10^{th} IMEKO TC7 International Symposium
June 30-July 2, 2004, Saint-Petersburg, Russia

SIGNAL PROCESSING ON FUZZY NOMINAL SCALES

Thomas ALLEVARD, Eric BENOIT, Laurent FOULLOY

LISTIC-ESIA, Université de Savoie, Annecy, France

Abstract - This paper presents some characteristics of fuzzy nominal scales, more specifically the existence of a distance operator on such scales. How this operator can be used for the processing of linguistic signals is presented. Numerical data provided by a data glove are transformed into linguistic data with fuzzy nominal scales. These linguistic data are then processed to recognize dynamic hand gestures.

Keywords: Fuzzy Nominal Scales, gesture recognition, measurement theory.

1. INTRODUCTION

The main objective of a measurement process is to convert the physical state of an entity into an information entity [1]. In order to manipulate information entities, relations that exist between physical states are also converted into relations between information entities. The goal of measurement scales is to define the set of possible relations that can be used on the information entities. In the case of nominal scales for example, the only relation that can be used is an equivalence relation.

Signal processing is commonly defined as the processing of information issuing from a measurement process. If numerical values are commonly used to represent measurement results, it is now admitted that some applications manipulate symbolic values or linguistic terms better. But the measurement process is then usually based on a nominal scale [2]. Hence only processings that use exclusively an equivalence relation can be used like the multicriteria binary classification [3][4] or rule-based classifications.

Introduced in [5] and [6], fuzzy nominal scales are based on the fuzzy descriptions of physical states. The set of possible relations on such scales contains a fuzzy equivalence relation also called proximity relation. This relation allows us to define a distance measure between information entities which cannot be done with a nominal scale. This paper shows the possibility of performing signal processing of information entities obained with a fuzzy nominal scale.

Linguistic data that is processed in this paper is provided by the fuzzy glove [7]. This fuzzy sensor gives a linguistic description of the hand configuration. Dynamic gestures are classified based on a distance operator that is defined between hand postures.

In the next section, fuzzy nominal scales and the distance operator defined on them are presented. Section 3 introduces the fuzzy glove and section 4 the dynamic gesture recognition algorithm. Experimental results are presented in sections 5 before the conclusion in section 6.

2. FUZZY NOMINAL SCALES

Let X be a set of object states. In order to characterize linguistically any measurement over X, let L be a set of linguistic terms, representative of the physical phenomenon. Let $F(L)$ be the set of fuzzy subsets of L.

The conversion of a physical state into its linguistic representation is called a fuzzy linguistic description mapping [9] or simply a fuzzy description mapping. It transforms an object x of X into a fuzzy subset of linguistic terms $D(x) \in F(L)$ called the *fuzzy description* of x. A dual mapping, called the *fuzzy meaning* mapping, associates a fuzzy subset of X to each term l of the lexical Set L. This fuzzy subset is the *fuzzy meaning* of l, $M(l)$. In this paper, fuzzy meanings are chosen to respect some characteristics of fuzzy sensors [9] and they form a strict fuzzy partition of X.

Considering the subset $LF(L)$ of $F(L)$ defined by:

$$A \in LF(L) \Leftrightarrow \sum_{l \in L} \mu_A(l) = 1 \qquad (1)$$

$LF(L)$ contains any fuzzy description $D(x)$ of a physical state x.

Whereas the only relation that can be defined on symbols is the equality relation, more interesting relations can be defined on fuzzy subsets of linguistic terms. One of them is the fuzzy equality which is defined from the fuzzy inclusion [10]:

$$[A \subseteq_t B] = \forall l(\mu_A(l) \rightarrow_t \mu_B(l)) \qquad (2)$$

Where t is a t-norm and $\rightarrow_t$ is the residuated implication asssociated to t. The fuzzy equality relation is then equal to the conjunction of the two reciprocal fuzzy inclusions. When t is the Lukasiewicz t-norm, its degree is equal to:

$$[A \equiv_{t} B] = \inf_{l \in L} 1 - |\mu_A(l) - \mu_B(l)| \qquad (3)$$

It has been shown in [11] that this relation has more interesting properties when it is extended to the closure of *L* by the *OR* operator. When *A* and *B* are elements of *FL(L)*, the degree of ≡ is then equal to:

$$\mu_{\equiv}(A, B) = \sum_{l \in L} min(\mu_A(l), \mu_B(l)) \qquad (4)$$

This similarity relation induces a similarity relation between elements of *X*:

$$\mu_{\sim}(x, y) = \mu_{\equiv}(D(x), D(y)) \qquad (5)$$

The duality between the fuzzy description mapping *D* and the fuzzy meaning mapping *M* allows us to write this relation as:

$$\mu_{\sim}(x, y) = \sum_{l \in L} min(\mu_{M(l)}(x), \mu_{M(l)}(y)) \qquad (6)$$

Equation (6) means that the degree of similarity between *x* and *y* is equal to the truth degree of the following proposition:

$$\exists l \in L, x \in M(l) \wedge y \in M(l) \qquad (7)$$

Hence the similarity relation defined in (5) is meaningful on X. The fuzzy nominal scale can then be defined by the symbolism:

$$C = \langle X, LF(L), D, \{\sim\}, \{\equiv\}, \{(\sim, \equiv)\}\rangle \qquad (8)$$

where:

- *X* refers to a set of object states with ~ the only relation on it.
- LF(L) is the subset defined by (1), with the fuzzy equivalence relation ≡, the only relation defined on it.
- *D* is the fuzzy description mapping.
- ~ and ≡ are the similarity relations defined in (4) and (5).

The relation ~ is L-transitive because it has been defined with the Lukasiewicz t-norm. Thus, the operator defined on symbols by (9) is a distance operator.

$$d(A, B) = 1 - \mu_{\equiv}(A, B) \qquad (9)$$

This means that if a fuzzy description of a physical state of an entity can be obtained, the associated scale contains a fuzzy equivalence relation on fuzzy descriptions and a distance between fuzzy descriptions.

The behaviour of this distance operator is illustrated on the following example of fuzzy nominal scale:

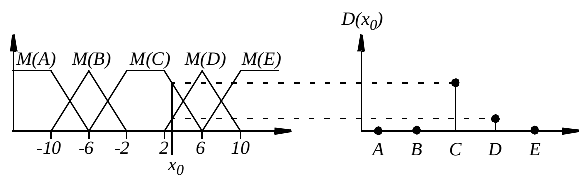

Fig. 1. An example of fuzzy nominal scale and its fuzzy description and meaning mappings

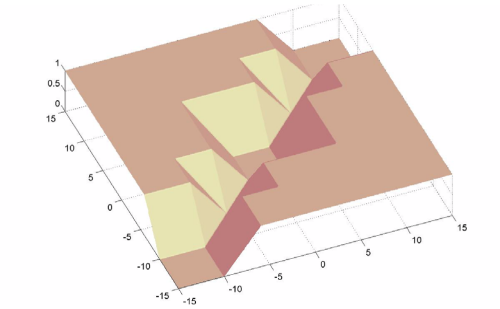

Fig. 2. $d(D(x),D(y))$ for x and $y \in [-15,15]$

Fig. 2. shows the result of $d(D(x),D(y))$ for *x* and *y* ranging from *-15* to + *15*. It can be seen that this distance is often equal to 1. This is the main drawback of this operator. Consequently, the distance between B and D (which is equal to $d(D(-6),D(6))$) is equal to the distance between A and E (which is equal to $d(D(-10),D(10))$) although these two latter terms have more "distant" meanings than the first ones.

But this drawback only arises with mono-dimensional fuzzy nominal scales having lexical sets of more than three terms. Such fuzzy nominal scales are not used in the application presented in next sections.

3. THE FUZZY GLOVE

In the field of Human Computer Interaction, there is a big interest in developing natural interfaces using human communication modalities. One of these modalities is communication through gestures. Gestures can be described with words. For example, the *pointing* hand configuration (see Fig. 3.) can be described quite simply as the one with a straight index, all other fingers being folded. This particularity of gestures makes it possible to describe them with fuzzy nominal scales. A fuzzy sensor called the *fuzzy glove* has been designed for this purpose in [7].

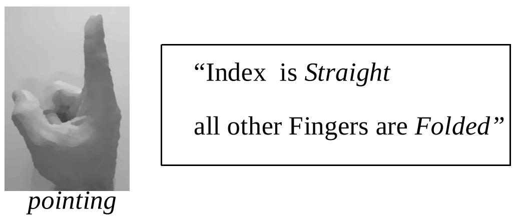

Fig. 3. The *pointing* gesture can be described with words

3.1. Structure of the fuzzy glove

The fuzzy glove provides a linguistic description of the hand configuration that is consistent with human perception of the hand. This fuzzy sensor is composed of two main elements as shown in Fig. 4.

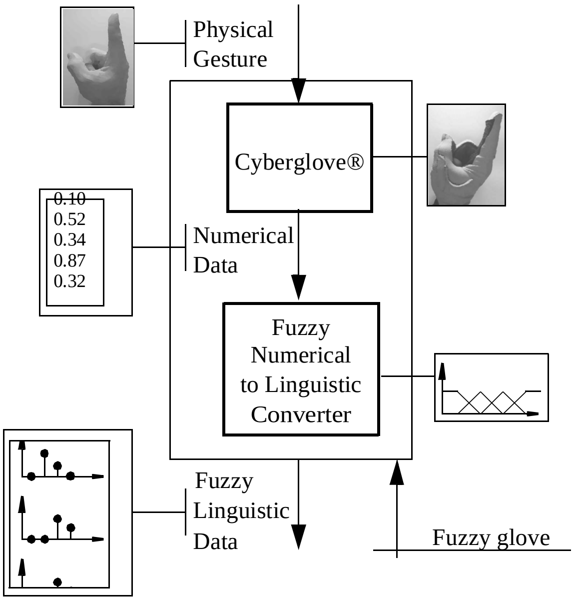

Fig. 4. Structure of the fuzzy glove

The acquisition of gestures is made with a data glove. The Cyberglove® developed by Kramer [8] has been chosen because it gives accurate and independent measurements of the different joint angles. It is made of 18 bending sensors (Fig. 5.). The wrist yaw and pitch sensors are useful for hand orientation measurements. They are not used here because the orientation of the hand is not taken into account. The palm arch sensor measures the cupping of the hand near the pinkie finger, a piece of information which is not used in our model of human perception of the hand.

Numerical data is entered in a numerical to linguistic converter which gives the linguistic descriptions of the finger configurations. The association of this numerical to linguistic converter with the CyberGlove® is called the fuzzy glove.

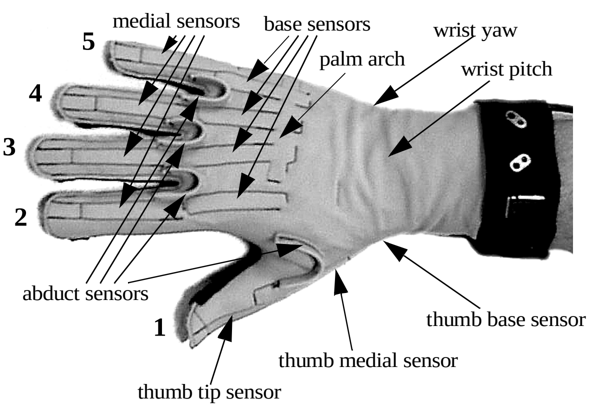

Fig. 5. The Cyberglove® and finger numbering

3.2. Linguistic variables

The linguistic description of the hand is made of nine linguistic variables respectively associated with nine fuzzy nominal scales. Lexical sets are chosen to have simple and easily understandable descriptions. This means that given the linguistic description of the hand configuration, anyone should be able to understand its shape or to realize it.

Two linguistic variables are used for the thumb description: the *thumbConfiguration* describes the stretching of the thumb and the *thumbOrientation* describes its orientation in the hand reference frame. Four variables are used for the description of the other fingers: the $finger_i Configuration$ (i = 2,...,5). They describe the shape of the corresponding finger. The last three variables are used for the description of the relative spacing of fingers 2/3, 4/3 and 5/4. They are respectively named *indexAbduct, ringAbduct* and *pinkieAbduct*.

The *thumbConfiguration* variable takes its values in the lexical set $L_{thumbConfiguration}$ = {*Folded*, *Straight*}.

The *thumbOrientation* variable takes its values in the lexical set $L_{thumbOrientation}$ = {*External*, *Aside*, *InternalBelow*, *InternalAbove*, *Ahead*}. These orientations are illustrated in Fig. 6..

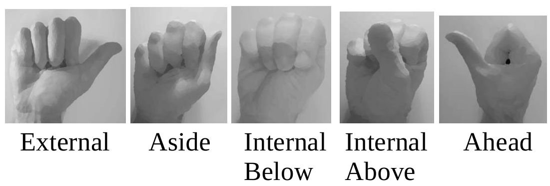

Fig. 6. . The five orientations of the thumb

The four $finger_i Configuration$ variables take their values in the lexical set $L_{fingerConfiguration}$ = {*Folded*, *Square*, *Round*, *Claw*, *Straight*}. These configurations are illustrated in Fig. 7..

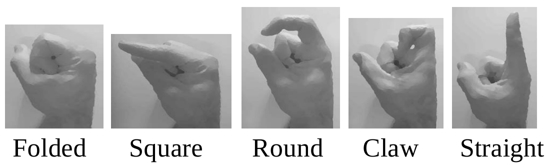

Fig. 7. . The five configurations of a finger

The three *fingerAbduct* variables take their values in the lexical set $L_{fingerAbduct}$ = {*Separated*, *Together*}.

3.3. Distance between hand postures

The configuration of the hand at a given time *t* is called a hand posture. The nine linguistic descriptions provided by the fuzzy glove at time *t* form the linguistic description of this hand posture. Hence, a hand posture can be written as:

$$P_i = (D_{i,1}, D_{i,2}, D_{i,3}, \ldots, D_{i,9}) \quad (10)$$

There are many ways to define a distance operator between two hand postures from the distance operators between their respective linguistic descriptions. In this paper, two distance operators will be studied:

$$d_{\Sigma}(P_1, P_2) = \sum_{j=1}^{9} d(D_{1,j}, D_{2,j}) \tag{11}$$

$$d_{sup}(P_1, P_2) = \sup_{j=1,\ldots,9} d(D_{1,j}, D_{2,j}) \tag{12}$$

These operators can have very different behaviours when comparing postures. The following example illustrates such a case. Three postures have been measured: *pointing*, *gunHand* and *flatHand*. Postures *pointing* and *gunHand* are quite similar, they differ only by the configuration of the thumb. Posture, *pointing* and *flatHand* are very dissimilar because they differ in all fingers except the index.

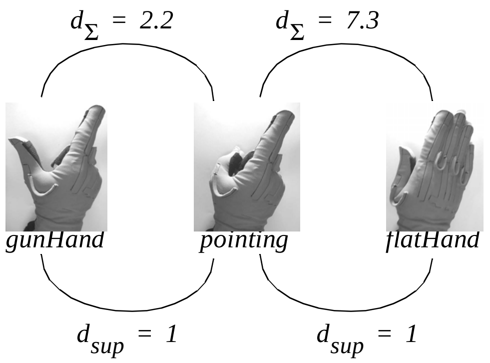

Fig. 8. Three postures illustrating the difference between d_{Σ} and d_{sup}.

Fig. 8. illustrates clearly the difference between the two operators: as soon as there is a difference between two postures, d_{sup} saturates at one and doesn't give information about the magnitude of this difference. Hence d_{Σ} seems to be more appropriate for the comparison of postures. However, it is proposed to classify gestures using these two distance operators.

4. DYNAMIC GESTURE RECOGNITION

An algorithm for speed independent motion recognition based on a resampling technique has been introduced by H. Sawada in [12]. A motion is considered as a sequence of 3D positions measured with equal time intervals. Such a sequence forms a trajectory in a 3D numerical space. The trajectories are normalized and can then be compared with prototypes. The comparison is made using a proximity relation.

This recognition procedure only uses distance operators and a proximity relation, so it can be applied on a fuzzy nominal scale.

4.1. Resampling technique

A gesture is a sequence of hand postures:

$$ges = \{P_1, P_2, \ldots, P_{n(ges)}\} \tag{13}$$

The size of the sequence depends on the corresponding gesture: some gestures are longer to execute than others. Moreover, the same gesture can be made more or less slowly by an operator. Thus, the size of the sequence also depends on the speed of execution of the corresponding gesture. Hence, to be able to compare these sequences, they first must be normalized to the same size.

A standard size *N* is defined and the sequences are then resampled in order to have *N* equal displacement intervals instead of equal time intervals (see Fig. 9.).

.

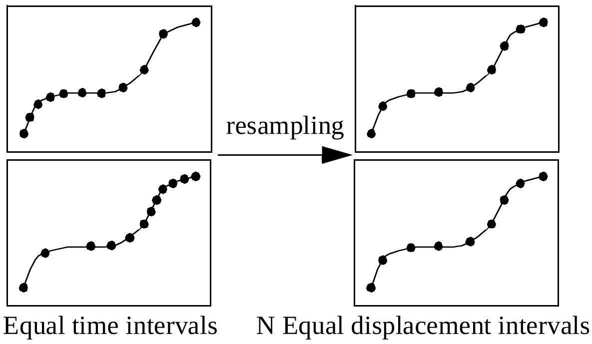

Fig. 9. Resampling principle

The length of a sequence is computed using (14).

$$L(ges) = \sum_{i=2}^{n(ges)} d(P_{i-1}, P_i) \tag{14}$$

A normalized segment length l_{normal} is calculated by dividing *L(ges)* by *N-1*. Then the normalized sequence is extracted from the original by selecting the postures the closest to $i \times l_{normal}$ for i = 0 to N-1.

Once all gestures have been normalized, they can be compared to each other using the proximity operator:

$$d(ges_1, ges_2) = \sum_{i=1}^{N} d(P_{1,i}, P_{2,i}) \tag{15}$$

where $P_{1,i}$ is the ith posture of the normalized sequence of gesture ges_1.

This resampling procedure works well if the sampling rate of the data glove is high compared to the execution speed of a gesture. The number of postures in the normalized gesture must also be big enough to satisfy a kind of Shannon sampling theorem.

A normal size N must be chosen and must be the same for all gestures. This obliges us to compare gestures having approximately the same length. Fig. 10. illustrates the case of two gestures having significantly different length.

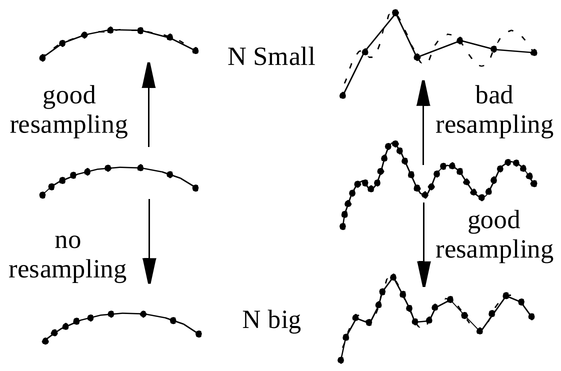

Fig. 10. resampling procedure enforce to have gestures of the same size.

5. EXPERIMENTAL RESULTS

The algorithm presented in section 4. is applied for the classification of eight dynamic gestures. These gestures are presented in Fig. 11.

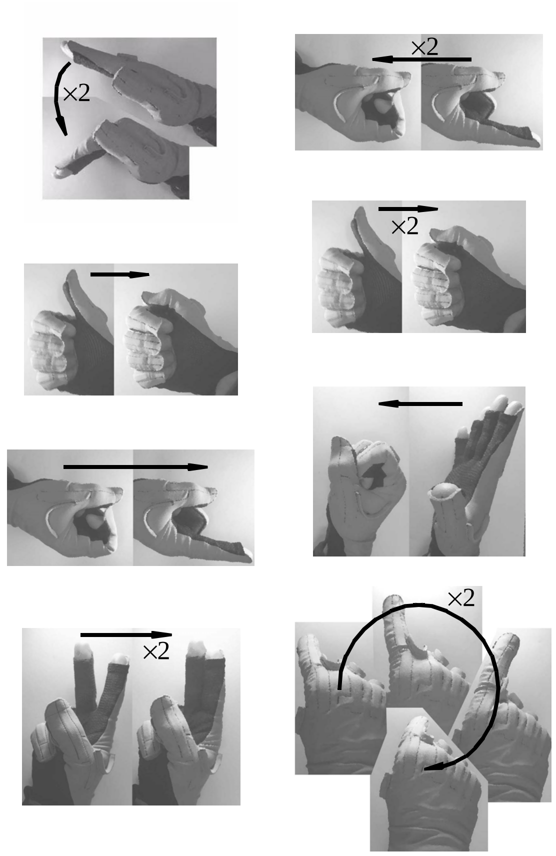

Fig. 11. The eight dynamic gestures to classify. "×2" means the movement is repeated two times

Three versions of each gestures are recorded. The first recording is the prototype. The second and third recording are the gestures to be classified. These two versions of the gestures are extracted from the continuous flow of postures provided by the fuzzy glove. This is the segmentation step.

The segmentation method is illustrated on Fig. 12. When the continuous flow of postures gets close to the beginning posture of a gesture, the recording starts. When it arrives close to the ending posture of this same gesture, the recording stops. This procedure is done manually but could be done automatically.

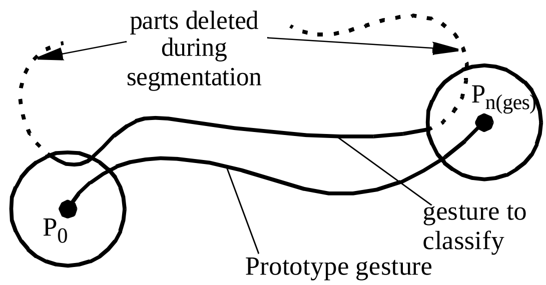

Fig. 12. Illustration of the segmentation procedure

Each gesture is compared to the eight prototypes with the d_Σ and the d_{sup} operators (see equations (11) and (12)). A gesture is classified as the nearest prototype. The experiment is repeated with two different sampling periods: the first one is 50ms and the second one is 100ms. This second sampling period is interseting because it quite long compared to the finger speed. It allows us to test the behaviour of the recognition system with a critical sampling.

The recognition system responded well because all the gestures were recognized, whatever the sampling period and the distance operator.

6. CONCLUSION AND FUTURE WORKS

It has been shown that it is possible to perform signal processing on fuzzy nominal scales. Classification of dynamic gestures has been performed succesfully with an algorithm based on distances between linguistic data. The problem of the aggregation of such distance operators was considered and two solutions have been envisaged. Hence to distance operators between hand postures have been designed. Surprisingly, there is no significant difference in the results given by the operators d_Σ and d_{sup} . They almost have the same behaviour when processing these data.

However, from the discussion of section 3, it seems that the d_{sum} operator is the most promising one . But these two solutions are quite empirical and it would be interesting to have more theoretical foundations for such operators. It would also be interesting to study other distance operators which would not saturate to one and thus would not have the drawback described in section 2.

The classification procedure applied in this paper used a resampling method based on the distance operator. The good recognition of gestures shows that this resampling procedure is efficient even with a sampling period of the gestures of 100 ms, which is quite long compared to the speed of fingers. Although this method has proved to be efficient on the gestures defined in Fig. 11., it could be less efficient for the classification of gestures having very different sizes (see section 4). In future works, other classifications algorithms will be tested like Dynamic Time Warping and Hidden Markov Models.

REFERENCES

[1] L. Mari, "The duality of sensing and coding as a fundamental topic of information science", *Proc. of the World Conf. on Systemics, Cybernetics and Informatics (SCI 2001)*, Vol. XVI, Orlando, USA, pp. 177-180, 2001.

[2] L. Finkelstein., "Representation by symbol systems as an extension of the concept of measurement", *Kybernetes*, Vol. 4, pp. 215-223, 1975.

[3] H. Watanabe,S. Ryumae,Y. Oka,K. Shigemasu,Y. Moritora, "A study of nominal scale", *Proc. of IMEKO TC7 int. symp. on AIMaC'91*, Kyoto, Japan, pp. 295-302, 1991.

[4] H. Watanabe, "Theory of classification to reinterpret the nominal scale for our better understanding of measurement information", *Proc. of the World Conf. on Systemics, Cybernetics and Informatics (SCI 2001)*, Vol. XVI, Orlando, USA, pp. 191-198, 2001.

[5] E. Benoit, L. Foulloy, "Fuzzy Nominal Scales", *IMEKO TC7 Symposium*, Cracow, Poland, pp. 21-25, 2002.

[6] E. Benoit, L. Foulloy, "Towards fuzzy nominal scales", *Measurement*, vol. 34, no. 1 (Fundamental of Measurement), pp. 49-55, 2003.

[7] T.Allevard, E. Benoit, L. Foulloy, "Fuzzy glove for gesture recognition", Proc. of the XVII IMEKO world congress, Dubrovnik, Croatia, pp. 2026-2031, 2003.

[8] J.F. Kramer, "The Talking Glove: hand-gesture-to-speech using an instrumented glove and a tree-structured neural classifying vector Quantizer", *PhD Thesis*, Stanford University, 1996.

[9] G. Mauris, E. Benoit, L. Foulloy, "Fuzzy symbolic sensors - From concept to applications", *Measurement*, no. 12, pp. 357-384, 1994.

[10]S. Gottwald, "Fuzzy Sets and Fuzzy Logic", Vieweg & Sohn, Wiesbaden, 216 p., 1993.

[11]E. Benoit, "Capteurs symboliques et capteurs flous : un nouveau pas vers l'intelligence", PhD Thesis, Université de Savoie, 160p., 1993.

[12]E. Benoit, T. Allevard, T. Ukegawa, H. Sawada, "Fuzzy Sensor for Gesture Recognition Based on Shape Recognition of Hand", *Proc. of Int. Symp. on Virtual Environments, Human-Computer Interfaces, and Measurement Systems (VECIMS)*, Lugnano, Switzerland, pp. 63-67, 2003.

Authors: Thomas ALLEVARD, thomas.allevard@univ-savoie.fr. Eric BENOIT, eric.benoit@univ-savoie.fr. Laurent FOULLOY, laurent.foulloy@univ-savoie.fr. LISTIC-ESIA, Université de Savoie, B.P. 806, 74016 Annecy cedex, France