Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 253 (2014) 101–121

FUZZY sets and systems

www.elsevier.com/locate/fss

Computational Perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon

Daniel Sanchez-Valdes *, Gracian Trivino

European Centre for Soft Computing, C/Gonzalo Gutierrez Quirós, s/n, 33600 Mieres, Asturias, Spain

Received 10 June 2013; received in revised form 22 October 2013; accepted 17 December 2013

Available online 24 December 2013

Abstract

As a contribution to the field of the Zadeh's Computational Theory of Perceptions, this paper deals with generating linguistic models of those phenomena that the designer cannot completely recognize at running time. We explore the possibility of generating linguistic descriptions including in their meaning the fact the complete interpretation of input data is not feasible. With this purpose, we extend our previous research in this field by introducing the concept of Computational Perception of uninterpretable data. This extension provides more expressiveness and flexibility to the currently available resources. We define a linguistic variable that explains the degree with which phenomena suit the model, i.e., if input signals can be interpreted. In order to demonstrate the advantages of this contribution, we apply it to create a definition of fuzzy set of quasi-periodic phenomena. We use these new linguistic models to analyze and linguistically describe some characteristics of the human gait, as a case study of quasi-periodic phenomenon.

Keywords: Uninterpretable data; Computing with perceptions; Linguistic modeling; Quasi-periodic signals; Gait analysis

1. Introduction

Human beings use Natural Language (NL) to organize the structure of their experience and to communicate with others [1]. The experience acquired in our everyday life environment using NL allows us to describe our perceptions to others by means of meaningful linguistic expressions that we share with them in specific contexts.

Many times, due to the limits of our personal experience, we do not have available all the resources needed to recognize and describe the relevant details of specific input data. In order to communicate the uncertainty produced by this lack of experience, we use expressions like "as far as I know...", remarking to the listener the lack of a complete knowledge about the meaning of input data. For example, after exploring a damaged car, the expert could tell us: "It seems to be a problem with the battery or the starter engine, but it may be other cause". Obviously, we

* Corresponding author. Tel.: +34 985 45 65 45.
E-mail addresses: daniel.sanchezv@softcomputing.es (D. Sanchez-Valdes), gracian.trivino@softcomputing.es (G. Trivino).

0165-0114/$ – see front matter
http://dx.doi.org/10.1016/j.fss.2013.12.010

understand that he/she is not very sure about the diagnosis. His/her statement indicates that the expert's experience is not enough to clearly recognize the cause of the car bad functioning.

There is a demand of computational applications able to convert data into relevant linguistic descriptions. According to each situation type, these tools should interpret and represent data in an understandable way, and therefore, generate descriptions so useful as possible to achieve the specific user's goals. Zadeh's principle of incompatibility [2] formulates that the more complex the phenomena, the more difficult to design precise computational models. In order to cope with this drawback, human beings use NL to build linguistic models of perceived phenomena in their environment.

In general, because the inevitable incompleteness of computational models, when monitoring complex phenomena, we only can recognize input data with a certain matching degree. This paper faces with the challenge of enriching the meaning of automatic linguistic descriptions of phenomena in order to contribute to overcome this limitation.

In recent years, there has been a growing interest for modeling approaches based on Fuzzy Logic due to its ability for linguistic concept modeling and system identification. On the one hand, semantic expressiveness, using linguistic variables and rules, is quite close to natural language, which reduces the effort of expert knowledge extraction. On the other hand, being universal approximators [3] fuzzy inference systems are able to perform nonlinear mappings between inputs and outputs. Thanks to these advantages, Fuzzy Logic has been successfully applied in classification [4,5], regression [6,7], control [8,9], system modeling [10,11] and linguistic summarization [12,13], achieving a good interpretability accuracy trade-off [14].

In this context, the idea of *uninterpretable data* arises naturally. We define *uninterpretable data* as input data that do not fully suit with the available model of the monitored phenomenon. The lack of complete knowledge is a typical characteristic of expert systems and data-driven systems, including rule base systems that combine expert rules with induced ones [15,16]. According to [17], "completeness means that for any possible input vector, at least one rule is fired, there is no inference breaking". The lack of complete knowledge or completeness varies depending on the applications. In automatic applications like fuzzy controllers, the rule base should be complete. However, in applications that involve interaction with humans, the interpretability is more appreciated than the completeness. In this case, the higher the number of rules, the lower the system interpretability. In classical fuzzy modeling, two options mainly exist to deal with this problem: to complete the rule base or to use a default rule [18]. Mencar and Fanelli define completeness as "a property of deductive systems that has been used in the context of Artificial Intelligence to indicate that the knowledge representation scheme can represent every entity within the intended domain" [19].

Quasi-periodic phenomena are a type of complex phenomena that provide signals with repetitive temporal patterns but including some variations in period and amplitude. Examples of this type of phenomena are electrocardiograms, accelerations produced during the human gait, vibrations of musical instruments, etc. They are good examples of phenomena that either cannot be modeled, or we do not want to model, with absolute precision. Popular approaches to deal with quasi-periodic signals vary from wavelets transform [20], classical curve fitting methodologies to Hidden Markov Models [21] and Neural Networks [22,23].

Our long term research line is based on the Computational Theory of Perceptions (CTP) introduced by Zadeh [24–26]. In previous works, we have developed computational systems able to generate linguistic descriptions of different types of phenomena. For example, we have generated assessing reports in truck driving simulators [27], reports about traffic evolution in roads [28], about the relevant features of the Mars' surface [29] and linguistic descriptions about visual double stars [30].

When Fuzzy Logic is combined with Finite State Machines, a more fluent modeling process is achieved. Fuzzy Finite State Machines (FFSM) are specially useful tools to model dynamical processes that are time-dependent, becoming an extension of classical Finite State Machines [31,32]. The main advantage of FFSMs is their ability to handle imprecise and uncertain data, which are inherent to real-world phenomena, in the form of fuzzy states and transitions. The theoretical basics of FFSMs were established in [33] and later developed in [34–36]. In previous research, we have learned that FFSMs are suitable tools to model quasi-periodic phenomena. We have applied FFSMs for pattern recognition tasks such as human gait recognition [37,38] and gesture recognition [39,40] based on accelerometer data. We used an FFSM to fuse body posture and WiFi positioning [41].

The context of our contribution can be defined using the following quote from Klir and Yuan [42]: "uncertainty in any problem-solving situation is a result of some information deficiency. Information (pertaining to the model within which the situation is conceptualized) may be incomplete, imprecise, fragmentary, no fully reliable, vague, contradictory, or deficient in some other way. In general, various information deficiencies may result in different