



ELSEVIER

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrecFuzzy clustering with semantically distinct families of variables:
Descriptive and predictive aspectsWitold Pedrycz^{a,b,*}, Andrzej Bargiela^c^aDepartment of Electrical and Computer Engineering, University of Alberta, Edmonton AB, Canada T6R 2G7^bSystem Research Institute, Polish Academy of Sciences, Warsaw, Poland^cSchool of Computer Science, The University of Nottingham, Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor Darul Ehsan, Malaysia

ARTICLE INFO

Article history:

Received 29 September 2008

Available online 11 July 2010

Communicated by F. Roli

Keywords:

Fuzzy clustering

Descriptive features

Functional features

Fuzzy C-means (FCM)

Reconstruction criterion

Granulation-degranulation

ABSTRACT

Fuzzy clustering being focused on the discovery of structure in multivariable data is of *relational* nature in the sense of not distinguishing between the natures of the individual variables (features) encountered in the problem. In this study, we revisit the generic approach to clustering by studying situations in which there are families of features of descriptive and functional nature whose semantics needs to be incorporated into the clustering algorithm. While the structure is determined on the basis of all features taken en-block, it is anticipated that the topology revealed in this manner would aid the effectiveness of determining values of functional features given the vector of the corresponding descriptive features. We propose an augmented distance in which the families of descriptive and predictive features are distinguished through some weighted version of the distance between patterns. The optimization of this distance is guided by a reconstruction criterion, which helps minimize the reconstruction error between the original vector of functional features and their reconstruction realized by means of descriptive features. Experimental results are offered to demonstrate the performance of the clustering and quantify the effect of reaching balance between semantically distinct families of features.

© 2010 Elsevier B.V. All rights reserved.

1. Introductory notes

It is evident that the concepts of closeness, similarity, distance and alike play a pivotal role in clustering as well as fuzzy clustering. Once some distance has been adopted, the structure revealed through clustering could exhibit a great deal of variation and might show different levels of relevance. Different features may also contribute at different levels to the formation of clusters. The relevance of the features in the clustering process can be addressed in several different ways. In spite of the existing variety of algorithmic underpinnings, there are several general categories of mechanisms involved. Quite often the relevance of the feature is associated with the mechanism of linear or nonlinear scaling. In other words, given the vector of features encountered in the clustering problem, say $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, each of the features enters into the calculations of the distance with its own scaling factor σ_i which is typically taken as a standard deviation of the corresponding feature. Then the Euclidean distance between two vectors \mathbf{x} and \mathbf{w} is computed in a straightforward manner as follows:

$$\|\mathbf{x} - \mathbf{w}\|^2 = \sum_{j=1}^n \frac{(x_j - w_j)^2}{\sigma_j^2}. \quad (1)$$

The weights (scaling factors) σ_i are determined by taking into account the underlying statistical properties of data. Conceptually, the weights being used in (1) allow us to carry out meaningful computing especially if the ranges of the original variables vary quite substantially. Some further augmentation of (1) may involve the use of additional weights f_i whose values are reflective of the relevance of the corresponding features as perceived from the perspective of the problem itself. The character of such weights is different from the nature of σ_i . Subsequently the origin of the numeric values of the weights is different from those capturing the statistical properties of the data (variables). In this sense, these weights need to accommodate domain knowledge quantified by experts. For instance, a pairwise comparison method (Saaty, 2000) or other elicitation techniques can be considered here. By considering these weights, the resulting distance can be computed in the following manner:

$$\|\mathbf{x} - \mathbf{w}\|^2 = \sum_{j=1}^n f_j \frac{(x_j - w_j)^2}{\sigma_j^2}. \quad (2)$$

The relationships between the individual features can be captured by considering a generalized distance of the form

* Corresponding author at: Department of Electrical and Computer Engineering, University of Alberta, Edmonton AB, Canada T6R 2G7. Tel.: +1 780 492 4661; fax: +1 780 492 1811.

E-mail address: pedrycz@ee.ualberta.ca (W. Pedrycz).

$(\mathbf{x} - \mathbf{w})^T A(\mathbf{x} - \mathbf{w})$ with A being a positive definite matrix. In case of the covariance matrix, $A = \Sigma$, we arrive at the Mahalanobis distance.

There are some practical situations in which the problem (data) comes quite naturally with blocks (ensembles) of variables (features) and therefore it becomes beneficial to accommodate such structural domain knowledge into the calculations of the distance function. This scenario gives rise to clustering with families of semantically distinct variables: patterns are described by features located in spaces which are *semantically* distinct. Here we offer a few representative examples in which a semantic grouping of variables comes into play.

1.1. Clustering image data

Each pixel comes with its geometric coordinates and the properties of brightness, color, texture, etc. As such, the object is described by features with two vector components such as the geometry (coordinates) and the image properties of the pixel. When dealing with 3D images, the vector pertaining to the geometry of the object (voxel) consists of three coordinates identifying a position of the voxel.

1.2. Clustering experimental data used in system modeling

The data consist of input–output pairs (\mathbf{x}, \mathbf{y}) with \mathbf{x} being a vector of inputs and \mathbf{y} forming the corresponding outputs of the system. Quite commonly we consider the output to be a single variable and in this case we allude to multiple input–single output system modeling.

1.3. Clustering of data with composite indicators

There could be situations in which the data are described by a large number of features along with a few synthetic indicators.

Depending upon the category of the problem at hand, we may encounter situations where the input space, in which the clustering of data takes place, exhibits regular topology, see Fig. 1(a) (for instance, when we are concerned with pixels or voxels) or a fairly irregular distribution of the objects as illustrated in Fig. 1(b) which is typical for a geo-referenced data. For instance, one may consider coordinates of some drilling sites while the objective of the prediction task is to estimate the effectiveness of drilling positioned at some specified geographic location.

All these problems exhibit some interesting similarity. In their formulation we may easily distinguish between *descriptive* and *functional* variables. The first group of variables is concerned with the description of the data themselves (say, location), while the second group is used to characterize the functional nature of the data. Altogether the variables can be referred to as semantic blocks (families) of features. Quite commonly we predict values of functional variables on the basis of the available functional component.

The choice of the conceptual blocks of features (descriptive or functional) is implied directly by the problem at hand. For instance, if we are concerned with modeling where we encounter inputs and outputs (and this distinction is apparent from the form of the data), then such inputs and outputs form the two blocks of the features. The section devoted to numeric experiments offers a number of insights and the main points regarding the formation of the sets of features have been elaborated in detail.

The objective of this study is to investigate the concept of fuzzy clustering in the context of data characterized by the semantic blocks of features. The crux of the proposed augmentation associates with a suitable adjustable distance function, which helps to distinguish between descriptive and functional features (variables) and quantify their role in the clustering process. To make this quantification possible, we introduce a *reconstructability* criterion whose minimization involves the adjustment of the parametric format of the augmented distance function. This criterion is applied to functional features. While there have been a great deal of problems in which semantic blocks of features are encountered, especially in the realm of image processing and clustering (Berget et al., 2008; Biosca et al., 2008; Cai et al., 2007; Carballido-Gamio et al., 2006; Cinque et al., 2004; Kang et al., 2009; Ozkan et al., 2008) the conceptual distinction between descriptive and functional features has not been addressed in the realm of clustering. In its very virtue, fuzzy clustering and clustering are *direction-free* constructs. There are neither conceptual nor algorithmic provisions to incorporate the aspect of directionality or any discriminative treatment of features. Numerous studies in image processing concerning techniques of fuzzy clustering (Chuang et al., 2006; Kannan, 2008; Maitra and Chatterjee, 2008; Peñaranda et al., 2006; Silva et al., 2008; Xia et al., 2007; Verikas et al., 2005) do not invoke any aspect of functional properties of pixels or blocks of pixels. The idea of directional clustering introduced by Hirota and Pedrycz (1996) can be viewed as a certain attempt to distinguish between input and output variables and in this manner the method brings a notion of *directionality* into the formulation of the clustering problem. The notion of conditional (context-oriented) clustering (Pedrycz, 1996, 1998) introduces a component of directionality through predefined information granules formulated in the output space, which are used to guide a search for structure completed in some multivariable input space.

The originality of the study realized here stems from the fact that in spite of the nature of the problems being tackled, the underlying clustering has never been spelled out and formulated formally so that we could take advantage of the quantification of the quality of clustering and endow the method with some parametric flexibility.

While there is a remarkably diversified suite of clustering and fuzzy clustering methods available in the literature, the issue of dealing with some subsets of features (attributes) which exhibit a clearly distinct semantics, has not been dealt with. Furthermore

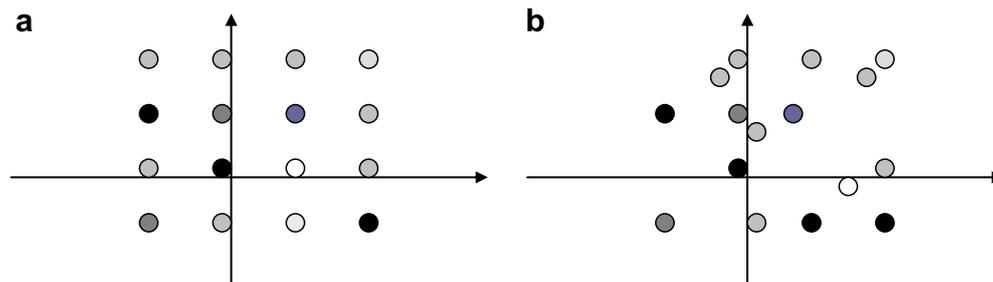


Fig. 1. Two categories of clustering problems of functional features with (a) regular distribution of data in the input space, and (b) irregular distribution in the input space. The level of brightness of individual points is reflective of the functionality in the input space.

the proposed concept of semantic blocks of features is augmented by a well-defined algorithmic framework and endowed with a constructive way of forming a sound balance between the role of descriptive and functional features. The reconstruction criterion plays a pivotal role with this regard.

The paper starts with a brief proposal on the formation of the distance function, which takes into consideration the semantic blocks of variables.

2. Distance function for semantic blocks of features

More formally, the patterns to be clustered are treated as $(n + m)$ -dimensional vectors described in terms of two blocks of variables (features) coming in the following form:

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}. \quad (3)$$

Alluding to image data, the first variables (described by \mathbf{x}) denote a location of a pixel (two-dimensional case, $n = 2$) or a voxel (three dimensional case, $m = 3$). The second part of the vector (\mathbf{y}) captures the functional properties of the given pixel (voxel). To effectively distinguish between the two blocks of features in (1) and emphasize their different character, we introduce the following distance between any two patterns \mathbf{z}_1 and \mathbf{z}_2 :

$$\|\mathbf{z}_1 - \mathbf{z}_2, \alpha\| = \|\mathbf{x}_1 - \mathbf{x}_2\| + \alpha\|\mathbf{y}_1 - \mathbf{y}_2\|. \quad (4)$$

The positive weight (scaling coefficient) α is used to establish a sound balance between the distance between the spatial location of the pixels (voxels) and their properties. Denote this distance by $\|\mathbf{z}_1 - \mathbf{z}_2, \alpha\|$ where this notation emphasizes the direct association with the weighting of the two blocks of the descriptors of the pixel. The scaling coefficient α brings a degree of flexibility but to take the full advantage of (4) there needs to be some effective scheme of its adjustment (calibration) of the numeric values of α .

It is worth stressing that the formation of the blocks of features (viz. descriptive and functional features) is reflective of the existing domain knowledge that is used to guide (navigate) the search for structure in data. In essence, we can regard the resulting clustering method as an example of the clustering algorithm falling under the umbrella of so-called knowledge-based clustering where the knowledge component is quantified via the split of the features into several semantically meaningful subsets.

3. The development of information granules through fuzzy clustering

The set of patterns consists of N vectors $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N$. Taking the distance (3) into account, the expressions governing the Fuzzy C-means (FCM) algorithm (Bezdek, 1981), using which we form information granules (clusters), are described as follows:

Partition matrix:

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|\mathbf{z}_k - \mathbf{v}_j, \alpha\|}{\|\mathbf{z}_k - \mathbf{v}_j, \alpha\|} \right)^{2/(p-1)}}. \quad (5)$$

Prototypes:

$$\mathbf{v}_i = \frac{\sum_{k=1}^N u_{ik}^p \mathbf{z}_k}{\sum_{k=1}^N u_{ik}^p}, \quad (6)$$

where $p > 1$, $i = 1, 2, \dots, c$, $k = 1, 2, \dots, N$. The fuzzification coefficient p is typically set up to be equal to 2. The derivations of these formulas are quite straightforward as the method is a certain minor modification of the standard version of the FCM algorithm. As it becomes apparent, the results of clustering depend directly upon the values of the scaling coefficient. While the boundary cases come with a straightforward interpretation ($\alpha = 0$ – only spatial location (descriptive features) is considered; $\alpha = 1$ all features are treated in the same way, high values of α – practically only the functional features of pixels are involved in the clustering process), this observation does not offer any hint as to an effective selection of the optimal value of α . To address the optimization of this clustering technique, in the next section we consider a minimization of the reconstruction criterion to quantify the prediction capabilities of the information granules constructed through fuzzy clustering.

4. The overall scheme of augmented clustering

The clear distinction made between descriptive and functional features brings forward a certain representation of the clustering problem as schematically portrayed in Fig. 2.

Let us elaborate on the main functional modules of the overall scheme:

- (a) Clustering (carried out with the use of the FCM endowed with the distance function (4)) gives rise to the structure described in terms of the partition matrix and the prototypes. These information granules are formed by minimizing the standard objective function as commonly encountered in the FCM clustering. Note that the results of clustering are impacted by the values of α . In this sense, by adjusting the values of this weight we have some control over the form of the resulting structure.
- (b) Based on the information granules revealed at the previous step, we engage this structure in predicting functional features on the basis of the descriptive features and the topology they entail. As illustrated in Fig. 2, the revealed structure which manifests in the descriptive features along with some given \mathbf{x} is used to predict the component of the functional features. The quality of prediction is quantified by means of some criterion (performance index) V . The

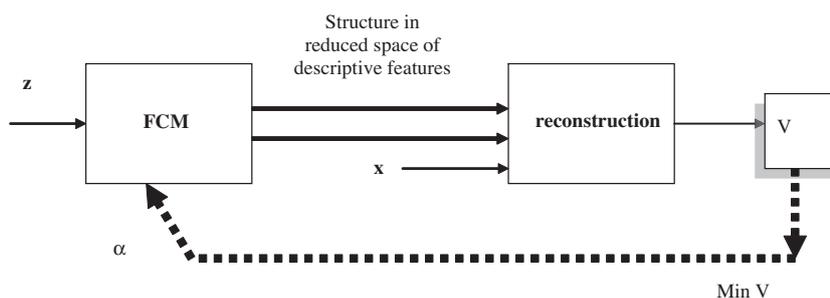


Fig. 2. Clustering with descriptive and functional features: a general flow of processing underlying the predictive facet of processing, which involves functional features.

minimization of V is realized by modifying the values of α as again illustrated in the same figure in terms of a certain feedback loop. The changes in the values of α impact the values of V and then the obtained minimum of V strikes a sound balance between the contributions made by descriptive and functional features.

5. Reconstruction criterion and its minimization

The concept of quality of reconstruction (Pedrycz and de Oliveira, 2008), which relates directly to the fundamentals of granular computing (Bargiela and Pedrycz, 2008) can be introduced as follows. As we indicated earlier, the predictive capability of the revealed structure is quantified by looking at the functional features. From this perspective, to implement the performance index V we introduce a concept of the reconstruction criterion.

Assume that the first block of the vector \mathbf{z} (namely \mathbf{x}) has been provided. For instance, the coordinates of the pixel are given or the input variables in case of system modeling are available. On the basis of the prototypes with coordinates available only in the same subspace as \mathbf{x} , we are interested to “reconstruct” the coordinates

in the m -dimensional space viz. the corresponding functional properties of \mathbf{x} . Denote this reconstruction (estimate) by \hat{y} . For the entire data set $(\mathbf{x}_k, \mathbf{y}_k)$ the overall reconstruction measure is expressed as

$$V = \sum_{k=1}^N \|\hat{\mathbf{y}}_k - \mathbf{y}_k\|^2. \quad (7)$$

The estimate $\hat{\mathbf{y}}_k$ is formed in an analogous way as presented by (1), (2) however all computations completed here are confined to the n -dimensional space. First, we compute the membership degrees of \mathbf{x}_k to the individual clusters:

$$\hat{u}_i = \frac{1}{\sum_{j=1}^c \left(\frac{\|\mathbf{x}_k - \tilde{\mathbf{v}}_i\|}{\|\mathbf{x}_k - \tilde{\mathbf{v}}_j\|} \right)^{2/(p-1)}}, \quad (8)$$

and then determine the estimate of the output (functional, viz. predictive features) in the following form:

$$\hat{\mathbf{y}}_k = \frac{\sum_{i=1}^c \hat{u}_i^p \tilde{\mathbf{v}}_i}{\sum_{i=1}^c \hat{u}_i^p}. \quad (9)$$

The notation $\tilde{\mathbf{v}}_i$ used in (6) underlines the fact that we consider only the first n -coordinates of the original prototype \mathbf{v}_i . The reconstruction error V is used to optimize the value of the weight α , namely:

$$\alpha_{\text{opt}} = \arg \min_{\alpha} V. \quad (10)$$

The tangible and easily quantifiable improvement offered by the proposed method can be captured in the form of the ratio $V(\alpha_{\text{opt}})/V(\alpha = 1)$. The lower the value of the ratio, the more significant the improvement offered by the clustering method. Note that the case $\alpha = 1$ corresponds to the situation where all features are treated in the same way, that is not distinction between their blocks exhibiting different semantics.

6. Experimental studies

In this section, we elaborate on the numeric aspects of the method by quantifying a beneficial aspect of the semantics of the sets of the variables and looking at some general trends between information granularity, optimal values of α and the ensuing values of the reconstruction criterion. The datasets considered here both synthetic data and those coming from the Machine Learning Repository (<http://archive.ics.uci.edu/ml/>).

As the clustering method augments the generic FCM, we used this particular method to assess the improvement achieved by distinguishing between the semantic blocks of features. The value of the fuzzification coefficient (p) is equal to 2; it is a typical value commonly encountered in the literature.

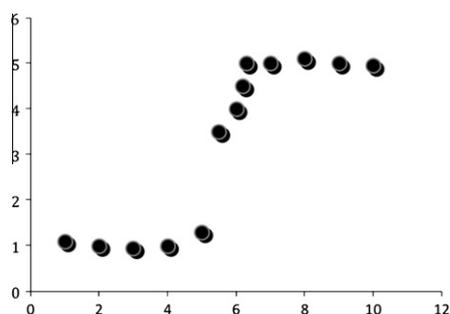


Fig. 3. Two-dimensional synthetic data set.

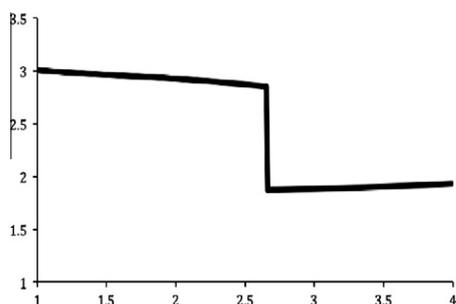


Fig. 4. V regarded as a function of α .

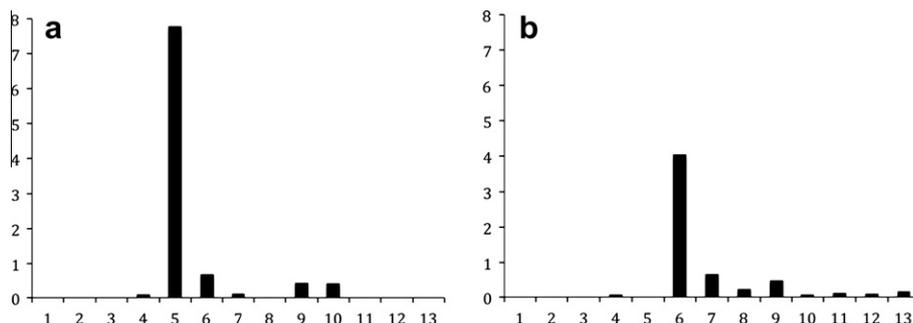


Fig. 5. Reconstruction error reported for individual data: (a) $\alpha = 1.0$ and (b) $\alpha = 2.7$.

6.1. Two-dimensional synthetic data

This small data set comprising 13 data (patterns), shown Fig. 3, is presented for illustrative purposes. The x -coordinate is a descriptive feature whereas the y coordinate is the functional feature. Note that in terms of the descriptive feature there is no apparent structure in the data as all patterns are distributed quite uniformly along this coordinate. Hence we may anticipate that the functional feature will become advantageous in capturing the structure in the data.

The optimization of V shows a clearly delineated minimum located at $\alpha = 2.7$, see Fig. 4 at which value there is a significant drop in the values of the performance index.

It becomes apparent that the optimization of the value of α is highly essential here. As visualized in Fig. 5, the distribution of the reconstruction error for each data point supports this claim even more profoundly. Overall, the ratio $V(\alpha_{opt})/V(\alpha = 1)$ is equal to $1.86/3.00 = 0.62$, which points a significant improvement over the case where no impact of semantic blocks is considered.

There is also a significant change in the distribution of the prototypes in these two cases:

$$\alpha = 1.0 \quad \mathbf{v}_1 = [8.97 \quad 4.99] \quad \mathbf{v}_2 = [6.14 \quad 4.33] \quad \mathbf{v}_3 = [2.86 \quad 1.05]$$

$$\alpha = 2.7 \quad \mathbf{v}_1 = [7.39 \quad 4.75] \quad \mathbf{v}_2 = [4.52 \quad 1.26] \quad \mathbf{v}_3 = [1.89 \quad 1.06]$$

Noticeably, for the optimal value of α , the prototypes start moving closer to the region in which there is a significant increase in the values of the functional variable. In contrast, if $\alpha = 1$ (in which case the first coordinate plays equally important role) the projection of the prototypes on the x -coordinate shows their quite uniform distribution in this space.

6.2. Auto-mpg data

We consider 6 variables ($n = 6$) as the first space while $m = 1$ and this space is formed by the fuel consumption variable (expressed in mpg). The distance function is specified as the weighted Euclidean distance with the weight coefficient being associated with the standard deviation of the corresponding variable. The plots of the reconstruction error regarded as functions of α for selected values of “ c ”, $c = 3, \dots, 7$ are included in Fig. 6. Some general comments are worth making here: (a) in each case there is an optimal value of α pointing at an ability to strike an optimal balance between structure discovery and the approximation (mapping) capabilities; (b) the reconstruction error is a highly asymmetric function of α with more substantial changes observed for lower values α . Once we have exceeded the optimal value of α , further increases of the values of this parameter do not translate into substantial increases in the values of the reconstruction criterion; (c) with the increase of the number of clusters the optimal values of α start moving towards higher values which is indicative of the need of higher impact from the structural hints reported for the variables of fuel consumption.

6.3. Fire

In this data set, by making use of the clustering results we predict the intensity of fire on a basis of its location and local weather condition. Table 1 shows the main results that is optimal values of α for different number of clusters and includes a ratio of the reconstruction error for α_{opt} and $\alpha = 1$ (the lower the value, the more effective the reconstruction). For higher number of clusters, optimal values of α get lower and the effectiveness of produced reconstruction tends to increase as well.

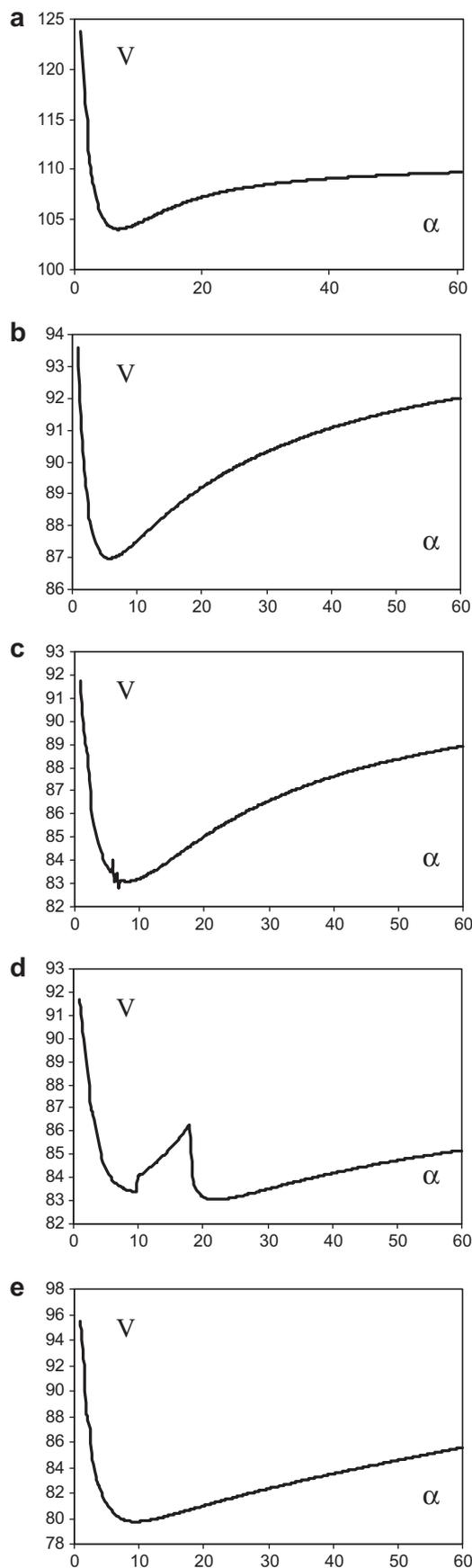


Fig. 6. Plots of $V = V(\alpha)$ for auto-mpg data and selected values of “ c ”: (a) $c = 3$, (b) $c = 4$, (c) $c = 5$, (d) $c = 6$, and (e) $c = 7$.

Table 1
Fire dataset – optimal values of α and the ratio of $V(\alpha_{opt})/V(\alpha = 1.0)$ for selected values of c , $c = 2, 3, \dots, 7$.

c	α_{opt}	$V(\alpha_{opt})/V(\alpha = 1.0)$
2	12.2	0.97
3	9.0	0.93
4	8.6	0.96
5	8.6	0.96
6	8.6	0.78
7	8.4	0.89

6.4. Boston housing

In this data, we consider the price of real estate to be the variable whose values are predicted on the basis of all remaining variables. The results are reported in the same manner as in the previous example with the values of the ratio $V(\alpha_{opt})/V(\alpha = 1.0)$ for selected number of clusters, see Table 2. Interestingly there is a certain tendency similar to the one reported so far: higher number of clusters, leads to the higher values of optimal α . The effectiveness of the functional features is more visible for the lower number of clusters (in this data we observe the ratio ranging from 0.75 for $c = 2$ to 0.50 when $c = 5$ or 6).

Table 2
Boston housing data – optimal values of α and the ratio of $V(\alpha_{opt})/V(\alpha = 1.0)$ for selected values of c , $c = 2, 3, \dots, 7$.

c	α_{opt}	$V(\alpha_{opt})/V(\alpha = 1.0)$
2	30.8	0.75
3	32.4	0.56
4	15.4	0.54
5	60.0	0.50
6	60.8	0.50
7	60.8	0.52

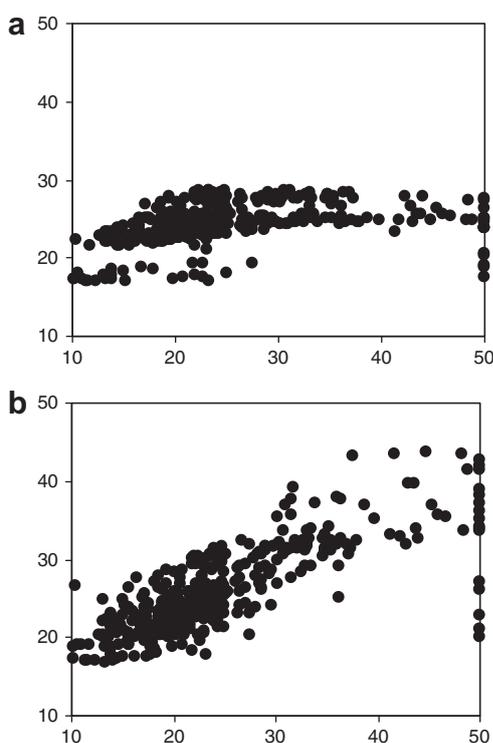


Fig. 7. Boston housing – reconstructed values of the functional feature versus the original one: (a) $\alpha = 0.0$ and (b) $\alpha = 60.0$.

Additional illustration of the performance of the method is visualized in Fig. 7 where we show the results of reconstruction vis-à-vis the values of the functional feature where $c = 5$ in case $\alpha = 1$ (no optimization – all features are treated in the same way), and $\alpha = 60$ (which is the optimal value of this coefficient). It is apparent that in the second case there is a far higher correlation between the output values and the corresponding reconstruction. In case of $\alpha = 1$, there is a significant “flattening” effect so the structure formed in the input space does not allow to fully reconstruct the outputs but rather reduces the resulting values down to a relatively narrow range (note that in the reconstruction the output does not exceed the upper limit of 30 whereas the original output goes up to 50).

The quantification of improvement realized by the clustering method proposed here can be visualized through the values of the ratio $V(\alpha_{opt})/V(\alpha = 1)$ obtained for successive values of “ c ”, see Fig. 8. In all cases, we see these values being below 1 with the more visible reduction achieved for the Boston housing data (in which case the reduction in the values of V was at the level of 50%).

6.5. Magic

The family of plots of the reconstruction criterion V is shown in Fig. 9; here we illustrated these relationships for selected values of “ c ”. Interestingly, in all cases we observe a minimal value of V which happens for some value of α . In other words, the reconstruction criterion is efficient in guiding the selection of the optimal tradeoff between descriptive and functional features. For this data

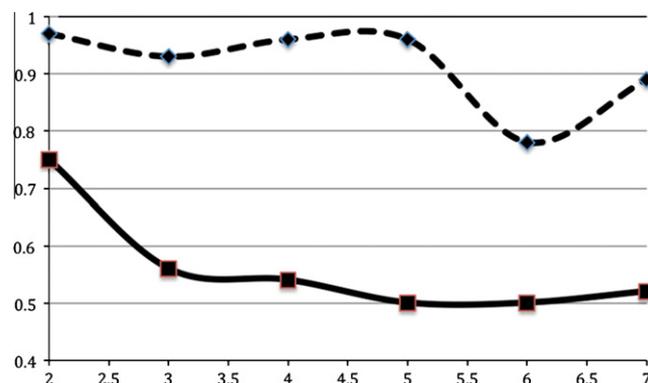


Fig. 8. The values of the ratio $V(\alpha_{opt})/V(\alpha = 1)$ versus the number of clusters: dotted line – fire, solid line – Boston housing.

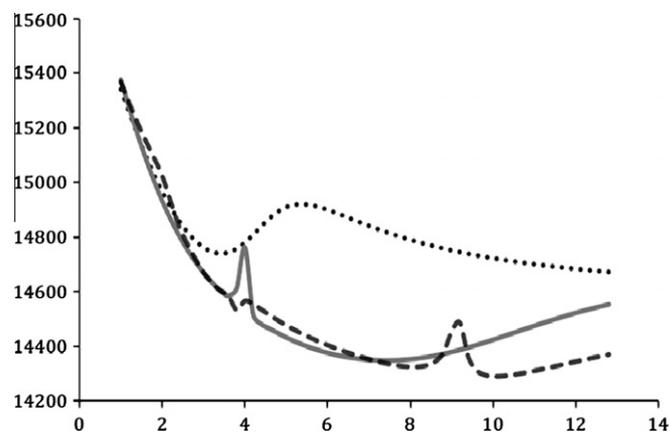


Fig. 9. V versus α for selected values of “ c ” for the Magic data, $c = 3, 4$ and 5 (the dotted curve corresponds to $c = 3$, $c = 4$ is related with the solid line and the lowest dashed line pertains to $c = 5$).

set, higher number of clusters associates with higher values of α and this demonstrates that in the determination of the structure we envision higher impact being associated with the functional feature.

7. Conclusion

Fuzzy clustering realized in presence of semantic groups of variables requires special treatment. The adjustments of the augmented distance function are necessary to minimize the reconstruction error, which leads to sound treatment of the functional features represented in the data. Experimental results reveal that the clustering of this nature is effective in the reconstruction of these features. This becomes particularly relevant when designing fuzzy models since in such modeling the functional features are their outputs. Here the reconstruction criterion reflects on the performance (accuracy) of the model. While this study has focused on a single vector of functional variables, this approach can be generalized to several families of such variables. In this case, one takes into account a vector of features composed of several sub-vectors, say $[\mathbf{x} \ \mathbf{y} \ \mathbf{w} \ \mathbf{t}]^T$, etc. Here one could regard some of them, say \mathbf{w} and \mathbf{t} as being of functional nature with the reconstruction problem involving all of them.

The proposed method can be sought as an example of the knowledge-based clustering. While the blocks of features coming as an essential component of domain knowledge are Boolean (two-valued) constructs, one can envision further relaxation by admitting a formation of fuzzy sets of descriptive and functional features.

References

- Bargiela, A., Pedrycz, W., 2008. Toward a theory of granular computing for human-centered information processing. *IEEE Trans. Fuzzy Syst.* 16 (2), 320–330.
- Berget, I., Mevik, B.J., Næs, T., 2008. New modifications and applications of fuzzy C-means methodology. *Comput. Stat. Data Anal.* 52 (5), 2403–2418.
- Bezdek, J.C., 1981. *Pattern Recognition with Fuzzy Objective Functions*. Plenum, New York.
- Biosca, J.M., Lerma, J.L., 2008. Unsupervised robust planar segmentation of terrestrial laser scanner point clouds based on fuzzy clustering methods. *ISPRS J. Photogramm. Remot. Sens.* 63 (1), 84–98.
- Cai, W., Chen, S., Zhang, D., 2007. Fast and robust fuzzy C-means clustering algorithms incorporating local information for image segmentation. *Pattern Recognit.* 40 (3), 825–838.
- Carballido-Gamio, J., Phan, C., Link, T.M., Majumdar, S., 2006. Characterization of trabecular bone structure from high-resolution magnetic resonance images using fuzzy logic. *Magn. Reson. Imaging* 24 (8), 1023–1029.
- Cinque, L., Foresti, G., Lombardi, L., 2004. A clustering fuzzy approach for image segmentation. *Pattern Recognit.* 37 (9), 1797–1807.
- Chuang, K.S., Tzeng, H.L., Chen, S., Wu, J., Chen, T.J., 2006. Fuzzy C-means clustering with spatial information for image segmentation. *Comput. Med. Imaging Graph.* 30 (1), 9–15.
- Hirota, K., Pedrycz, W., 1996. Directional fuzzy clustering and its application to fuzzy modelling. *Fuzzy Sets and Systems* 80 (3), 315–326.
- Kang, J., Min, L., Luan, Q., Li, X., Liu, J., 2009. Novel modified fuzzy C-means algorithm with applications. *Digit. Signal Process* 19 (2), 309–319.
- Kannan, S.R., 2008. A new segmentation system for brain MR images based on fuzzy techniques. *Appl. Soft Comput.* 8 (4), 1599–1606.
- Maitra, M., Chatterjee, A., 2008. Hybrid multiresolution slantlet transform and fuzzy c-means clustering approach for normal-pathological brain MR image segregation. *Med. Eng. Phys.* 30 (5), 615–623.
- Ozkan, I., Türkşen, I.B., Canpolat, N., 2008. A currency crisis and its perception with fuzzy C-means. *Inform. Sci.* 178 (8), 1923–1934. 15.
- Pedrycz, W., 1996. Conditional fuzzy C-means. *Pattern Recognit. Lett.* 17, 625–632.
- Pedrycz, W., 1998. Conditional fuzzy clustering in the design of radial basis function neural networks. *IEEE Trans. Neural Networ.* 9, 601–612.
- Pedrycz, W., Valente de Oliveira, J., 2008. A development of fuzzy encoding and decoding through fuzzy clustering. *IEEE Trans. Instrum. Meas.* 57 (4), 829–837.
- Peñaranda, J.A., Flores, C., Corzo, L.G., Gutierrez, I., 2006. Colour and proportion of grains measurement in through-body porcelain tiles by means of image processing. *J. Eur. Ceram. Soc.* 26 (15), 3373–3381.
- Saaty, T.L., 2000. *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*. RWS, Pittsburgh, PA.
- Silva, S., Júnior, M.D., Lopes Jr., V., Brennan, M.J., 2008. Structural damage detection by fuzzy clustering. *Mech. Syst. Signal Process.* 22 (7), 1636–1649.
- Xia, Y., Feng, D., Wang, T., Zhao, R., Zhang, Y., 2007. Image segmentation by clustering of spatial patterns. *Pattern Recognit. Lett.* 28 (12), 1548–1555.
- Verikas, A., Malmqvist, K., Bergman, L., 2005. Detecting and measuring rings in banknote images. *Eng. Appl. Artif. Intel.* 18 (3), 363–371.