# Recent Advances in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
# Volume II: Applications


**Editors**

Krassimir T. Atanassov
Władysław Homenda
Olgierd Hryniewicz
Janusz Kacprzyk
Maciej Krawczak
Zbigniew Nahorski
Eulalia Szmidt
Sławomir Zadrożny


SRI PAS IBS PAN

# Recent Advances in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics Volume II: Applications

**Systems Research Institute**
**Polish Academy of Sciences**

# Recent Advances in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
# Volume II: Applications

Editors

Krassimir T. Atanassov
Władysław Homenda
Olgierd Hryniewicz
Janusz Kacprzyk
Maciej Krawczak
Zbigniew Nahorski
Eulalia Szmidt
Sławomir Zadrożny

IBS PAN SRI PAS

Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
www.ibspan.waw.pl

ISBN 9788389475367

# Structure of a lip-reading system based on IF-Sets

**Krzysztof Dyczkowski**
Faculty of Mathematics and Computer Science,
Adam Mickiewicz University,
Umultowska 87, 61-614 Poznań, Poland
chris@amu.edu.pl

**Abstract**

This paper presents the structure of a lip reading system. A method of encoding lip shapes is presented and a similarity/dissimilarity measure, based on IF-Sets, for comparison of lip shapes in a recognition process is proposed. The use of IF-Sets methodology makes it possible to compare various lip shapes in a bipolar way involving both similarities and dissimilarities. It also allows the estimation of a hesitation degree for lip similarity, which expresses ignorance in comparing them.

**Keywords:** Atanassov IF-Sets, lip reading, lip shape encoding and comparison, similarity/dissimilarity measure.

## 1 Introduction

Recognizing and comparing lip shapes is very important for applications in audio-visual speech recognition (cf. [12], [13]), lip reading (cf. [6], [9]) and phonetics research (cf. [8]). This article presents an outline of a project which deals with visual speech recognition based on the movements of the lips.

An initial version was implemented to check establish whether speech recognition from the movement of the lips can be achieved using a computer. In this system we assumed that we would not be translating from video to an orthographical record of the text, but only converting into merged audio, human would assess for comprehensibility.

***Recent Advances in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications*** *(K.T. Atanassow, W. Homenda, O. Hryniewicz, J. Kacprzyk, M. Krawczak, Z. Nahorski, E. Szmidt, S. Zadrożny, Eds.), IBS PAN - SRI PAS, Warsaw, 2010.*

The system consists of three main processes: Learning, Recognizing and Encoding. The recognizing and learning processes use the encoding process to extract numerical representations from frames of video input. The encoding process consists of a face detection module and lip-shape detection and encoding. In the learning process a knowledge base is built up. This means that groups of video frames with encoded lip-shapes are associated with their interpretations – recorded sounds. In the recognition process similarity measures are used to find best fitting lip-shapes representations and reproduce associated audio records. Figure 1 shows a schematic diagram of the entire lip-reading system.

We will present concepts relating to the numerical modelling of lip shapes and their comparison, employing the idea of IF-Sets. Using such concepts it is possible to compare lip shapes in a bipolar view. We can see how two lip shapes are similar to each other and simultaneously how they are dissimilar. We can also estimate our hesitation degree for lip similarity which expresses our ignorance in comparing them.

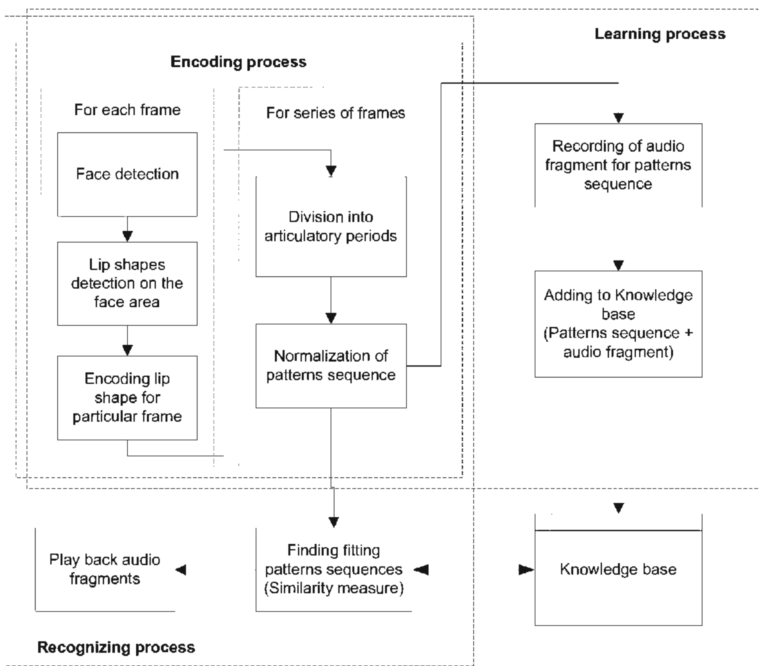



Figure 1: Architecture of the lip-reading system

## 2 Preliminaries

The notion of intuitionistic fuzzy sets (IF-Sets) was introduced by Atanassov in 1986 ([1]). It can be viewed as a generalization of fuzzy sets as introduced by Zadeh in 1965 ([16]). They deliver tools for modelling simultaneously positive and negative information approaches to different phenomena and enable us to estimate our level of ignorance about them.

**Definition 1** *An intuitionistic fuzzy set (IF-Set) $A$ in the universe $X$ is an object of the form $A = \{\langle x; \mu(x); \nu(x)\rangle : x \in X\}$ where $\mu$ and $\nu$ are degrees of membership and non-membership of each $x \in X$, respectively, and $0 \leq \mu(x) + \nu(x) \leq 1$ for each $x \in X$.*

*The number $\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$ is called the hesitation margin (or intuitionistic fuzzy index) of the IF-Set $A$.*

More details on the theory of Atanassov intuitionistic fuzzy sets can be found in [1], [2], [7], [11] and [14, 15].

Very interesting applications of IF-Sets for image processing and pattern recognition can be found in [3], [5] and [4].

## 3 The lip shape encoding

The main task in the process of building a lip shape model is to construct an encoded pattern of a given lip shape, that is representative and easy to compute and compare. The whole process is performed in an automatic manner. We used the OpenCV library (see [10]) which provided the necessary machineries of image processing for the implementation.

The procedure of lip shape pattern encoding is as follows.

In the first step, on the picture of the face we mark two points denoting the lip corners, which are denoted by $c_1$ and $c_5$. The segment linking the two points is divided into four parts limited by points $c_2$, $c_3$, $c_4$ through which segments parallel to the segment $c_1, c_5$ are drawn limited by the edges of the lips, setting test points $u_1$ $u_2$ for the upper lips and $l_1$, $l_2$ $l_3$ for the lower lips. In this manner a pattern for the lip arrangement is constructed (see Fig. 2).

In order to make this model independent of different sizes of lips or of the picture the ratios of the lengths of vertical and horizontal segments are encoded in the pattern, rather than the distances between individual marked points. (see Fig. 3).

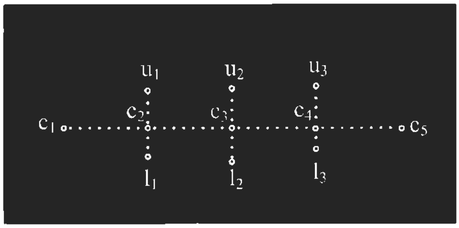



Figure 2: Control points on the lip shape

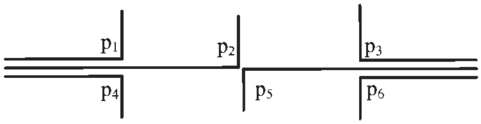



Figure 3: Coding pattern of the lip shape

**Definition 2** *Lip pattern $P$ is defined as the vector:* $P = (p_1, p_2, p_3, p_4, p_5, p_6)$ *where* $p_i \in [0,1]$ *and* $p_1 = \frac{|u1,c2|}{|c1,c2|}$, $p_2 = \frac{|u2,c3|}{|c1,c3|}$, $p_3 = \frac{|u3,c4|}{|c4,c5|}$, $p_4 = \frac{|l1,c2|}{|c1,c2|}$, $p_5 = \frac{|l2,c3|}{|c1,c2|}$, $p_6 = \frac{|l3,c4|}{|c4,c5|}$.

*Here $p_i$ will be called the i-th pattern part of $P$.*

It should be emphasized that an individual element of the lip shape pattern $p_i$ takes values from the period $[0,1]$, and if $p_i$ is closer to 1 then the vertical and horizontal distances are more equal. If $p_i = 0$ then there is no vertical deviation from the line connecting the lip corners.

# 4 The lip movements - division into articulatory periods

The proper division of continuous video input into informative periods is an important component of the system. Not only lip shapes, but also their changes in a time, are essential. The input video consists of frames. Every frame is described by the system with a given lip pattern. We obtain a conversion of the video into a sequence of numeric data. Then this data should be grouped into minimal informational sequences which will constitute the knowledge base. In performing this process we assume that the sequence of lip patterns is formed only from monotonic items, i.e. the movement of the lips takes place vertically only in one established direction.

Finally, sequences of lip patterns prepared in this way are normalized and audio covering this sequence is recorded. They constitute pairs: articulatory periods – audio representation, which are added to the knowledge-base.

# 5 Similarity measure

The appropriate structure of the similarity measure is the next very important element of the system, playing a large role in the learning and recognition processes.

The construction of a lip shape similarity measure involves three steps:

1. calculation of a distance between particular pattern parts of two lip patterns;
2. interpretation of the distance by means of similarity and dissimilarity degree (constructing the IF-Set of each distance between corresponding pattern parts);
3. aggregation to one similarity and one dissimilarity measure value.

**Definition 3** *Let $a_i, b_i$ be the $i - th$ pattern parts of the lip patterns $A$, $B$ respectively. The similarity degree $sim_i(A,B)$ of the $i - th$ pattern parts is defined as $sim_i(A,B) = (1 - |a_i - bi|)^2$ and the dissimilarity degree of the $i - th$ pattern parts is defined as $dis_i(A,B) = |a_i - b_i|^2$.*

**Definition 4** *Let $A$, $B$ be lip patterns. The IFS-LIP measure $L(A,B)$ is a triple $(SIM, DIS, HES)$ such as $L : A, B \to (SIM, DIS, HES)$ and*

$$SIM(A,B) = \frac{\sum_{i=1}^{6}(w_i * sim_i(A,B))}{\sum_{i=1}^{6} w_i},$$

$$DIS(A,B) = \frac{\sum_{i=1}^{6}(w_i * dis_i(A,B))}{\sum_{i=1}^{6} w_i},$$

$$HES(A,B) = 1 - SIM(A,B) - DIS(A,B),$$

$\forall_{A,B} \;\; 0 \leq SIM(A,B) + DIS(A,B) \leq 1$ *where $w_i$ is the weight representing the importance of the $i - th$ lip pattern.*

As can be noticed, we have applied the weighted average of membership degrees as an aggregation operator. The selection of appropriate values of $w_i$ is very important. If all $w_i$ values are equal to 1, all fragments of the pattern will affect the measure in the same way. However experimental results have suggested that central elements of the lip pattern should have greater weights assigned. Therefore we suggest using the vector of weights $w = (1, 2, 1, 1, 2, 1)$.

# 6 Example

Table 1 presents four pictures of lip arrangements with the corresponding patterns.

Table 1: Lip shapes with encoded patterns $P_1$ - $P_4$

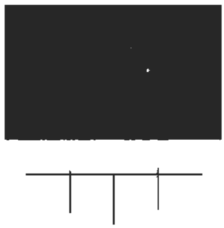
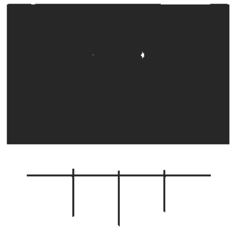
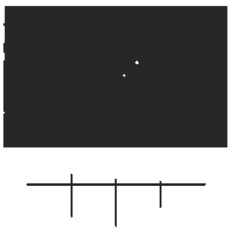
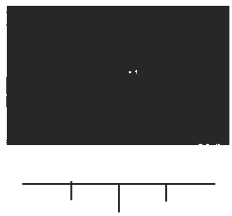
| | |
|---|---|
|  $P_1 = (0.05, 0, 0.14, 0.89, 0.58, 0.78)$ |  $P_2 = (0.15, 0.05, 0.12, 0.88, 0.54, 0.78)$ |
|  $P_3 = (0.22, 0.05, 0.05, 0.73, 0.47, 0.51)$ |  $P_4 = (0.13, 0, 0.03, 0.62, 0.44, 0.56)$ |

Table 2 presents values of the IFS-LIP measure between the lip arrangements in Table 1.

Table 2: Computed values of IFS-LIP measure

| | $p_1$ | $p_2$ | $p_3$ | $p_4$ |
|---|---|---|---|---|
| $p_1$ | (1,0,0) | (0.93,0,0.07) | (0.75,0.02,0.23) | (0.75,0.03,0.22) |
| $p_2$ | (0.93,0,0.07) | (1,0,0) | (0.81,0.02,0.17) | (0.78,0.02,0.2) |
| $p_3$ | (0.75,0.02,0.23) | (0.81,0.02,0.17) | (1,0,0) | (0.88,0,0.11) |
| $p_4$ | (0.75,0.03,0.22) | (0.78,0.02,0.2) | (0.88,0,0.11) | (1,0,0) |

After verification of the results, it can be concluded that the constructed measure provides a more compressive form of information about the relations between lip arrangements. Apart from the information about similarities between lip patterns (e.g. the relatively high similarity between $p_1$ and $p_2$, $p_6$ and $p_8$) we additionally obtain essential information about hesitation in determining the power of similarity (e.g. $p_1$ vs. $p_6$ and $p_1$ vs. $p_8$ where the hesitation margin is higher than the degree of similarity).

# 7 Summary

We have presented an overview of the lip-reading system on which we are currently working. The idea of encoding lip shapes into patterns and comparing such patterns to the suggested measure IFS-LIP appears to be very promising. We have attempted to utilize the approach suggested by Atanassov, which involves taking a dual view of the data. Our work presents the initial stage of more complex research on the modelling of lip shapes. The first results show that the applied solution can work in the process of automatic lip reading.


### Acknowledgment

The research project is supported by Ministry of Science and Higher Education grant NN519 384936.


# References


[1] K. Atanassov. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 20:87–96, 1986.

[2] K. Atanassov. *Intuitionistic fuzzy sets: Theory and Applications*. Springer-Verlag, 1999.

[3] H. Bustince, M. Pagola, P. Melo-Pinto, E. Barrenechea, and P. Couto. Use of Atanassov's intuitionistic fuzzy sets for modelling the uncertainty of the thresholds associated to an image. In *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, Intelligent Systems from Decision Making to DataMining, Web Intelligence and Computer Vision. Springer, Berlin, 2008.

[4] P. Couto, H. Bustince, V. Filipe, E. Barrenechea, M. Pagola, and P. Melo-Pinto. Image threshold using A-IFSs based on bounded histograms. In *Foundations of Fuzzy Logic and Soft Computing*, volume 4529/2007 of *Lecture Notes in Computer Science*, pages 96–103. Springer Berlin / Heidelberg, 2007.

[5] P. Couto, M. Pagola, H. Bustince, E.Barrenechea, and P. Melo-Pinto. Image segmentation using a-ifs. In *12th Conf. Information Processing and Management of Uncertainty in Knowledge-Based Systems, Malaga,Spain*, pages 1620–1627, 2008.

[6] S. Cox, R. Harvey, Y. Lan, J. Newman, and B. Theobald. The challenge of multispeaker lip-reading. In *International Conference on Auditory-Visual Speech Processing*, 2008.

[7] D. Dubois, S. Gottwald, P. Hajek, J. Kacprzyk, and H. Prade. Terminological difficulties in fuzzy set theory - the case of intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 156(485-491), 2005.

[8] K. Dyczkowski, P. Nowakowski, N. Kordek, and K. Stronski. Computational tools in the analysis of phonetic grammar. *Speech and Language Technology*, 11:145–155, 2008.

[9] H. McGurk and J. MacDonald. Hearing lips and seeing voices. *Nature*, 264:746–748, Dec. 1976.

[10] OpenCV. Open source computer vision official web site. `http://opencv.willowgarage.com/wiki/`.

[11] A. Pankowska and M. Wygralak. On hesitation degrees in if-set theory. In L. Rutkowski, J. Siekmann, R. Tadeusiewicz, and L. Zadeh, editors, *Artificial Intelligence and Soft Computing. LNAI*, volume 3070, pages 338–343. Springer, Heidelberg, 2004.

[12] G. Potamianos, C. Neti, G. Gravier, A. Garg, and A. Senior. Recent advances in the automatic recognition of audio-visual speech. In *Proceedings of the IEEE*, volume 92, pages 1306–1326, 2003.

[13] G. Potamianos, C. Neti, J. Luettin, and I. Matthews. *Audio-Visual Automatic Speech Recognition: An Overview*, chapter 10. MIT Press, 2004.

[14] E. Szmidt and J. Kacprzyk. Distances between intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, 114:505–518, 2000.

[15] E. Szmidt and J. Kacprzyk. A similarity measure for intuitionistic fuzzy sets and its application in supporting medical diagnostic reasoning. In *Artificial Intelligence and Soft Computing - ICAISC 2004*, volume 3070 of *LNCS*, pages 388–393. Springer-Verlag Berlin Heidelberg, 2004.

[16] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

The papers presented in this Volume 2 constitute a collection of contributions, both of a foundational and applied type, by both well-known experts and young researchers in various fields of broadly perceived intelligent systems.
It may be viewed as a result of fruitful discussions held during the Ninth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2010) organized in Warsaw on October 8, 2010 by the Systems Research Institute, Polish Academy of Sciences, in Warsaw, Poland, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences in Sofia, Bulgaria, and WIT - Warsaw School of Information Technology in Warsaw, Poland, and co-organized by: the Matej Bel University, Banska Bystrica, Slovakia, Universidad Publica de Navarra, Pamplona, Spain, Universidade de Tras-Os-Montes e Alto Douro, Vila Real, Portugal, and the University of Westminster, Harrow, UK:

Http://www.ibspan.waw.pl/ifs2010

The consecutive International Workshops on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGNs) have been meant to provide a forum for the presentation of new results and for scientific discussion on new developments in foundations and applications of intuitionistic fuzzy sets and generalized nets pioneered by Professor Krassimir T. Atanassov. Other topics related to broadly perceived representation and processing of uncertain and imprecise information and intelligent systems have also been included. The Ninth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2010) is a continuation of this undertaking, and provides many new ideas and results in the areas concerned.

We hope that a collection of main contributions presented at the Workshop, completed with many papers by leading experts who have not been able to participate, will provide a source of much needed information on recent trends in the topics considered.

ISBN-13 9788389475367
ISBN 838947536-7
9 788389 475367