Signal Processing 141 (2017) 1–15

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

A two-phase method of forming a granular representation of signals

Shuai Liu[a,*], Witold Pedrycz[b,c,d], Adam Gacek[e], Yaping Dai[a]

[a] School of Automation, Beijing Institute of Technology, Beijing 100081, P. R. China
[b] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, T6R 2V4 AB, Canada
[c] Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, 21589, Saudi Arabia
[d] Systems Research Institute, Polish Academy of Sciences Warsaw, Poland
[e] Institute of Medical Technology and Equipment (ITAM) 118, Roosevelt Street, Zabrze 41-800, Poland

ARTICLE INFO

Article history:
Received 4 January 2017
Revised 12 April 2017
Accepted 25 May 2017
Available online 26 May 2017

Keywords:
Granular Computing
Principle of justifiable granularity
Coverage and specificity
Granular representation of signals

ABSTRACT

This study focuses on a granular representation of signals. The development process dwells upon the use of the principle of justifiable granularity encountered in Granular Computing and the least square error method. This process consists of two phases where the construction of granular representatives of a family of signals (temporal data) is realized by invoking the design at the local and global level. At the local design level involving individual elements of the universe of discourse (time moments), the principle of justifiable granularity is applied to construct (a vertical part) information granules. At the global level, the least square error method is invoked to develop the bounds (envelopes) of the information granules already formed at the local level. Experimental studies are reported for the granular representation of synthetic data and publicly available ECG signals. Furthermore we demonstrate that the proposed approach can be used to construct fuzzy sets of type-2.

1. Introduction

For the signal processing, we are faced with several fundamental types of issues such as classification [48,53,54,67], prediction [50–52], compression and representation [45–47,66]. Owing to the uncertainties and the nonlinearities of the signals as well as the diversity of signals belonging to the same class, it becomes difficult or impractical to establish the exact mathematical models of signals. In order to cope with these issues, Computational Intelligence has been investigated as a viable conceptual and algorithmic framework. For example, some research is carried out with a focus on the technology of fuzzy sets with applications reported in forecasting time series, identification of non-linear dynamic systems and so on [35–39,62,64], as it could capture uncertainty and diversity of signals and ensuing systems, for instance, uncertainty of the rules and the noisy data. Signal processing is also carried out with the aid of neural networks, evolutionary algorithms (such as particle swarm optimization, differential evolution and others) and Granular Computing (owing to the development and usage of granular models) for signal representation, compression, classification and prediction [40–44,49,63,65]. The importance and usefulness of Granular Computing stem from the fact that information granules involved in the modeling temporal data help capture and quantify a diversity of signals and build their abstract representation.

As an emerging paradigm of information processing, Granular Computing plays a pivotal role in the construction of human-centric models [5,7,14,15,59–61]. Let us recall that Granular Computing is about acquiring, processing, interpreting and communicating information granules [3,4]. Information granules are formalized as sets, fuzzy sets, rough sets and the like [2]. Note that the central thought behind information granules is a notion of abstraction: instead of considering individual elements exhibiting some closeness or resemblance, we arrange them together by forming a single information granule. Information granules are abstract entities which arise through the process of formation of an abstract view at a certain real-world phenomenon or system. While designing granular models, the construction of information granules is of paramount relevance and constitutes a challenging problem. This problem has been intensively studied [1,9,29–34,55–57], however a number of issues are still open. A way of forming granular representatives of a collection of data or signals is discussed in [9] and [29]. In [9], a granular representative of a collection of signals is obtained by using fuzzy clustering, more specifically fuzzy C-means (FCM) [16]. While in [29], the method just concentrate on the construction of interval granules.

The key objective of this study is to establish a two-phase development process of a granular representation of signals. In the design, we engage the principle of justifiable granularity (when constructing individual information granules for the corresponding time moments; such granules are locally formed information granules) and the Least Square Error (LSE) approximation (leading

* Corresponding author.
E-mail address: liushuai870327@163.com (S. Liu).

http://dx.doi.org/10.1016/j.sigpro.2017.05.026
0165-1684/

to the refinement and construction of an information granule of a global nature). There are three evident advantages for the proposed method: at every time moment, the local information granule is reflective of the locally available data and in the design we capture the essential properties of the data expressed in terms of the measures of coverage and specificity. The formation of the global information granule is completed with the least square error method, in which way the properties of locally formed information granules are retained and reconciled. In comparison to the methods introduced in [9] and [29], the proposed method could be applied to construct both interval and fuzzy granules and is not impacted by a choice of numeric values of some parameters of the clustering method.

In what follows, we assume that we are concerned with space $\boldsymbol{X} = \{x_1, x_2, \ldots, x_N\}$, $\boldsymbol{X} \subseteq \boldsymbol{R}$ and we are also concerned with n signals located in some real-valued space $\boldsymbol{Y} \subseteq \boldsymbol{R}$ for each x_i ($x_i \in \boldsymbol{X}$) being provided in the following format:

$x_1 \quad y_{11}, y_{12}, \ldots, y_{1n}$
$x_2 \quad y_{21}, y_{22}, \ldots, y_{2n}$

$x_N \quad y_{N1}, y_{N2}, \ldots, y_{Nn}$

In numerous real-world problems, we encounter signals belonging to a certain class. The intent is to find a representative of such signals so that this representative captures the essence of the data to the highest extent. We advocate that when a collection of numeric signals is characterized, the description (prototype) has to account for the inherent diversity existing within the family of signals and be made at the higher level of abstraction in the form of an information granule rather than a numeric signal (for example, an average of all signals belonging to the same class).

The rest of this paper is structured as follows: In Section 2, the underlying concepts are briefly reviewed; the formulation of the problem along with its detailed two-phase solution is presented in Section 3; experimental studies are reported in Section 4; in Section 5, the conclusions are drawn.

2. Preliminaries

This section serves as a brief introduction to the concepts used throughout the paper. In particular, we bring forward the principle of justifiable granularity [8,17].

The principle of justifiable granularity, originally introduced in [8], serves as a fundamental design vehicle to form information granules [13] realized in the presence of experimental evidence (data). It constitutes a generic method for the designing of information granules in Granular Computing [23–27]. The principle exhibits a significant deal of generality. First, it applies to a variety of formal settings of Granular Computing. Second, it applies to various formats of experimental evidence that could be composed of information granules coming as intervals, fuzzy sets, probability density functions and others not being exclusively confined to numeric data.

In its generic version, the construction of information granule is carried out in the presence of numeric one-dimensional data. It translates into a certain optimization problem, in which we maximize the following intuitively appealing criteria: (i) high data coverage and (ii) high specificity of the constructed information granule. Concisely speaking, coverage, cov(.), expresses an extent to which the constructed information granule “covers” (includes) the available experimental data. The higher the coverage, the more legitimate (justifiable) the formed information granule tends to become. The specificity, sp(.), as the name stipulates, quantifies how detailed the information granule is. Intuitively, the highest specificity is achieved for an information granule consisting of a single element, $\mathrm{sp}(\{z\}) = 1$. The more elements contribute to the information granule, the lower its specificity tends to become. It is easy to note that the increase in the coverage yields lower specificity values and vice versa. As both of the requirements have to be satisfied, one can consider a product of coverage and specificity and regard it as a suitable performance index whose value have to be maximized with respect to the parameters of the information granule. In case of interval information granule, these parameters are the bounds of the interval. In situation when we focus on fuzzy sets, those are the parameters of fuzzy sets (for instance, in case of a triangular fuzzy set, the parameters are modal value and lower and upper bounds).

3. The two-phase design of granular representation of signals

The development of the granular representation of signals is completed in two phases. The local phase is concerned with the formation of an information granule for individual elements of the universe (viz. discrete time moments). The global phase focuses on the development of envelopes (bounds) completed over all information granules already produced at the local basis.

3.1. Formulation for the two-phase optimization method

As noted, the development process consists of the two essential construction phases:

(i) the determination of the parameters of each vertical information granule. For each x_i in $\boldsymbol{X}$, we construct the corresponding information granule A_i by invoking the principle of justifiable granularity. Because of the focus of this construct, we refer to A_i as the vertical information granule. It is referred to as the local granule also, as the vertical information granule constructed for each x_i is part of the final result. The analytical form of A_i is predetermined and its parameters are estimated. Obviously, the parameters depend on the type of the considered information granule. For instance, in case of an interval, we have a triple of numbers (a_i, m_i, b_i) (the detailed algorithm about how to obtain these parameters is given in Section 3.2) describing the lower bound, the central point and the upper bound of the interval. If a fuzzy set, which is described by a triangular, parabolic or square root membership function is considered, the individual coordinates concern the lower bound, modal value, and the upper bound of the fuzzy set.

(ii) the determination of the envelope functions (envelopes) of the vertical granules formed over $\boldsymbol{X}$. Having (a_i, m_i, b_i) constructed at the local phase, the parameters of the envelopes are determined through curve fitting. This optimization problem falls under the rubric of the least square error approximation. More specifically, given the parametric form of the fitting functions, their parameters are estimated. With this regard, we form three functions pertaining to the bounds and modal values of the vertical information granules. For the data (x_i, a_i), ($i = 1, 2, \ldots, N$), one constructs f^-. The formation of f is carried out on a basis of (x_i, m_i) whereas $f^+$is constructed on a basis of (x_i, b_i). Furthermore the obvious inequality $f^- \le f \le f^+$ implies a constraint-based optimization, which has to be taken into consideration in the overall development process. It is noticeable that now the parameters of the individual fitting functions are considered together in the approximation of the envelopes of the information granules. This step realizes the global phase and the final result produces an information granule of a global character. The essence of this two-phase design is depicted in Fig. 1.