windowing. Then, we present the automatic mapping of the biomechanical and clinical data relative to the patients, using “forests” of fuzzy decision trees induced from sets of examples. The rule base, derived from the trees, permits the clinical state of a given patient to be explained in terms of his/her biomechanical characteristics, by means of a satisfiability measure.

Our methods are then applied to obtain the objective explanation of the subjective self-evaluation of the functional state of patients with a shoulder prosthesis. Eventually the performances of the explanation method are evaluated using an explanation rate and a stratified tenfold cross validation. There are both biomechanical and medical reasons for developing this application. Although technological and methodological improvements in biomechanical measurement procedures have recently been applied to upper limb biomechanical analysis [9], the data obtained have often been difficult to interpret. These interpretation difficulties stem partially from the newness of the techniques, but also from the complexity of upper limb biomechanics that is due, in part, to the redundancies of the kinematic chain. In addition, shoulder arthroplasty is a relatively new field of medicine. Given the significant increase of the number of interventions since the 1990s, a method for objectively evaluating results has become necessary, particularly in view of the current social and juridical pressure put on surgeons. So the present work is original for two reasons: it first presents advanced Artificial Intelligence methods that, to our knowledge, have not been applied to orthopedics yet; second, these methods are used within the analysis of the upper limb biomechanics and pathologies, which have been understudied until recently.

II. METHODS

A. Characterizing and Coding Biomechanical Data

Traditionally, biomechanical data include kinematic data obtained, for example, by means of a 3-D optoelectronic system; acceleration data provided by position data derivatives or by accelerators; and force measurements produced by force plates. These data provide temporal and 3-D information. For a given subject, such data can be obtained simultaneously, for several experimental conditions and for several trials under the same experimental conditions. These data are related to measurement variables that are not always directly exploitable. In fact, analysis variables, such as hip flexion or acceleration of the body’s center of mass, have to be derived from the measured variables, in order to provide an intelligible description, in a biomechanical context, of the phenomenon studied.

These analysis variables must then be statistically characterized and coded [10]: characterized in order to summarize the data, and coded to make them homogeneous and compatible with a particular statistical method. However, biomechanical data, and especially movement measurements by means of optoelectronic systems and external markers, are very sensitive to both measurement artifacts [11]–[14] and the poor repeatability of the human movement. In this context, fuzzy sets theory appears to be an appropriate tool for dealing with the variability and the inaccuracy of the biomechanical data in one hand, and the characteristics of human reasoning on the other hand. For these reasons, we propose to characterize and to code biomechanical data by computing their membership values with regard to fuzzy space–time windows.

1) *Space–Time Windows*: Space–time windowing was proposed by Loslever and Bouillard [10] for characterizing and coding biomechanical data and for use in Multiple Correspondence Analysis. These characterizing and coding steps are performed simultaneously in space–time windowing. The difference between Loslever *et al.*’s work [10] and ours lies in the choice of the data analysis method applied and in the definition of the space–time windows that, in our application, are linked to the functional characteristics of human movement (see Section III).

The time domain of each analysis variable is considered through a fuzzy window set $T = \{T_1, \dots, T_j, \dots, T_{N_T}\}$ such that $\mu_{T_j}(t_q)$, i.e., the membership value of the q th time sample t_q , falls between $[0, 1]$ and meets the condition $\sum_j \mu_{T_j}(t_q) = 1$. In the same way, the range of amplitudes (spatial domain) of each analysis variable is considered through N_A spatial fuzzy windows with the same properties.

Of course, in the temporal and the spatial domains, the choice of the shape and number of the fuzzy subsets depends on both the application and the objectives of the study. These issues will be described and explained in the application section.

2) *Membership Value Computation*: Let $V_n(t_q)$ be the value taken by the n th analysis variable at time unit t_q .

Let $\mu_{T_j}(t_q)$ (respectively, $\{\mu_{A_{i,n}}(V_n(t_q))\}$) be the membership value of the j th time window T_j for the q th time unit (respectively, of the i th space window $A_{i,n}$ for the $V_n(t_q)$ value).

TABLE I
PART OF THE LEARNING SET AND NOTATIONS USED FOR THE DECISION TREES INDUCTION WITHIN OUR APPLICATION

	Kinematics data					Functional evaluation	
	Variables	V ₁ :Elbow flexion			... V _n	To do ones hair	
	Fuzzy time windows	T ₁ :“Starting”			... T _i	“Impossible”	“NOT-Impossible”
	Fuzzy space windows	A _{1,1} :“Low”	A _{2,1} :“Medium”	A _{3,1} :“High”	... A _{j,n}		
Example 1 (Patient 1)	μ _{S₁⁽¹⁾(1) = 0.53}	μ _{S₂⁽¹⁾(1) = 0.47}	μ _{S₃⁽¹⁾(1) = 0.00}	... μ _{S_j⁽¹⁾(1)*...}	C ₁ (1) = 1	C ₂ (1) = 0	
Example 2	μ _{S₁⁽¹⁾(2) = 0.42}	μ _{S₂⁽¹⁾(2) = 0.35}	μ _{S₃⁽¹⁾(2) = 0.23}	... μ _{S_j⁽¹⁾(2) ...}	C ₁ (2) = 0	C ₂ (2) = 1	
...	
Example k	μ _{S₁⁽¹⁾(k)}	μ _{S₂⁽¹⁾(k)}	μ _{S₃⁽¹⁾(k)}	... μ _{S_j⁽¹⁾(k) ...}	C ₁ (k)	C ₂ (k)	

V_n is the analysis variable n, T_i the fuzzy time window i and A_{j,n} the fuzzy space window j related to the variable n.

μ_{S_i^(l)(k) is the membership value of the fuzzy subset S_i^(l) for the kth example (patient) and the lth linguistic variable* (see section II.A. for details).}

* In our case, as N_A is the same for all the variables, we have l = (n - 1) · N_A + j, i.e., as N_A = 3, l = (n - 1) · 3 + j.

The membership value of the space–time window W_{j,i}ⁿ for a given time series (or signal) related to the analysis variable V_n, is defined as in Loslever and Bouilland [10]

$$\mu_{W_{j,i}^n} = \frac{1}{\sum_{q=1}^Q \mu_{T_j}(t_q)} \cdot \sum_{q=1}^Q \mu_{T_j}(t_q) \cdot \mu_{A_{i,n}}(V_n(t_q)) \quad (1)$$

with Q being the number of time units, and μ_{W_{j,i}ⁿ, the weighted average of the space membership values with the time membership values as weights. Note that for a given variable V_n and a given temporal window j, μ_{W_{j,i}ⁿ verifies: ∑_i μ_{W_{j,i}ⁿ = 1. This property is required to maintain the statistical context and to allow μ_{W_{j,i}ⁿ to be interpreted as the frequency of the signal’s appearance in the space–time window W_{j,i}ⁿ.}}}}

N_V and N_T are the number of analysis variables and the number of time windows, respectively. The universe of objects is described by N_L = N_V · N_T linguistic variables (or attributes) L^(l) (l ∈ [1, N_L]). Each attribute L^(l) is divided into N_A fuzzy subsets S_i^(l) (i ∈ [1, N_A]), corresponding to the space windows. We note μ_{S_i^(l) the membership value of the fuzzy subset S_i^(l) for a given time series related to the lth variable. An example of such a fuzzy coding, with the notations used, is presented in Table I in Section III.}

B. Fuzzy Decision Trees (FDT), Forests of FDT

Some fuzzy decision tree (FDT) induction methods join the tree (FDT) building and the computation of the membership values. In Boyen *et al.* [15] for example, attributes are not fuzzified beforehand but the tree is induced by finding, at each node, a locally optimal fuzzy dichotomy of the subset of examples. Parameters of the tree are adjusted afterwards in order to find a global optimum. In our case, finding optimal fuzzy subsets would provide rules difficult to read and interpret by the users. To facilitate the interpretation of the results, we prefer to keep explicitly, in the resulting model, the expert’s background knowledge that permitted the definition of the space–time windows. So the method we use is derived from Quinlan’s approach [16] and the ID3 algorithm adapted to fuzzy membership values [17]–[21]. First, a learning set is built of examples/observations, characterized both by their fuzzy member-

ship values in relation to the modalities of a set of categorical variables (or attributes) and by their fuzzy or crisp membership values in relation to the classes of a known categorization (see Table I for an example of learning set and for notations within our application). This learning set is subjected to a four-step procedure.

- 1) A *discrimination* measure is used to determine which categorical variable best explains the repartition of the patients among the classes, and a *node* is created.
- 2) The data set is partitioned to build as many subsets as there are modalities of the variable chosen in step 1.
- 3) A *termination condition* is tested with the help of a termination criterion.
- 4) If the termination condition is verified, then the subset is considered to be a *leaf* of the tree, which means that it has an acceptable proportion of examples belonging to one class, as defined by the termination criterion. If the termination condition is not verified, then steps 1) to 3) are repeated.

The most popular discrimination measure is based on information entropy as defined by Shannon. The fuzzy conditional information entropy, denoted E^(l), is defined by the following expression: for the lth attribute

$$E^{(l)} = \sum_{i=1}^{N_s} \frac{M(\chi^{N/S_i^{(l)}})}{\sum_{j=1}^{N_s} M(\chi^{N/S_j^{(l)}})} \cdot \sum_{c=1}^C -p^{N/S_i^{(l)}}(c) \cdot \log(p^{N/S_i^{(l)}}(c)) \quad (2)$$

M(S_i^(l)) is the cardinality of the fuzzy set S_i^(l) defined as

$$M(S_i^{(l)}) = \sum_{k=1}^K \mu_{S_i^{(l)}}(k) \quad (3)$$

where K is the number of examples and μ_{S_i^(l)(k) the membership value of the fuzzy subset S_i^(l) for the data relative to the lth variable and the kth example (see Table I).}

$$\forall k \in [1, K], \chi^{N/S_i^{(l)}}(k) = \min(\chi^N(k), \mu_{S_i^{(l)}}(k)) \quad (4)$$

where χ^N is a vector that indicates to what extent the examples verify the conjunction of the fuzzy restrictions from the root of the tree to the node N [18]. At the root level, $\forall k \in [1, K], \chi^0(k) = 1$. The min operator completes the logical conjunction AND.

In (2), $p^{N/S_i^{(l)}}(c)$ can be considered as the fuzzy conditional probability of belonging to the class c , given the fuzzy restriction $S_i^{(l)}$ at the node N [22], [23]. $p^{N/S_i^{(l)}}(c)$ is defined as

$$p^{N/S_i^{(l)}}(c) = \frac{M(\chi^{N/S_i^{(l)}}) \cdot C_c}{M(\chi^{N/S_i^{(l)}})} \quad (5)$$

where C_c is the set of membership values of the patients' learning set for the class c (see Table I). In fact, because the occurrences of the examples are assumed to be equally probable and the product operator is used in the numerator expression, $p^{N/S_i^{(l)}}(c)$ meets the conditions: $\sum_c p^{N/S_i^{(l)}}(c) = 1$.

The variable chosen to split the observations of the training set at a node is the one that minimizes $E^{(l)}$, the conditional information entropy.

After selecting an attribute, examples must be shared out among the subnodes of the tree. The membership value of an example in the subnode that corresponds to the fuzzy subset $S_i^{(l)}$ is defined as follows:

$$\begin{cases} \chi^{N/S_i^{(l)}}(k) = \min(\chi^N(k), \mu_{S_i^{(l)}}(k)), \\ \quad \text{if } \mu_{S_i^{(l)}}(k) = \max_{j \in [1, N_A]}(\mu_{S_j^{(l)}}(k)) \\ \chi^{N/S_i^{(l)}}(k) = 0, \text{ otherwise.} \end{cases} \quad (6)$$

This definition implies that an example belongs to only one sub node. While this is not an essential property for the fuzzy decision trees [18], [19], the computation time and the size of the tree are significantly reduced given this restriction.

The termination criterion is based on a threshold denoted $\beta \in [0, 1]$ and is defined as follows: if, at the node N , $p^N(c) \geq \beta$, then the node N is a leaf. $p^N(c)$ is denoted p at the leaf level.

The conditional information entropy is particularly adapted to cases where there are two classes [19]. Consequently, we induce as many decision trees as there are classes, and each tree infers the membership and the non-membership of a class. So for each clinical classification with C ($C > 2$) classes, a "forest" of decision trees (i.e., a set of C trees) is induced, as proposed by Marsala [19]. This principle is graphically presented in Fig. 1.

C. Defining Fuzzy Rules

Each path from the root to the leaf of a fuzzy decision tree (i.e., each branch of a tree) can be converted into a rule. The fuzzy conditions of these rules are comprised of the fuzzy modalities of the attributes along the path. The rule concludes, at the leaf level, with the membership or the nonmembership in the class for which the tree was induced and p corresponds to the rule's fire strength.

Note that in the present study, the main goal of the rule base is not to make predictions, but rather to act as a knowledge depository that would allow new situations to be explained.

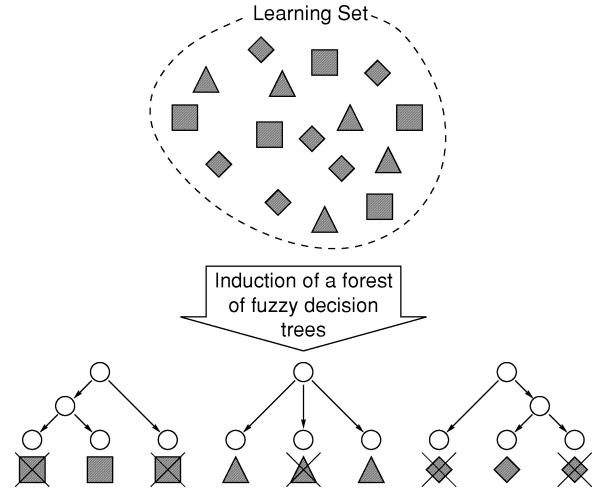


Fig. 1. Induction of a forest of decision trees for a classification with classes \square , \triangle , and \diamond . One tree is induced per class. A branch of a tree infers the membership or the nonmembership (crossed symbols) of the class.

For this purpose, the knowledge extracted from the data, and modeled by the fuzzy rules, has to be as general as possible so that it can be exploited by the final user, which means that the rules must be concise and intelligible. However, the knowledge also has to be specific enough to permit rare clinical situations to be explained. Clearly, if such a situation is observed with a new patient, a set of rules that infer this situation must exist in the rule base. To manage these contradictory criteria, β is initially set at an initial value β_0 and is increased by increments until both the membership *and* the nonmembership of the class under consideration are explained.

D. Rule Base Optimization

The discrimination measure and the tree induction process are heuristically defined to produce the smallest tree. Still, each individual rule can be optimized by removing all fuzzy conditions that penalize the rule's fire strength. The method is similar to that of Yuan and Shaw [21]: as each condition is removed from the *IF* part of the rule, a new fire strength p' is computed. If $p' \geq p$, the condition is removed once and for all. The result is shorter and more general rules with higher fire strengths.

E. Objective Explanation of Patient's Clinical State

This section proposes a method to provide a set of hypothesis that could explain a clinical characteristic of a given patient by exploiting the fuzzy rule bases induced beforehand. Proposing an explanation of a situation is possible if this situation was part of the learning process. Given the definition of the parameter β , a set of fuzzy rules leading to this clinical situation should exist. The explanation is found by editing the rules that lead to this clinical situation and whose fuzzy conditions (*IF* part) are best satisfied by the patient's biomechanical data.

Let $\{R_1, R_2, \dots, R_N\}$ be the set of fuzzy restrictions (conditions) of the rule R . $\forall j \in [1, N], \mu^{R_j}(k)$ is the membership value of the condition R_j for the patient k and represents the extent to which the biomechanical data of the k th patient satisfy

R_j . Let $S_R(k)$ be a satisfiability measure defined as

$$S_R(k) = \min_{j \in [1, N]} (\mu^{R_j}(k)). \quad (7)$$

The rule(s) that can explain the given clinical characteristic of the patient is (are) the one (those) that maximizes $S_R(k)$. In (7), the min operator is used instead of the product operator to make $S_R(k)$ independent of the length of the rule premises. In fact, in practice, all trees are not the same size, and for any given tree, branch length can vary, due to the different sizes of the classes in the learning set.

III. APPLICATION

The previously described data analysis method was applied in order to objectively explain the functional state of patients with a shoulder prosthesis.

A. Population of the Study

Forty two patients were included in the study. They were about to undergo shoulder arthroplasty or had already been operated on. Their average age was 65 (standard deviation was 9), and they suffered from various pathologies, the most frequent being osteoarthritis and arthritis. At present, all these patients have been operated on and have participated in a rehabilitation program in the department of orthopedic surgery run by Dr. De. Bouttens of the *Groupe Hopale*.

B. Biomechanical Data

In this application, we considered only the kinematic measurements of the upper limb. Several procedures for the measurement of the upper limb kinematics have been proposed in order to provide a routinely usable protocol in the clinical domain (see [9], [13], [14], [24], and [25]). However, we had to develop a protocol more suitable to the specificities of the patient's state of health and to the results we wanted to obtain. So the measurement protocol and the data preprocessing method that we used are original [12]. They are routinely employed at the Hopale Group's Movement Analysis Laboratory.

Measurements are carried out as follows. A 6-camera optoelectronic system is used to measure the 3-D positions of reflective markers fixed on the patients' skin [12]. Movement is measured during three different gestures that are specific to our study: the elevation of the arm in the scapular plane (SP), the movement of the hand to the nape of the neck (HN), and the elevation of a light load from table height (LE). These gestures, considered to be representative of everyday activities, are repeated from three to five times, depending of the patient's state of health. Patients are seated and are asked to perform the tasks as naturally as possible. Instructions are given concerning the beginning and the end of the gestures: the subject is asked to look straight ahead with both arms outstretched vertically and with the palms turned medially. The body segments taken into account are the head, the trunk, the arm, the forearm, and the hand. Euler's angles are used to describe the relative movement of the body segments since this representation is easily interpretable by clinicians. For each patient, a preliminary pre-

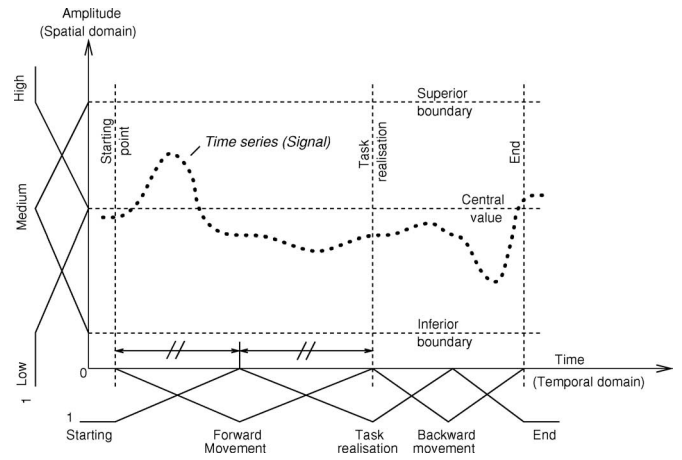


Fig. 2. Fuzzy space–time windowing for the movements: Elevation of the arm in the Scapular Plane (SP) and movement of the Hand to the Nape of the neck (HN) (see Section III-C of the text for details).

operative measurement is performed, followed by a series of postoperative measurements, performed three and six months after surgery, and then once a year.

In our present application, the actual number of the measurements varied from one patient to another, depending on the length of the patient's follow-up period. In the present paper, data were not considered as longitudinal. In fact, for a given patient, each evaluation defined a new observation that was added to the learning set (see Section III-E). Nineteen ($N_V = 19$) analysis variables were retained to describe the kinematics, including the relative orientations of the studied body segments, the angular velocities and accelerations of shoulder and elbow flexion, and the linear velocity and acceleration of the hand. These 19 variables were the result of a compromise between information saving and interpretability of movement description. This set of variables was chosen in cooperation with the medical team, and a preliminary multivariate data analysis permitted certain highly correlated variables to be removed. We do not further develop the whole process of the choice of the variables as it is not properly the issue of the present paper.

C. Fuzzy Space–Time Windowing

For each measurement trial, significant periods of movement were defined. For SP and HN movements, five fuzzy periods, i.e., five fuzzy time windows, were defined according to the following five instants: the starting point, the instant of task completion (maximal elevation of the arm or hand in contact with the nape of the neck), the return to the initial position (movement's end) and the two midpoints between each couple of consecutive instants (see Fig. 2). LE movement was segmented into nine fuzzy periods by defining the following nine instants: the starting point, the grasping of the load prior to lifting, the maximal elevation of the load, the placing of the load after lifting, the movement's end and the four midpoints between each couple of these consecutive instants. Note that these instants were defined from a functional point of view, i.e., in relation to the completion of a task. In our application, analysis variables were coded

using N_T ($N_T \in \{5, 9\}$) fuzzy triangular time windows centered on characteristic instants (see Fig. 2). Note that our time windows are trial dependant as they depend on characteristic instants. This approach is quite different from the one proposed by Loslever *et al.* [10], in which time windows are of equal width. It provides a higher temporal abstraction level and a functional representation of the movement that is of major interest within the evaluation of the functional recovery. As an illustration, with such a movement characterization, we do not refer to the “shoulder flexion between X% and Y% of the movement duration,” but to the “shoulder flexion during the lifting of the load,” for example.

All analysis variables were coded using $N_A = 3$ fuzzy triangular space windows. Loslever *et al.* [3] showed this characterization accomodates data summarizing and information saving. These windows were defined according to three characteristic values for each variable (see Fig. 2). These values were established according to: i) statistical properties of the variables (e.g., -25 and 40 were the extrema of the values for the rotation of the head and thus corresponded to the inferior and superior boundaries, respectively); ii) anatomical joints’ limits (e.g., 0 and 150 are the anatomical minimum and maximum elbow flexion and consequently defined the inferior and superior boundaries for this variable, respectively); iii) preferences of the clinicians (e.g., more than 150 for the shoulder flexion was considered as healthy, so 150 defined the superior boundary for this variable). For a given variable, the space windows were the same for all the gestures (SP, HN, and LE) in order to assign an unique meaning to a space modality and thus to facilitate the interpretation of the results.

For both time and space domains, the triangular shape of the windows was chosen because it only requires three parameters, facilitates and speeds up computation and has been proved to correctly characterize the data [3].

D. Functional Evaluation

The functional state of the patients was determined by the American Shoulder and Elbow Surgeons (ASES) questionnaire [26]. Of course, any functional evaluation method could be used. The choice of the ASES questionnaire is justified here by its routine application in the orthopedic department of the *Groupe Hopale*. The ASES questionnaire is a typical subjective self-evaluation composed of ten questions concerning the ability to perform everyday tasks: “to put a coat on,” “to do one’s hair,” “to sleep on the shoulder,” “to wash one’s back,” “to get washed,” “to reach a high shelf,” “to lift 500 g above the shoulder,” “to throw a ball,” “to practice professional activities,” and “to practice a sport.” Patients answer each question by choosing the appropriate modality among the following four: “Impossible,” “Very difficult,” “Quite easy,” and “Easy.” The last two items of the questionnaire, related to professional and sports activities, were ignored in our application. In fact, due to their age and functional deficiencies, most of the patients did not have professional and sports activities and consequently did not respond to these two questions. In any case, the situations with regard to these two items are not the same from an

individual to another. Therefore, inducing general knowledge from these highly specific situations would not make sense. So eight items out of ten were studied. For each patient, the functional evaluation was performed the same day as the kinematic measurements.

E. Induction of the Fuzzy Rule Base

To facilitate the interpretation of results, the three gestures [*Hand to the Nape* (HN), *arm elevation in the Scapular Plane* (SP) and *Load Elevation* (LE)], were considered separately. The data training pairs (or examples) from which the rules were induced were constructed using the membership values of the kinematic data for a given gesture and the responses to a given question of the ASES functional evaluation (see Table I). A patient could define more than one example, since he/she could be evaluated both cinematically and functionally several times (pre- and postoperatively). By doing this, we assume that the relations between the kinematic behavior and the functional abilities are independent of the phase of the rehabilitation program in which the patients are. The training sets gathered 87 examples for the gesture HN, 90 examples for the LE and 93 examples for the SP. The set of modalities associated with an item of the ASES questionnaire was considered as a classification of the set of patients. Given the principle of a forest of decision trees, four trees (the same number as the modalities) were induced for each ASES item.

IV. RESULTS

A. General Results

Ninety-six fuzzy decision trees were induced, corresponding to three gestures and eight items of the ASES functional evaluation with four response modalities. β_0 was fixed at 0.7. We obtained 1330 rules. The average number of rule conditions was two, with a minimum of one and a maximum of 11; 95.8% of the rules had at most five conditions.

As an example, the following rules link the gesture HN to the impossibility and the possibility of *doing one’s hair*:

- 1) **IF** *Hand velocity is Medium during Forward movement (during HN gesture)*, **THEN** *it is Possible (to do one’s hair)* ($p = 0.96$);
- 2) **IF** *Hand velocity is High during forward movement*, **THEN** *it is Possible* ($p = 1$);
- 3) **IF** *Shoulder Flexion is Low at the instant of contact*, **THEN** *it is Impossible* ($p = 1$);
- 4) **IF** *Hand velocity is Low during forward movement AND Wrist is in Extension during forward movement*, **THEN** *it is Impossible* ($p = 0.77$);
- 5) **IF** *Hand velocity is Low during forward movement AND Shoulder Flexion is Low during forward movement*, **THEN** *it is Impossible* ($p = 0.86$);
- 6) **IF** *Wrist Flexion is Low during forward movement AND Shoulder Flexion is Medium during forward movement*, **THEN** *it is Possible* ($p = 0.91$).

The fuzzy rule bases act as a knowledge repository from which different kind of information can be retrieved. In fact,

the fuzzy rules map functional deficiencies and insufficient mobility ranges, movement velocities and/or accelerations that are directly related to the pathological joints and muscles. It is the case, for example, in the two following rules: 1) *During LE gesture realization, IF the Shoulder Flexion is “Low” at the instant of maximum load elevation, THEN it is “NOT-Easy” to reach a high shelf* ($p = 1$); and 2) *IF the Shoulder Flexion Velocity is “Low” during the load elevation, THEN it is “NOT-Easy” to reach a high shelf* ($p = 0.90$). Rules can also map functional deficiencies and compensatory strategies characterized by kinematic behaviour of (supposed) healthy joints. In fact, the wrist is a healthy joint in the following rule: *During HN gesture realization, IF the Hand Velocity is “Low” AND IF the Wrist is in “Extension” during forward movement, THEN it is Impossible to do one’s hair* ($p = 0.77$). In fact, the hyper-extension of the wrist is a frequently observed compensatory strategy when patients have difficulties to reach the nape of their neck due to restrictions in shoulder mobility. Eventually, fuzzy rules mapping gesture characteristics and the facility to perform a given task often describe the different “healthy” strategies used to complete the considered gesture. It is the case for the following rule: *During SP gesture, IF the Shoulder Flexion is “Important” at the maximum arm elevation in the scapular plane, THEN it is “Easy” to reach a high shelf* ($p = 0.92$).

B. Explanation Rate

In order to evaluate the explanation method for use with patients that did not belong to the training set, a stratified tenfold cross validation was performed on each of the 96 fuzzy rule bases [27]. For each patient in the test sets, we determined which rule was most likely to explain his/her functional self-evaluation, i.e., which rule produced the maximum satisfiability value. If all the conditions of the explanatory fuzzy rule were met by the kinematic characteristics of the patient, the functional state of a patient was considered explained. The condition $[V_n \text{ is } S_i^{(l)}]$ was satisfied by the kinematic data relative to the patient k if: $\max_{j \in [1, N_A]} \mu^{S_j^{(l)}}(k) = \mu^{S_i^{(l)}}$. An *explanation rate* index, intended to evaluate the explanation capacities, was defined as the proportion of patients in a test set whose functional state had been explained.

For each patient in a test set, we determined whether or not his/her functional state had been explained, and we computed the explanation rate relative to the set. Then, the ten explanation rates, relative to the ten test sets, were averaged. Ninety-six *Mean Explanation Rates* (MER) were thus obtained. The MER evaluates the capacity of the method to explain a given response of the ASES questionnaire (“Easy” for instance) *or* its negation (“NOT-easy”). Because of the large number of decision trees involved, we have chosen to describe our results globally, by presenting the proportions of the 96 rule bases (i.e., of the 96 decision trees) that reached some given performances. Thus, our computations demonstrate that 54% of the 96 fuzzy rule bases present a MER at least equal to 80% and that only 6% of the rule bases provide a MER below 70%. We do not detail the results of the 90 rule bases for which the MER exceed 70%. However, in order to critically discuss the method, Table II presents the

TABLE II
THE SIX LEAST EFFECTIVE ASSOCIATIONS GESTURE-FUNCTIONAL STATE FOR THE EXPLANATION

Gesture	Functional state	MER
		in % (standard deviation in %)
SP	To put a coat on / Easy or NOT-Easy	68.9 (8.9)
HN	To throw a ball / Quite Easy or NOT-Quite Easy	68.5 (6.5)
SP	To sleep on the shoulder / Easy or NOT-Easy	64.8 (11.5)
LE	To sleep on the shoulder / Easy or NOT-Easy	60.0 (15.9)
LE	To wash one’s back / Impossible or NOT-Impossible	50.4 (20.8)
SP	To wash one’s back / Impossible or NOT-Impossible	47.6 (21.9)

six *kinematic data-functional state* associations that resulted in a rule base with a MER below 70%.

We also defined a more restrictive evaluation index, denoted MER^+ , by assessing the method capacity to explain a given response without considering its negation. The MER and the MER^+ could be intuitively compared to the *precision* and the *sensitivity* in a classification problem, respectively.

As it is developed in the *Discussion* section, the MER^+ was particularly restrictive in our application as some modalities of the ASES questionnaire have been rarely answered by the patients. For instance, the case “*It is Very Difficult to reach a high shelf*” appeared only four times. So the learning process was performed with a class composed of only three examples. Moreover, the MER^+ could thus only be computed for the four test sets that comprised one example for which “*It is Very Difficult to reach a high shelf*.” In this case, the explanation rate was then equal to one or zero for each test set. However, we chose to present the results that correspond to the limits of the applicability of the method. In fact, these results do not discard the method by itself and permits to discuss the functional evaluation. Fig. 3 presents the cumulative histogram of the MER^+ values. It shows that 16.7% of the MER^+ values exceed 80% for the modalities *Impossible* and *Easy*, but no decision tree is able to provide a MER^+ superior to 50% for the two other modalities. These results are discussed and put into perspective in the next section.

V. DISCUSSION

In this section, some aspects of the general method are discussed in terms of the application and of the results presented above, as methodological choices have been justified throughout the paper.

The general results show that the model provided by fuzzy rules is readable. In fact, only 4.2% of the fuzzy rules are potentially difficult to exploit by the final user, in that they have more than five conditions. Still, some results remain quite difficult for the surgeon to interpret. Certain functional states are more related to the kinematic characteristics of joints exhibiting no previously established pathology (wrist, neck, etc.) than they are to problems in the shoulder joint. These kinematic characteristics reflect compensatory strategies, and should not be considered as the explanation for the functional deficiency but rather as the consequence of the deficiency. This remaining difficulty does, however, not detract from the value of our

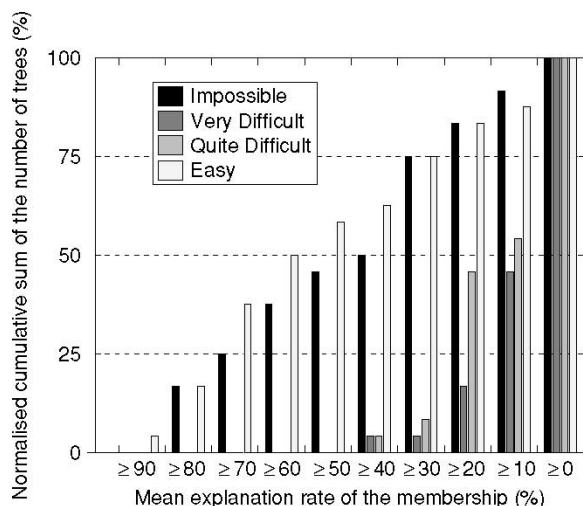


Fig. 3. Histogram of the normalized cumulative sum of the number of trees associated with the MER^+ values (MER values for the “positive answers”) as a function of the modalities. For instance, 38% of the 24 trees (rule bases) that infer the modality “Easy” have a MER^+ superior or equal to 70. Of course, 100% of the trees have an explanation rate superior or equal to zero.

proposed method, since the mapping of the objective biomechanical data and the clinical data (possibly subjective) can still direct clinicians’ attention to important biomechanical information related to particular clinical features, i.e., relative to a given context of interpretation. However, it does bring the term “explanation” into question. (Despite this, we will continue to use the term, at least for the time being.) Note that identifying compensatory strategies can be of major interest as the definition of an individualized rehabilitation program can be derived from the identified compensations. In fact, some rehabilitation programs proposed to immobilize healthy joints used for compensation, in order to force the pathological joints to be mobilized.

The stratified tenfold cross validation method permits a quantitative evaluation of the explanation method. Good results were obtained for the MER . Although Table II indicates that our method is less effective at explaining difficulties in performing certain tasks, such as *To put a coat on* and *To wash one’s back*. The fact that these tasks require an internal rotation of the shoulder that cannot be specifically measured through the three gestures analyzed here, could account for this lack of effectiveness. Patients present great difficulties to perform shoulder rotation and specific gestures intended to obtain rotation amplitude would be often unrealizable. So adding such gestures to the existing measurement protocol seems unconceivable. In addition, the low MER value for the task, *To sleep on the shoulder*, could be explained by the absence of a direct link between this task and a specific movement. *To throw a ball*, on the other hand, requires movements that can easily be evaluated through the three gestures discussed in this paper. Despite these seeming discrepancies, a detailed study of the other results shows that similarities, or differences, between the measured gestures and those required to complete the ASES questionnaire tasks do not seem to in-

fluence the performances of our explanation method significantly.

The factor that most influences explanation performances is the size of the classes in the training set. The poor results (Fig. 3) concerning the explanation of the two modalities, *Very difficult* and *Quite easy*, can be explained by the low proportions of these two responses. These two modalities, *Very difficult* and *Quite easy*, are also particularly prone to subjective interpretation and uncertainty. Obviously, identifying a task as “Impossible” or “Easy” is simpler than judging it “Very difficult” or “Quite Easy.”

Still, given that the evaluation method only considers fuzzy rules whose conditions are totally satisfied by the kinematic data, these poor results do not invalidate the use of this explanation method for these two functional deficiencies. Clearly, some rules can provide partial explanations if only a subset of their conditions is satisfied. Moreover, fuzzy rules can also be considered absolutely, providing an explanation of a functional deficiency that is not relative to a given patient, i.e., providing a general knowledge. Other discrimination measures, such as the ambiguity measure [21] or the measure of fuzziness [19], could improve performances by better managing class size differences, but the results could be less intelligible for clinicians, who are more familiar with probabilities.

In particular, our results call into question the efficiency of the ASES evaluation. Given that the two modalities, *Very difficult* or *Quite Easy*, are rarely selected, using three response modalities (for example: *Easy*, *Difficult*, *Impossible*) might be more appropriate to the way the patients evaluate their own functional state. Other evaluation methods do exist. For instance, the Shoulder Disability Questionnaire (SDQ) [28] is used to evaluate functional status limitations in patients with shoulder disorders. It covers 16 items related to patient activity over the 24 h preceding the evaluation, and offers three response modalities: *Yes*, if the patient is restricted with respect to the given activity; *No*, if he/she is not restricted; and *NA* (Not Applicable), if the given activity has not been performed during the past 24 h. These response options tend to focus patients on the activities they actually can perform and not on those they would be able to do, or even on those they would like to be able to do.

SDQ answers are less uncertain and subjective than those solicited by the ASES evaluation. However, they gather information on the person’s ability to perform certain tasks, without evaluating the difficulties encountered during their performance.

Another evaluation method could propose a continuous scale for each item of the functional evaluation, with the patient choosing a position between two extreme functional states like *absolutely easy* and *completely impossible*. The functional status could be characterized using an *a posteriori* definition of fuzzy modalities, which would allow the subjectivity and the uncertainty of the answers to be modeled. Such a functional evaluation should be exploitable using the data analysis method proposed in this article, without any modifications. However, one of the drawbacks of such a functional evaluation would be the computerization necessary to facilitate answer collection and processing.

VI. CONCLUSION

The method we propose in this article for the objective “explanation” of clinical characteristics is a step towards taking objective biomechanical measurements into account in medical practice, especially in orthopedics. We have shown that by connecting kinematic data with an easily understandable functional evaluation, our method can support the contextual interpretation of movement characteristics, thus helping the practitioner interpret the movement from a functional point of view. In the future, functional deficiencies could be reduced or eliminated by acting on the movement characteristics likely to explain them. Rules obtained using fuzzy decision trees can also be seen as classification rules, which would provide an objective functional evaluation based solely on kinematic data.

This work is original in a sense that it proposes first a new method within biomechanical data analysis and secondly deals with the upper extremity on which research efforts begin to focus. The method presented in this paper is part of the effort to improve the quality of medical practice and of its objective evaluation.

ACKNOWLEDGMENT

The authors would like to thank O. Remy-Neris, M.D., Ph.D., the director of the Laboratoire d’Analyse du Mouvement du Groupe Hopale, as well as the laboratory staff, for their help in performing this study. The authors would also express their sincere thanks to A. Hernández for his friendly help and to L. Spencer for her professionalism in correcting this paper.

REFERENCES

- [1] S. R. Simon, “Quantification of human motion: Gait analysis—benefits and limitations to its application to clinical problems,” *J. Biomech.*, vol. 37, no. 12, pp. 1869–1880, 2004. [Online]. Available <http://www.sciencedirect.com/science/article/B6T82-4CN9KK8-l/2/9162919271964102ee9a1346ba9156ee>
- [2] M. G. Benedetti, F. Catani, A. Leardini, E. Pignotti, and S. Giannini, “Data management in gait analysis for clinical applications,” *Clin. Biomech.*, vol. 13, no. 3, pp. 204–215, 1998.
- [3] P. Loslever, J. Popieul, and P. Simon, “From classic statistical characterization to fuzzy windowing based characterization for the exploratory analysis of miscellaneous time variables: Example in the field of car driving studies,” *Fuzzy Sets Syst.*, vol. 137, pp. 271–296, 2003.
- [4] T. Chau, “A review of analytical techniques for gait data, part 1: Fuzzy, statistical, and fractal methods,” *Gait Posture*, vol. 13, no. 1, pp. 49–66, 2001.
- [5] —, “A review of analytical techniques for gait data, part 2: Neural network and wavelet methods,” *Gait Posture*, vol. 13, no. 2, pp. 102–120, 2001.
- [6] A. R. Morris, S. Naumann, G. M. T. D’Eleuterio, and J. Wedge, “Simulating cerebral palsy locomotion for surgery,” in *Proc. VIIIth Int. Symp. Computer Simulation in Biomechanics*, Calgary, Canada, 1999.
- [7] T. Bylander, M. Weintraub, and S. R. Simon, “Quawds: Diagnosis using different models for different subtasks,” in *Second Generation Expert Systems*, R. Simmons, J. David, and J. Krivine, Eds. Berlin, Germany: Springer-Verlag, 1993.
- [8] O. Kuchar, Y. Lazukova, D. Riordan, and J. Leahey, “A case-based reasoning system involving a quantized method for gait disorder diagnosis,” in *Proc. IASTED Int. Conf. Artificial Intelligence and Soft Computing*, Cancun, Mexico, 1998, pp. 367–370.
- [9] G. Rau, C. Disselhorst-Klug, and R. Schmidt, “Movement biomechanics goes upwards: From the leg to the arm,” *J. Biomech.*, vol. 33, no. 10, pp. 1207–11216, 2000.

- [10] P. Loslever and S. Bouillard, “Marriage of fuzzy sets and multiple correspondence analysis: Examples with subjective interval data and biomedical signals,” *Fuzzy Sets Syst.*, vol. 107, pp. 255–275, 1999.
- [11] A. Cappozzo, F. Catani, A. Leardini, M. G. Benedetti, and U. D. Croce, “Position and orientation in space of bones during movement: Experimental artefacts,” *Clin. Biomech.*, vol. 11, no. 2, pp. 90–100, 1996.
- [12] E. Roux, S. Bouillard, D. Bouttens, and A.-P. Godillon-Maquinghen, “Evaluation of the global optimisation method within the upper limb kinematics analysis,” *J. Biomech.*, vol. 35, no. 9, pp. 1279–1283, 2002.
- [13] R. Schmidt, C. Disselhorst-Klug, J. Silny, and G. Rau, “A measurement procedure for the quantitative analysis of the free upper-extremity movements,” in *Proc. 5th Int. Symp. 3-D Analysis of Human Movement*, Chattanooga, TN, 1998, pp. 47–50.
- [14] —, “A marker-based measurement procedure for unconstrained wrist and elbow motions,” *J. Biomech.*, vol. 32, pp. 615–621, 1999.
- [15] X. Boyen and L. Wehenkel, “Automatic induction of fuzzy decision trees and its application to power system security assessment,” *Fuzzy Sets Syst.*, vol. 102, pp. 3–19, 1999.
- [16] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann, 1992.
- [17] H. Ichihashi, T. Shirai, K. Nagasaka, and T. Miyoshi, “Neuro-fuzzy id3: A method of inducing fuzzy decision trees with linear programming for maximizing entropy and an algebraic method for incremental learning,” *Fuzzy Sets Syst.*, vol. 81, no. 1, pp. 157–167, 1996.
- [18] C. Z. Janikow, “Fuzzy decision trees: Issues and methods,” *IEEE Trans. Syst., Man, Cybern.*, vol. 28, no. 1, pp. 1–14, Feb. 1998.
- [19] C. Marsala, “Apprentissage inductif en présence de données imprécises: Construction et utilisation d’arbres de décision flous,” Ph.D. dissertation, University of Paris VI, 1998.
- [20] X. Wang, B. Chen, G. Qian, and F. Ye, “On the optimization of fuzzy decision trees,” *Fuzzy Sets Syst.*, vol. 112, no. 1, pp. 117–125, 2000.
- [21] Y. Yuan and M. J. Shaw, “Induction of fuzzy decision trees,” *Fuzzy Sets Syst.*, vol. 69, no. 2, pp. 125–139, 1995.
- [22] D. Dubois and H. Prade, “Fuzzy sets and probability: Misunderstandings, bridges and gaps,” in *Proc. 2nd IEEE Conf. Fuzzy Systems*, San Francisco, CA, 1993.
- [23] L. Zadeh, “Probability measures of fuzzy events,” *J. Math. Anal. Applicat.*, vol. 10, pp. 421–427, 1968.
- [24] I. Murray and G. Johnson, “Definition of marker positions and technical frames for studying the kinematics of the shoulder,” in *Proc. 3rd Conf. Int. Shoulder Group*, H. Chadwick, F. Veeger, E. Van Der Helm, and J. Nagels, Eds. Newcastle Upon Tyne, UK: DUP Science/2001, 2000.
- [25] G. Rab, K. Petuskey, and A. Bagley, “A method for determination of upper extremity kinematics,” *Gait Posture*, vol. 15, no. 2, pp. 113–119, 2002.
- [26] R. Richards *et al.*, “A standardized method for the assessment of shoulder function,” *J. Shoulder Elbow Surg.*, vol. 3, pp. 347–352, 1994.
- [27] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Proc. Int. Joint Conf. Artificial Intelligence (IJCAI)*, Montreal, QC, Canada, 1995.
- [28] G. J. Van der Heijden, P. Leffers, and L. M. Bouter, “Shoulder disability questionnaire design and responsiveness of a functional status measure,” *J. Clin. Epidemiol.*, vol. 53, no. 1, pp. 29–38, 2000.



Emmanuel Roux received the Ph.D. degree in industrial and human automation and computer science from the University of Valenciennes, France, in 2002.

He worked as a Postdoctoral Fellow in the Epidemiology and Biostatistics unit of the INSERM (U472-IFR 69), Villejuif, France. He is currently a Postdoctoral Fellow in the Laboratory of Signal and Image Processing of the INSERM (U642), Rennes, France.

Anne-Pascale Godillon-Maquinghen, photograph and biography not available at the time of publication.



Patrice Caulier received the Ph.D. degree in industrial and human automation and computer science from the University of Valenciennes, France, in 1997.

He is currently an Assistant Professor at the Laboratory of Industrial and Human Automation, Mechanic and Computer Science (UMR CNRS 8530) of the University of Valenciennes. His research work relates to the artificial intelligence techniques applied to the monitoring, diagnosis, and prognosis in the medical and industrial fields.

Stéphane Boulland received the Ph.D. degree in industrial and human automation and computer science from the University of Valenciennes, Valenciennes, France, in 1998.

He is currently an Engineer in a clinical center, where he is the specialist in measurement and diagnosis based on video-computer systems. His research in areas involving imaging systems includes gait analysis, and back and shoulder movement analysis.

Denis Bouttens, photograph and biography not available at the time of publication.