Recognizing Hand Gestures Using a Fuzzy Rule-Based Method, and Representing them with HamNoSys *

Antônio C. R. Costa[1], **Graçaliz P. Dimuro**[1] **and Benjamin C. Bedregal**[2]

[1]Programa de Pós-Graduação em Informática, Universidade Católica de Pelotas, Pelotas, Brazil
[2]Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte, Natal, Brazil
{rocha,liz}@atlas.ucpel.tche.br, bedregal@dimap.ufrn.br

Abstract

This paper introduces a fuzzy rule-based method for the recognition of hand gestures acquired from a data glove, and a way to show the recognized hand gesture using the graphical symbols provided by the HamNoSys notation system. The method uses the set of angles of finger joints for the classification of hand configurations, and classifications of segments of hand gestures for recognizing gestures. The segmentation of gestures is based on the concept of "monotonic" gesture segment, i.e., sequences of hand configurations in which the variations of the angles of the finger joints have the same tendency (either non-increasing or non-decreasing), separated by reference hand configurations that mark the inflexion points in the sequence. Each gesture is characterized by its list of monotonic segments. The set of all lists of segments of a given set of gestures determine a set of finite automata that recognize such gestures. For each gesture, a sequence of HamNoSys symbols representing the reference hand configurations of the gesture is produced as an output.

1. Introduction

Fuzzy sets were introduced in 1965 by Zadeh (1965) for representing vagueness in everyday life, providing an approximate and effective means for describing the characteristics of a system that is too complex or ill-defined to be described by precise mathematical statements. In a fuzzy approach the relationship between elements and sets follows a transition from membership to non membership that is gradual rather than abrupt.

Fuzzy set theory is the oldest and most widely used theory for soft computing, which deals with the design of flexible information processing systems (Mitra and Pal, 2005). A fuzzy system implements a function (usually nonlinear) of *n* variables, given by a linguistic description of the relationship between those variables.

Figure 1 illustrates the architecture of standard fuzzy systems. The *fuzzificator* computes the membership degrees of the crisp input values to the linguistic terms (fuzzy sets) associated to each input linguistic variable. The *rule base* contains the inference rules that associate linguistic terms of input linguistic variables to linguistic terms of output linguistic values. The *information manager* is responsible for searching in the rule base which rules are applicable for the current input. The *inference machine* determines the membership degrees of the output values in the output sets, by the application of the rules selected in the rule base. The *defuzzificator* gives a single output value as a function of the output values and their membership degrees to the output sets.

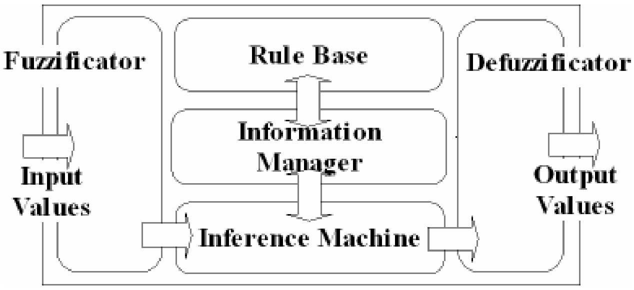


Fig. 1. Architecture of standard fuzzy systems.

The HamNoSys notation system (Prillwitz, S. et al.; 1989) is a graphical system for the symbolic representation of linguistic features of sign languages. In particular, it has a flexible set of symbols for the representation of hand configurations.

In this paper, we propose a *fuzzy rule-based method* for the recognition of hand gestures acquired from a data glove, and the representation of the recognized gestures as sequences in HamNoSys. We apply the method to the recognition of a sample hand gesture of LIBRAS, the *Brazilian Sign Language* (Brito, 1995).

The paper is organized as follows. In Sect.2, we introduce our fuzzy rule-based method for hand gesture recognition. A case study is discussed in Sect.3, with the recognition of a LIBRAS hand gesture. Section 4 shows the procedure for the automatic representation of hand gestures in HamNoSys. Section 5 is the Conclusion.

2. The Fuzzy Rule Based Method for Hand Gesture Recognition

The idea is to recognize some hand gestures with data obtained from a data glove with 15 sensors, as shown in Fig. 2. It is assumed that the data has been smoothed for jitter noise. The fingers are labelled as: F1 (little finger), F2 (ring finger), F3 (middle finger), F4 (index finger) and F5 (thumb). The joints in the fingers are labelled as J1 (the knuckle), J2 and J3, for each finger. Separations between fingers Fi and Fj are labelled as Sij.

Since any hand gesture can be represented as a sequence of frames, a hand gesture using a data glove is represented as a sequence of hand configurations, one for each discrete time instant. That is, at each time instant, the data glove sensors should provide the set of angles of joints and finger separation that characterizes a hand configuration.

In order to simulate this data transfer, a generator of hand configurations was implemented, generating monotonic sequences of handshapes: at each instant one hand configuration represented by a tuple of angles corresponding to each sensor shown in Fig. 2:

```
((F1J1,F1J2,F1J3), S12, (F2J1,F2J2,F2J3), S23,
 (F3J1,F3J2,F3J3), S34, (F4J1,F4J2,F4J3), S45,
 (F5J1,F5J2,F5J3) )
```

Given a hand configuration *c* and a sensor *s*, denote the value of a sensor angle by $s(c)$, e.g., F1J1(c), S45(c) etc.

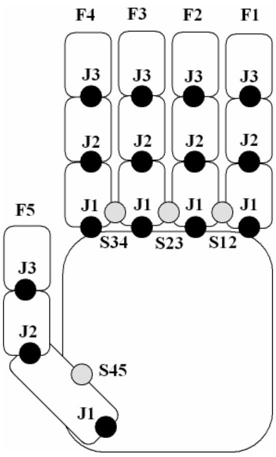


Fig. 2. Localization of sensors in the data glove.

2.1. Fuzzification

To each sensor corresponds a linguistic variable, whose values are linguistic terms representing typical angles of joints and separations. For the joints in the fingers (linguistic variables F1J1, F1J2, F1J3 etc.) the linguistic terms are: STRAIGHT (St), CURVED (Cv) and BENT (Bt). For the separations between fingers F1 and F2, F2 and F3, F4 and F5 (linguistic variable S12, S23, S45), the linguistic terms are: CLOSED (Cd), SEMI-OPEN (SOp) and OPEN (Op). For the separations between fingers F3 and F4 (linguistic variable S34), the linguistic terms are: CROSSED (Cr), CLOSED (Cd), SEMI-OPEN (Sop) and OPEN (Op). Figures 3, 4, 5, 6 and 7 show the fuzzification adopted for those variables.

2.2. The Recognition Process

The hand gesture recognition process is divided into four steps: (1) recognition of finger configurations, (2) recognition of hand configurations, (3) segmentation of the gesture in monotonic hand segments and (4) recognition of the sequence of monotonic hand segments.

For the step 1 (*recognition of finger configurations*), 27 possible finger configurations are considered. These configurations are codified in the following format: XYZ, where X is the value of the linguistic variable corresponding to the first joint J1, Y is the value of the linguistic variable corresponding to the second joint J2 and Z is the value of the linguistic variable corresponding to the third joint J3. For example, StStSt is used to indicate that the three joints are STRAIGHT, StCdCd indicates that the first joint is STRAIGHT whereas the others are CURVED etc.

The hand configuration is the main linguistic variable of the system, denoted by HC, whose linguistic terms are names of hand configurations, which names are application dependent. For instance, in Sect. 3, names of Brazilian Sign Language (LIBRAS) hand configurations (see Fig. 9) were used for such linguistic terms.

The 27 possible finger configurations determine 27 inference rules that calculate membership degree of each finger to each configuration. For example:

```
If F4J1 is STRAIGHT and F4J2 is CURVED and
   F4J3 is CURVED
Then F4 is StCdCd
```

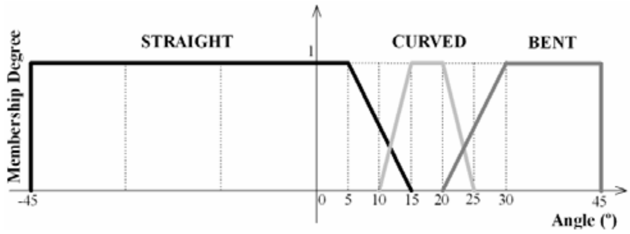


Fig. 3: Fuzzification of the linguistic variable of the joint F5J2 in the thumb finger F5.

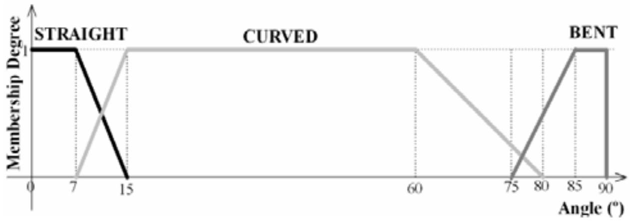


Fig. 4. Fuzzification of the linguistic variables of remaining finger joints

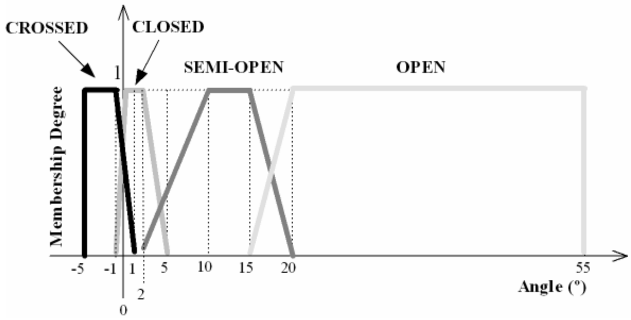


Fig. 5. Fuzzification of the linguistic variable of the separation S34 between the middle finger F3 and the index finger F4.

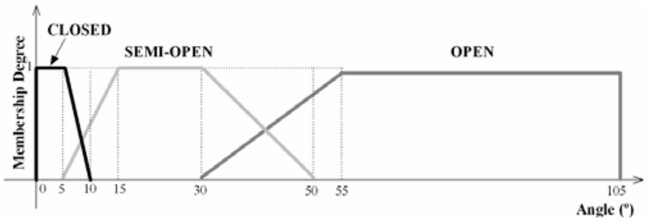


Fig. 6. Fuzzification of the linguistic variable of the separation S45 between the index finger F4 and the thumb finger F5.

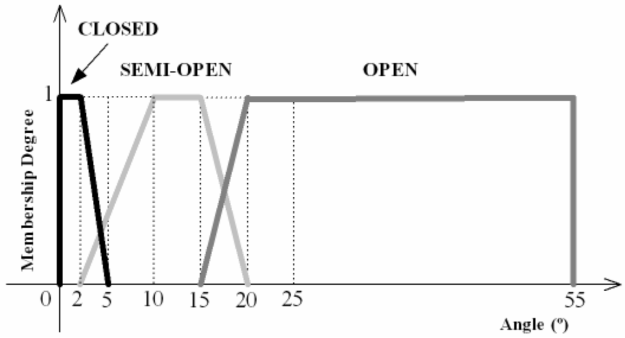


Fig. 7. Fuzzification of the linguistic variables of the separations between remaining fingers.

The next step is 2 (*recognition of hand configurations*), where the hand configuration is determined, considering each finger configuration and each separation between fingers. For example, the rule for the hand configuration [G] of LIBRAS (see Fig. 9) is described below:

```
If F1 is BtBtSt and S12 is Cd and F2 is BtBtSt
   and S23 is Cd and F3 is BtBtSt
```

```
    and S34 is Cd and F4 is StStSt
    and S45 is Cd and F5 is StStSt
Then HC is [G]
```

We note that since the hanshape is recognized as a unity, no co-articulation problem arises for its recognition.

In 3 (*segmentation of the gesture in monotonic hand segments*), we divide each gesture in a sequence of *k* limit hand configurations $l_1,\ldots,l_k$, where l_1 is the initial configuration and l_k is the terminal configuration. The limit configurations are such that, for each *c* between l_i and l_{i+1}, and for each sensor *s*, $s(c)-s(l_i)$ has the same sign of $s(l_{i+1})-s(l_i)$, for $i=1,\ldots,k-1$ (a difference of 0 is compatible with both negative and positive signs).

The limit hand configurations are the points that divide the gesture into monotonic segments, that is, segments in which each sensor produces angle variations with constant (or null) sign. For each monotonic segment l_il_{i+1}, l_i and l_{i+1} are its initial and terminal hand configurations, respectively.

The procedure for step 3 is the following. To find any monotonic segment l_il_{i+1}, the next *n* configurations sent by the data glove after l_i are discarded, until a configuration c_{n+1}, such that the signs of $s(c_{n+1})-s(c_n)$ and $s(c_n)-s(l_1)$ are not the same (or, c_{n+1} is the last configuration of the gesture). Then, c_n (resp., c_{n+1}) is the terminal hand configuration l_{i+1} of the considered monotonic segment, and also coincides with the initial configuration of the next segment $l_{i+1}l_{i+2}$ (if there is one). The process starts with $l_1 = l_1$, which is the initial gesture configuration, and is repeated until the end of the gesture, generating the list of *k* limit hand configurations.

In 4 (*recognition of the sequence of monotonic hand segments*), the recognition of each monotonic segment l_il_{i+1} is performed using a list of reference hand configurations $r_1, r_2, \ldots, r_m$ that characterizes the segment, where r_1 and r_m are the initial and terminal hand configurations of the segment, respectively. A monotonic segment is recognized by checking that it contains its list of reference hand configurations. The process is equivalent to a recognition based on a linear finite automaton, where $l_i = r_1$ and $l_{i+1} = r_m$ and the transition function is shown in Fig. 8.

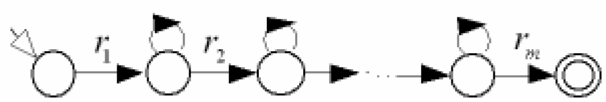


Fig. 8. Automaton for the recognition of monotonic segments.

3. Case Study: a Hand Gesture of LIBRAS

LIBRAS is the Brazilian Sign Language. As in any other sign language, the main parameters that characterize its phonological units are: the configurations of the hands used in the gestures, the main spatial location (relative to the persons who is signing) where the movements of the gestures are performed, the different movements (of the fingers in the hand, of the hands and arms in the space, of the whole body) that constitute the gesture, the facial expressions that express different syntactic, semantic and pragmatic marks during the production of the signs etc.

To support that recognition process, a reference set of hand configurations is usually adopted, driven either from the linguistic literature on sign languages, or dynamically developed by the experimenters with an ad hoc purpose. For our purposes, we have chosen a standard set of hand configurations (some of them shown in Fig. 9, taken from the linguistic literature on LIBRAS (Brito, 1995).

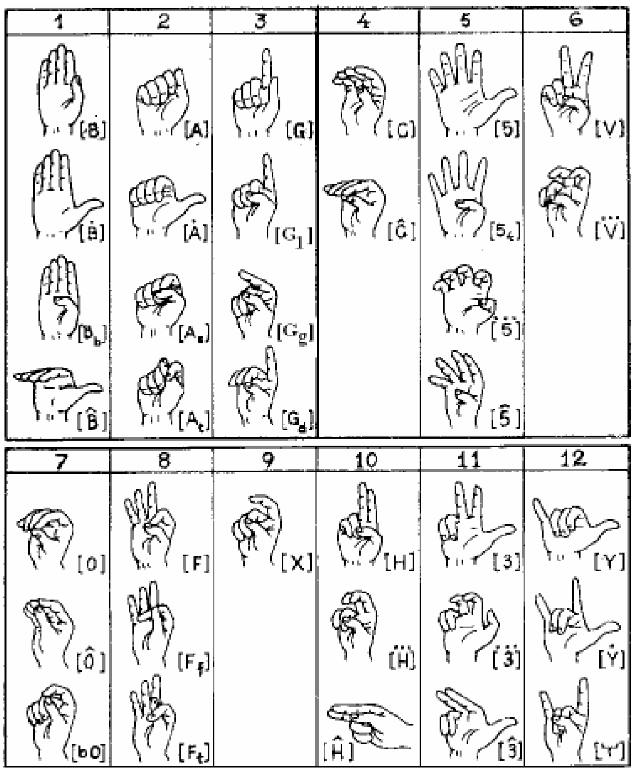


Fig.9. Some LIBRAS hand configurations.

Since we take the set of hand configurations from the literature, our method requires that each sign be thoroughly characterized in terms of its monotonic segments and the sequences of hand configurations that constitute such segments, and that the identification of the monotonic segments and hand configurations be manually provided to the system. Of course, a capture device such as a data glove can be used to help to identify the typical values of the angles of the finger joints, but the final decision about the form of the membership functions that characterize the linguistic terms used in the system has to be explicitly taken and manually transferred to the system.

We illustrate here the application of the method by the definition of the necessary parameters for the recognition of the hand gestures that constitute the sign CURIOUS, in LIBRAS. CURIOUS is a sign performed with a single hand placed right in front of the dominant eye of the signer, with the palm up and fingers oriented in the forward direction. The initial hand configuration is the one named [G1] in Fig. 9. The gesture consists of the monotonic movement necessary to perform the transition from [G1] to [X] and back to [G1] again, such movements been repeated a few times (usually two or three). Thus, a possible analysis of the hand gestures that constitute the sign CURIOUS in LIBRAS is:

Initial configuration: [G1]
Monotonic segment S1: [G1]-[G1X]-[X]
Monotonic segment S2: [X]-[G1X]-[G1]
State transition function for the recognition automaton: see Fig. 10.

To support the recognition of the monotonic segments of CURIOUS, we have chosen to use one single intermediate hand configuration, [G1X]. It is an intermediate configuration that does not belong to the reference set (Fig. 9) and whose characterization in terms of the set of membership functions for linguistic terms was defined in an ad hoc fashion, for the purpose of the recognition of CURIOUS. Together with [G1] and [X], it should be added to the list of hand configurations used by the recognition system.

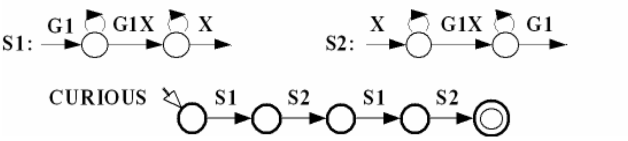


Fig. 10. Automaton for the recognition of hand gestures of the sign CURIOUS.

4. Representation of Hand Gestures in HamNoSys

HamNoSys (Hamburg Notation System) is a graphical system for the representation of linguistic features of sign languages. In particular, it has a flexible set of symbols for the representation of hand configurations.

Figure 11 presents the HamNoSys representation of some LIBRAS hand shapes.

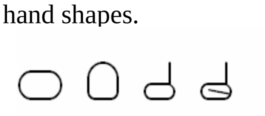


Fig. 11. The [A], [B], [G] and [G1] hand shapes respectively represented in HamNoSys.

It is easy to see how one can associate in a table the lists of features characterizing hand configurations with their representation in HamNoSys, so that the hand gesture recognition system can produce as output a sequence of HamNoSys notations.

Figure 12 shows the representation of the hand gesture of the sign CURIOUS.

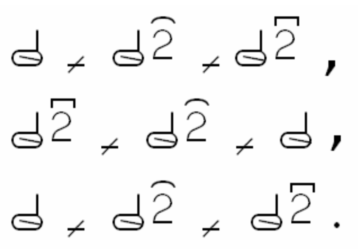
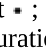


Fig. 12. The three monotonic segments of the LIBRAS hand gesture of the sign CURIOUS represented in HamNoSys.

Notice the following conventions in Fig. 12: the reference hand configurations within a monotonic gesture segment are separated by the symbol  while the segments themselves are separated by the comma  and ended by the point  ; the initial configuration is [G1]; the other hand configurations are variations of [G1], denoted by indicating the finger whose configuration varies:

 indicates that the index figure is curved, while

 indicates that the index figure is bent.

Notice also that the order for the enumeration of fingers used in this paper (Fig. 2) is different from that adopted in HamNoSys.

5. Conclusion and Final Remarks

We presented a fuzzy rule-based for the recognition of hand gestures. The method is highly dependent on a detailed previous analysis of the features of the gestures to be recognized, and on the manual transfer of the results of that analysis to the recognition system. This makes it suitable for the application to the recognition of hand gestures of sign languages, because of the extensive analysis that linguists that have already done of those languages.

Prototypes of a random gesture generator and of the gesture recognizer were implemented in the programming language Python. In the fuzzification process, we considered only trapezoidal fuzzy sets and the minimum (or Gödel) t-norm, motivated by simplicity. The output of the recognition system was fed into a simple translator able to render the recognized hand gestures as they are annotated in the HamNoSys notation system.

Future work should develop in two directions: the recognition of arm gestures, by including the analysis of the angles of arm joints; the application of the recognizer to support computer systems controlled by pre-defined sets of signs, phonologically specified in HamNoSys.

6. References

Al-Jarrah, O., Halawani, A. (2001) Recognition of gestures in arabic sign language using neuro-fuzzy systems. *Artificial Intelligence* **133**: 117-138.

Bimber, O. (1999) Continuous {6DOF} gesture recognition: A fuzzy-logic approach. In: *Proc. of 7th. Intl. Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media* (pp. 24-30}.v.1

Binh, N.D., Ejima, T. (2005) Hand gesture recognition using fuzzy neural network. *In: Proc. of ICGST Intl. Conf. Graphics, Vision and Image Proces*. (pp. 1-6), Cairo.

Brito, L.F. (1995) *Por uma Gramática de Línguas de Sinais.* Rio de Janeiro: Tempo Brasileiro.

Fang, G., Gao, W., Zhao, D. (2004) Large vocabulary sign language recognition based on fuzzy decision trees. I*EEE Trans. Syst., Man and Cybern*. **34**: 305-314

Hanmandlu, M., Yusof, M.H.M., Madasu, V.K. (2005) Off-line signature verification and forgery detection using fuzzy modeling. *Pattern Recognition* **38:**341-356.

Kwak, K., Pedrycz, W. (2005) Face recognition using a fuzzy fisherface classifier. *Pattern Recognition* **38**: 1717-1732.

Lin, C.T., Lee, C.S.G. (1996) Neural Fuzzy Systems: A neuro-fuzzy synergism to intelligent systems. Upper Saddle River: Prentice Hall.

Mitra, S., Pal, S.K. (2005) Fuzzy sets in pattern recognition and machine intelligence. *Fuzzy Sets and Systems* **156**: 381-386.

Prillwitz, S. et al. (1989) HamNoSys: Hamburg Notation System for Sign Languages. Hamburg: Signum.

Su, M. (2000) A fuzzy rule-based approach to spatio-temporal hand gesture recognition. *IEEE Transactions on Systems, Man and Cybernetics, Part C* **30**: 276-281

Zadeh, L.A. (1965) Fuzzy sets. *Information Control* **8**: 338-353.

* This work is partially supported by FAPERGS and CNPq (Proc. 470871/2004-0, Proc. 470556/2004-8).