Interpretable fuzzy system allowing to be framed in a profile photo through linguistic expressions

Patricia Conde-Clemente[1] Jose M. Alonso[1] Gracian Trivino[1]

[1]European Centre for Soft Computing
Gonzalo Gutiérrez Quirós s/n, 33600 Mieres, Asturias, Spain
Email: {patricia.conde,jose.alonso,gracian.trivino}@softcomputing.es

Abstract

This work explores the possibilities of building interpretable fuzzy control systems including human-in-the-loop by using natural language. For this purpose, fuzzy computational perceptions are embedded in a highly interpretable linguistic granular model. We present a very simple prototype that allows a person with visual disabilities to take their own profile photos.

Keywords: Interpretable Fuzzy Systems, Computational Theory of Perceptions, Linguistic Summarization of Data, Human-in-the-loop.

1. Introduction

The rapid evolution of the technology embedded in smartphones [1], combined with the increase in the use of these devices, broaden the scenarios in which it is possible to use systems that interact with humans [2, 3]. Although this interaction is not a simple task, fuzzy logic can facilitate knowledge extraction and representation in order to model complex problems [4, 5]. According to Michalski's "Comprehensibility Postulate": *the results of computer induction should be symbolic descriptions of given entities, semantically and structurally similar to those a human expert might produce observing the same entities* [6]. By using Highly Interpretable Linguistic Knowledge methodology (HILK) [7] we are able to represent the extracted knowledge in highly interpretable fuzzy rule-based systems. The human-centric character of interpretable fuzzy systems [8] is highly appreciated in many applications, especially in those involving high interaction with humans [9]. Today we can find interpretable fuzzy systems able to provide decision support [10] in medicine [11] or agriculture [12, 13].

The Computational Theory of Perceptions (CTP) [14, 15] is based on the widely known capability of fuzzy systems to model complex phenomena. CTP provides a framework to implement computational systems with the capacity of computing with the meaning of natural language expressions, i.e. with the capacity of computing with imprecise descriptions of the world in a similar way that humans

do [16, 17]. That aim is only feasible when the underlying fuzzy systems are highly interpretable. The effectiveness of CTP relies on human-centric interpretability of the designed models. There are CTP based systems that offer automatic linguistic reports of traffic evolution in roads [18] or in driving simulation environments [19]. The data generated by these systems can be graphics, tables, simple linguistic variables or linguistic descriptions of complex phenomena.

With the aim of using the new technologies integrated in smartphones, in this paper we introduce an interpretable fuzzy system in which the user is not a simply data summaries recipient, but he/she can interact directly with the system. The developed prototype is called `inProfilePhoto` because its goal is aiding a user with visual disability to get framed correctly when he/she desires to take his/her own profile photos.

`inProfilePhoto` provides a seemingly simple task that may be impossible otherwise. To accomplish this task, the system will interact in real-time with the user through linguistic commands related to the movements he/she has to perform in order to get framed. When the user hears the instruction, he/she is expected to react with the movement advised. Then, the system will analyze the last movement made by the user and will inform him/her about the effectiveness of the movement. Once the user is properly framed, the picture is taken automatically.

The rest of this paper is structured as follows. Section 2 describes the concept of granular linguistic model of phenomena (GLMP) and Section 3 describes how to enhance GLMP with HILK methodology. Then, the proposed methodology is used in the design of the prototype `inProfilePhoto` which is described in Section 4. Finally, Section 5 presents conclusions and sketches future work.

2. Granular linguistic model of phenomena

According to CTP, our perception of world is granular. A granule underlies the concept of a linguistic variable [20]. Zadeh defined linguistic variables as *"variables whose values are not numbers but words*

or sentences in a natural or artificial language" [21]. In this approach, a fuzzy linguistic label can be viewed as a linguistic summary of numerical data.

Our approach based on CTP, for developing computational systems able to generate linguistic descriptions of data, is called granular linguistic model of phenomena (GLMP) [19]. It consists of a network of perception mappings (PMs). Each PM receives a set of computational perceptions (CPs) and transmits upwards a CP. In the network, each CP covers specific aspects of the phenomenon with a certain degree of granularity. Using different aggregation functions and different linguistic expressions, the GLMP paradigm allows the designer to model computationally her/his perceptions. As an illustrative example, Fig. 1 shows the GLMP for the developed prototype. It will be further explained in Section 4.
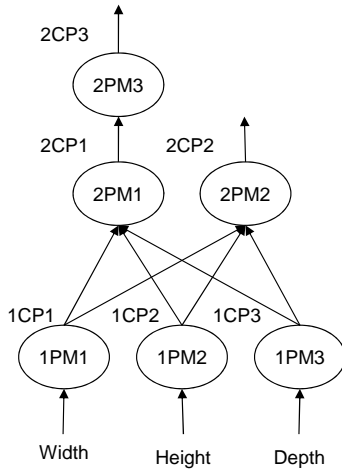

Figure 1: GLMP of `inProfilePhoto` prototype.

2.1. Computational perception (CP)

A CP is the computational model of a unit of information acquired by the designer about the phenomenon to be modeled. In general, CPs correspond with specific parts of the phenomenon at certain degrees of granularity. A CP is a couple (A,W) where:

- $A = (a_1, a_2, .., a_n)$ is a vector of linguistic expressions (words or sentences in natural language) that represents the whole linguistic domain in CP.

- $W = (w_1, w_2, .., w_n)$ is a vector of validity degrees $w_i \in [0,1]$ assigned to each a_i in the specific context.

2.2. Perception mapping (PM)

We use PMs to create and aggregate CPs. A PM is a tuple (U,y,g,T) where:

- $U = (u_1, u_2, .., u_n)$ is a vector of n input CPs u_i = (A_{ui}, W_{ui}). In the special case of first order

perception mappings (1PMs), these are the inputs to the GLMP and they are values $z \in R$ being provided either by a sensor or obtained from a database.

- $y = (A_y, W_y)$ is the output CP.

- g is an aggregation function employed to calculate $W_y = g(W_{u1}, W_{u2}, ..., W_{un})$ from the input CPs. In Fuzzy Logic, many different types of aggregation functions have been developed. For example, g might be implemented using a set of fuzzy rules. In the case of 1PMs, g is built using a set of membership functions.

- T is a text generation algorithm that allows generating the sentences in A_y. In simple cases, T is a linguistic template, e.g., "It has been taken a {bad, middle, good} framed photo", but it can be customized according to user preferences, mood, etc.

The following section explains how to guarantee the interpretability of a GLMP using the fuzzy modeling methodology HILK.

3. Highly interpretable GLMP

HILK is a fuzzy modeling methodology that was conceived for carefully integrating expert and induced knowledge under the fuzzy logic formalism, producing compact and robust classifiers easily comprehensible by human beings. It enables the user to follow a step-by-step procedure in the generation of all elements involved in a fuzzy knowledge base, starting from the design of fuzzy partitions, going through the rule-based learning and ending up with a knowledge base improvement stage which iteratively refines both partitions and rules [7].

For each step, HILK offers a number of alternative methods so that the user can select the best choice depending on each specific problem. In the prototype developed in this paper, only expert knowledge was considered. For ensuring interpretability of the GLMP we have carried out the following steps:

- We have defined the CPs as linguistic variables with small odd number of linguistic terms. Each linguistic variable is characterized by a Strong Fuzzy Partition (SFP) in its universe of discourse, as recommended by HILK. SFPs satisfy most constraints (coverage, distinguishability, overlapping, etc.) demanded to have interpretable partitions. Increasing the granularity of the underlying fuzzy partitions produces an increase in the number of linguistic expressions given to the user. This point is very important, because the GLMP should contain only the strictly necessary and sufficient information to describe the phenomenon.

- Once global semantics is defined, then we have defined CP linguistic rules of form "IF premise THEN conclusion". Both premise and conclusion are made up of linguistic propositions like "V is a_i" where one of the previously defined linguistic terms, is assigned to one of the selected variables. The absence of one variable in a rule means such variable is not considered in the evaluation of the selected rule.

4. inProfilePhoto: linguistic commands for taking framed profile photos

The architecture of the developed prototype is depicted in Fig. 2. It follows the human-in-the-loop approach which is popular in control system like [3, 22]. inProfilePhoto aids a person with visual disability to get framed correctly in order to obtain a good profile photo. To ensure that the user is positioned correctly, the system will indicate the movements he/she has to perform by voice messages.
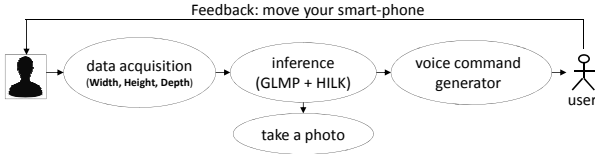

Figure 2: Architecture of inProfilePhoto prototype.

The application receives a profile photo in which a face is detected. From this we extract, the values of Width, Height and Depth which determine the user's position as depicted in Fig. 3. Then, relying on the developed GLMP (Fig. 1), we infer the most suitable motion the user should perform in order to get framed. Then, we generate the linguistic command that is conveyed to the user who is expected to move the smart-phone accordingly. Once the user is properly positioned, the photo is taken and another linguistic expression informs him/her that the photo was taken successfully.
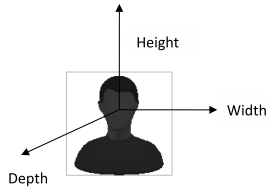

Figure 3: Variables which determine the user's position in a given photo.

4.1. GLMP for taking profile photos

The design of the developed GLMP is based on expert knowledge extracted and represented by HILK methodology. In the Section 4.3 you can find some additional implementation details about the prototype.

4.1.1. First order perceptions

As pointed out, inProfilePhoto uses three input variables (corresponding to the three dimensions Width, Height, and Depth) to take a profile photo in which the user is properly framed. We have defined the 1PM and 1CP for each of them.
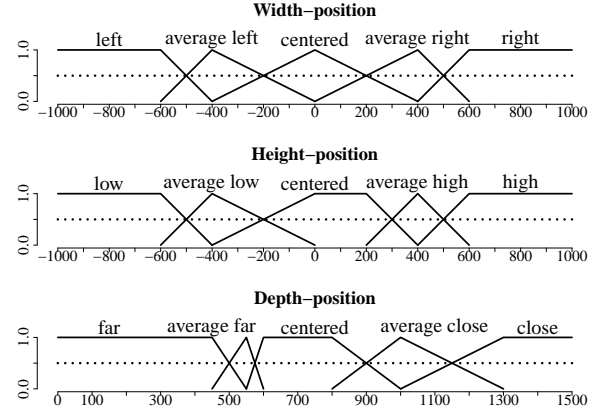

Figure 4: Strong fuzzy partitions for first order perceptions

Width-position (1PM1) is defined as (U_{1PM1}, y_{1PM1}, g_{1PM1}, T_{1PM1}) where:

U_{1PM1} is the X value of the center point of the user's face. It allows us to know the user's position in the width direction.

y_{1PM1} is the variable 1CP1 describing the Width position. It is expressed by the linguistic expressions A_{1CP1} = (left, average left, centered, average right, right) along with the corresponding validity degrees.

Height-position (1PM2) is defined as (U_{1PM2}, y_{1PM2}, g_{1PM2}, T_{1PM2}) where:

U_{1PM2} is the Y value of the central point of the user's face. It allows us to know the position regarding the height of the user in the photo.

y_{1PM2} is the variable 1CP2 describing the Height position. It is expressed by the linguistic expressions A_{1CP2} = (high, average high, centered, average low, low) along with the corresponding validity degrees.

Depth-position (1PM3) is defined as (U_{1PM3}, y_{1PM3}, g_{1PM3}, T_{1PM3}) where:

U_{1PM3} is the subtraction of the left and right points in the rectangle surrounding the user's face (see Section 4.3). This value is smaller when the user is afar the smartphone and is larger when the user is close.

y_{1PM3} is the variable 1CP3 describing the Depth position. It is expressed by the linguistic expressions A_{1CP3} = (far, average far, centered,

average close, close) along with the corresponding validity degrees.

The g_{1PMi} of these first order perceptions is constructed with respect to the SFPs of Width-position, Height-position and Depth-position (Fig.4). The membership function associated to each linguistic term provides the validity degree of the related linguistic expression for the given input value U_{1PMi}.

T_{1PMi} are simple templates. For instance T_{1PM3} = "The user's depth position in the photo is {far | average far | centered | average close | close}".

4.1.2. Second order perceptions

Move (2PM1) is defined as (U_{2PM1}, y_{2PM1}, g_{2PM1}, T_{2PM1}) where:

U_{2PM1} = (1CP1, 1CP2, 1CP3)

y_{2PM1} is the variable 2CP1 which represents the possible movements the user can perform. It is expressed by the linguistic expressions A_{2CP1} = (up, slightly up, slightly down, down, left, slightly left, slightly right, right, closer, slightly closer, slightly afar, afar, no move) along with the corresponding validity degrees.

g_{2PM1} is implemented using a set of 13 expert fuzzy rules like:

IF (Height-position is high) and (Width-position is left) and (Depth-position is far) THEN (Move is up)

T_{2PM1} = "Move your smartphone {to the left | slightly to the left | slightly to the right | to the right | up | slightly up | slightly down | down | closer | slightly closer | slightly afar | afar}".

Frame (2PM2) is defined as (U_{2PM2}, y_{2PM2}, g_{2PM2}, T_{2PM2}) where:

U_{2PM2} = (1CP1, 1CP2, 1CP3)

y_{2PM2} is the variable 2CP2. It summarizes the user's position in the photo regarding Width, Height, and Depth. It is expressed by the linguistic expressions A_{2CP2} = (bad, middle, good) along with the corresponding validity degrees.

g_{2PM2} is implemented using a set of 10 expert fuzzy rules like:

IF (Height-position is centered) and (Width-position is centered) and (Depth-position is centered) THEN (Frame is good)

T_{2PM2} = "It has been taken a {bad | middle | good} framed photo".

Motion Evaluation (2PM3) is defined as (U_{2PM3}, y_{2PM3}, g_{2PM3}, T_{2PM3}) where:

U_{2PM3} = (2CP1)

y_{2PM3} is the variable 2CP3. It makes a comparison between previous and current movements. The goal is to provide the user with feedback about the level of fulfillment of the previous command. It is expressed by the linguistic expressions A_{2CP3} = (negative, small negative, zero, small positive, positive) along with their corresponding validity degrees.

g_{2PM3} takes the previous ($Move_{pre}$) and current ($Move_{cur}$) movement (2CP1) as inputs of 48 expert fuzzy rules. Table 1 includes a subset of the defined fuzzy rules.

IF $Move_{pre}$ IS	AND $Move_{cur}$ IS	THEN Motion E. IS
up	up	zero
up	slightly up	small positive
up	slightly down	small negative
up	down	negative
slightly up	up	positive
slightly up	slightly up	zero
slightly up	slightly down	small negative
slightly up	down	negative
slightly down	up	negative
slightly down	slightly up	small negative
slightly down	slightly down	zero
slightly down	down	positive
down	up	negative
down	slightly up	small negative
down	slightly down	small positive
down	down	zero

Table 1: A set of fuzzy rules which compares previous and current movements.

T_{2PM3} is produced by the following algorithm:

If the linguistic expression in 2CP3 with the greatest validity degree is *negative*, then the sentence is "You are Away".

Else if it is *small negative*, then the sentence is "You are near the target point".

Else if it is *zero*, then the sentence is "You are in the same position".

Else if it is *small positive*, then the sentence is "You have passed the target point a bit".

Otherwise (it is *positive*), then the sentence is "You have clearly over-passed the target point".

Notice that each g_{2PMi} is implemented by a set of fuzzy rules with the usual Min-Max fuzzy reasoning mechanism.

4.2. Generating voice commands

Using the GLMP defined in the previous section, we can generate a set of valid sentences describing the phenomenon "take a profile photo" in different levels of detail. The interaction with the user is made through voice commands (customizable through the user's language). The generation of such messages is made in two steps:

1. Choose the linguistic expression in 2CP1 with the greatest validity degree. Outgoing messages are prioritized by 1) Height-position, 2) Width-position, and 3) Depth-position.
2. If the selected linguistic expression indicates there is no movement (*no move*), then we use the template T_{2PM2}.

   Else if there is no previous movement i.e., it is the first command, then we use the template T_{2PM1}.

   Else if the previous movement was related to different dimensions (Fig. 3), then the template for the command is {"Well done. Now" $\cup$ T_{2PM1} }.

   Otherwise, the template for the command is {T_{2PM3} $\cup$ "Now, " $\cup$ T_{2PM1}}.
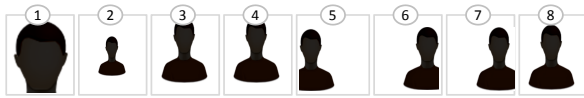

Figure 5: Simulation of `inProfilePhoto` prototype.

As an illustrative study case the Fig. 5 shows the simulation of a profile photo capture, from an initial state to the final state, in which the photo is taken. The prototype developed give the following orders for each capture:

Step 1: Move your smartphone afar.

Step 2: You have passed the target point a bit. Now, move your smartphone slightly closer.

Step 3: Well done. Now, move your smartphone slightly up.

Step 4: You are in the same position. Now, move your smartphone slightly up.

Step 5: Well done. Now, move your smartphone to the right.

Step 6: You have clearly over-passed the target point. Now, move your smartphone to the left.

Step 7: You are near the target point. Now, move your smartphone slightly to the left.

Step 8: It has been taken a good framed photo.

4.3. Implementation details

GLMP has been developed with the help of two free software tools, GUAJE[1] [23] and Xfuzzy[2] [24].

- GUAJE implements the HILK methodology. We have used this tool to design the involved partitions and rules. Then, the designed fuzzy systems is exported to Xfuzzy.

---

[1] http://www.softcomputing.es/guaje
[2] https://forja.rediris.es/projects/xfuzzy/

- Xfuzzy integrates a set of tools for designing fuzzy logic-based inference systems. One of these tools allows translating designed system into different programming languages. We have used this tool to translate our fuzzy system to Java code, easy to integrate with the rest of the control architecture.

The prototype runs on a smartphone equipped with front camera and android software for detecting person's faces and for converting the generated text messages into voice messages that can be delivered to the user through the speakers of the smartphone. When the camera software detects a face, it draws a rectangle around it. The capture image has a range (-1000, 1000), where the center is the point (0, 0).

5. Conclusions

This paper presents `inProfilePhoto` a prototype capable of interacting with a human to perform a seemingly simple task but really hard for people with visual disabilities, such as taking a profile photo correctly framed. With this aim we have merged two approaches related to two different fields. Namely, the HILK methodology regarding Interpretability of fuzzy systems and the GLMP approach with respect to Computational theory of perceptions.

The prototype must be still further developed up to be converted in a commercial application. To do so, the next step is testing the prototype with real users. As a result we can gather real data to refine partitions and rules using the automatic learning methods provided by HILK. Anyway, after this initial exploration, the preliminary results are very promising. The interpretability-guided design of GLMP can be very useful in the development of systems that interact with humans in real-time, where interpretability is a main concern. The presented new technology could be successfully applied to develop more complex control systems including human-in-the-loop. For instance, we are thinking about a system able to take group photos.

References

[1] N. Palmer, R. Kemp, T. Kielmann, and H. Bal. The case for smartphones as an urgent computing client platform. *Procedia Computer Science*, 9:1667–1676, 2012.

[2] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Fox, H. Kautz, and D. Hahnel. Inferring activities from interactions with objects. *IEEE Pervasive Computing*, 3(4):50–57, 2004.

[3] R. Stanciu and P. Oh. *Human-in-the-Loop Control for a Broadcast Camera System.* InTech, 2010.

[4] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, 3(1):28–44, 1973.

[5] F. Barrientos and G. Sainz. Interpretable knowledge extraction from emergency call data based on fuzzy unsupervised decision tree. *Knowledge-Based Systems*, 25(1):77–87, 2012.

[6] R. S. Michalski. A theory and methodology of inductive learning. *Artificial intelligence*, 20(2):111–161, 1983.

[7] J. M. Alonso and L. Magdalena. HILK++: an interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 15(10):1959–1980, 2011.

[8] W. Pedrycz. Human centricity in computing with fuzzy sets: an interpretability quest for higher order granular constructs. *Journal of Ambient Intelligence and Humanized Computing*, 1(1):65–74, 2010.

[9] J. M. Alonso and L. Magdalena. Special issue on interpretable fuzzy systems. *Information Sciences*, 181(20):4331–4339, 2011.

[10] J. M. Alonso, C. Castiello, M. Lucarelli, and C. Mencar. Modelling interpretable fuzzy rule-based classifiers for Medical Decision Support. In R. Magdalena-Benedito, E. Soria-Olivas, J. Guerrero-Martínez, J. Gómez-Sanchis, and A. J. Serrano-López, editors, *Medical Applications of Intelligent Data Analysis: Research advancements*, pages 255 – 272. IGI Global, 2012.

[11] I. Gadaras and L. Mikhailov. An interpretable fuzzy rule-based classification methodology for medical diagnosis. *Artificial Intelligence in Medicine*, 47(1):25–41, 2009.

[12] S. Guillaume and B. Charnomordic. Learning interpretable fuzzy inference systems with FisPro. *Information Sciences*, 181(20):4409–4427, 2011.

[13] S. Guillaume and B. Charnomordic. Fuzzy inference systems: An integrated modeling environment for collaboration between expert knowledge and data using FisPro. *Expert Systems with Applications*, 39(10):8744–8755, 2012.

[14] L. A. Zadeh. From computing with numbers to computing with words - from manipulation of measurements to manipulation of perceptions. *IEEE Transactions on Circuits and Systems*, 45(1):105–119, 1999.

[15] L. A. Zadeh. Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. *Journal of statistical planning and inference*, 105(1):233–264, 2002.

[16] R. R. Yager. Approximate reasoning as a basis for computing with words. *Computing with words in information/intelligent systems*, 1:50–77, 1999.

[17] L. A. Zadeh. Computing with words (CW) and its application to decision support and systems analysis. In *IEEE International Symposium on Intelligent Signal Processing*, pages 1–2, 2003.

[18] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Trivi no, Á. Sánchez, and P. D. Suárez. Automatic linguistic report of traffic evolution in roads. *Expert Systems with Applications*, 39(12):11293–11302, 2012.

[19] L. Eciolaza, M. Pereira-Fariña, and G. Trivino. Automatic linguistic reporting in driving simulation environments. *Applied Soft Computing*, doi:10.1016/j.asoc.2012.09.007, 2012.

[20] L. A. Zadeh. Toward human level machine intelligence - is it achievable? the need for a paradigm shift. *IEEE Computational Intelligence Magazine*, 3(3):11–22, 2008.

[21] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning-I. *Information Sciences*, 8(3):199–249, 1975.

[22] K. M. Passino and S. Yurkovich. *Fuzzy Control.* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1997.

[23] J. M. Alonso and L. Magdalena. Generating understandable and accurate fuzzy rule-based systems in a Java environment. In *Lecture Notes in Artificial Intelligence - 9th International Workshop on Fuzzy Logic and Applications*, volume LNAI6857, pages 212–219. Springer, 2011.

[24] I. Baturone, F. J. Moreno-Velo, S. Sánchez-Solano, Á. Barriga, P. Brox, A. A. Gersnoviez, and M. Brox. Using Xfuzzy environment for the whole design of fuzzy systems. In *IEEE International Conference on Fuzzy Systems*, pages 1–6, 2007.