Fast Algorithm for Finding True Number of Clusters. Applications to Control Systems

Mikhail Morozkov, Oleg Granichin, Zeev Volkovich and Xuming Zhang

Abstract— **One of the most difficult problems in cluster analysis is the identification of the number of groups in a given data set. In this paper we offer the randomized approach in the rate distortion framework. A randomized algorithm has been suggested to allocate this position. The scenario approach is used to significantly reduce the computational complexity. With ability to determine the true number of clusters and perform clustering in real-time operational mode we outline several applications in control systems and decision-making problems that can benefit from algorithm in question essentially. We also provide simulation results to show considerable speed optimization with guaranteed level of probability.**

Index Terms— **Clustering, Randomized Algorithms, Adaptive Control, Optimitzaion**

I. INTRODUCTION

More and more recently emerging prolems in control systems require algorithms to be able to give a solution in real-time. Such algorithms have been widely used in such fields as robotics, model predictive control or supply chain optimization, but sample times are measured once in several minutes, which is not exactly a real-time operational mode. But, of course, there are a lot of systems with much faster dynamics, that call for the execution times to be measured in seconds at the most, which is why the majority of nowadays "real-time" algorithms are simply inappropriate. Such state of affairs results in very high demand for really fast, online approaches. Hopefully, the problem has been acknowledged, for example, prof. Boyd concerned this problem in his plenary speech at the last Multi-Conference on Systems and Control [1]. Their implementation of the automatic code generation system CVXGEN, which scans a description of the problem family and performs much of the analysis and optimization of the algorithm, yields an extremely fast solution with execution times measured in milliseconds or microseconds for small and medium size problems.

But generally one has to deal with many complex problems like disturbances, uncertainty, NP-hardness etc. In order to significantly reduce computational costs and therefore drastically decrease computing time, it is reasonable to exploit stochastic and randomized approaches, such as novel Randomized Scenario Approach (first introduced by Campi [2] and Calafiore [3]), which is developed for solving RCP (Robust Convex Problem). The primary idea of this technique is to take into account only a certain and relatively small amount of sampled instances of uncertainty (the scenarios) affecting the system in question and then solve corresponding problems. The main result is that the number of samples to be considered is reasonably small and solution obtained from corresponding convex problems satisfies with up to a guaranteed level of probability that can be set arbitrarily close to one to all unseen scenarios as well.

One of a widely unsolved problem in cluster analysis is finding the true number of clusters in given data set. There have been suggested a great number of approaches during recent decades, such as different index function-based techniques (Hartigan [4], Sugar and James [5]), stability models (Levine and Domany [6], Jain and Moreau [7]) and probability models (Cuevas et al [8] and [9], Volkovich et al [10]). Elbow methodology is often employed as a stopping rule to determine the number of clusters in data set. A number of clusters is chosen such that adding another cluster doesnt give much better modeling of the data set. The number of clusters is chosen at such elbow point. Unfortunately, many of the approaches concerned were developed for a specific problem and are somewhat ad hoc.

Recently, rate distortion theory has been applied to choosing true number of clusters due to sufficient mathematical background provided in [5]. In this version of jump method, the procedure is based on distortion curve, which is a measure of within cluster dispersion. The distortion curve is then transformed by a negative power selected based on the dimension of the data. Largest jump of obtained curve signifies the best alternative for the true number of clusters. Nevertheless, on top of its advantages such technique implies high computational complexity. From future control systems point of view it is important to develop algorithms which are able to work in real-time. With the ability to perform clustering in the real-time one can extend its application to other related fields, such as logistics, multi-agent systems, sensor networking, wireless communications, dynamic coverage control, bioinfarmatics and others.

In this paper we spread the ideas of scenario approach to one of this challenging problem of finding true number of clusters in order to significantly reduce computational costs and introduce applications of our algorithm to control systems as a decision-making tool. The paper is organized as follows: we describe our implementation of randomized algorithm of finding true number of clusters in Section II.

M. Morozkov and O. Granichin are with Department of Mathematics and Mechanics, Saint Petersburg State University, 198504, Universitetskii pr. 28, St. Petersburg, Russia mmorozkov@gmail.com, oleg_granichin@mail.ru

Z. Volkovich is with Software Engineering Department, ORT Braude College of Engineering, Karmiel 21982, Israel vlvolkov@braude.ac.il

X. Zhang is with Department of Biomedical Engineering, School of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China xmboshi.zhang@gmail.com

978-1-4577-2072-7/12/$26.00

Then in Section III we provide different applications of algorithm in question. Further we introduce obtained results in Section IV and briefly list the conclusions and future work directions in Section V.

II. RANDOMIZED ALGORITHM

Based on the inner index approach proposed by Sugar and James [5] in the framework of rate distortion theory, the task of determining the true number of clusters can be formally be represented as follows. Let x be a n-dimensional random variable having a mixture distribution of k^* components, each with covariance Γ. Then the minimum achievable distortion associated with fitting k centers to the data is

$$G_k = \frac{1}{n} \min_{C_1,\ldots,C_k} E\left[(x - C_x)^T \Gamma (x - C_x),\right],$$

where $C_1, \ldots, C_k$ is a set of k cluster centers obtained by running a standard clustering procedure; C_x is the nearest cluster center to x. Note that in the case where Γ is the identity matrix distortion is simply mean squared error. Given the distortions G_k, a "jumping differential" curve is constructed according to the following rule:

$$J_k = G_k^{-\lambda} - G_{k-1}^{-\lambda},$$

where λ is the transformation power. According to asymptotic results obtained from distortion rate theory [5], its preferred value is

$$\lambda = n/2.$$

Moreover, for rather high values of n, the differential distortion J_k asymptotically tends to zero subject to the number of clusters k is less than the number of components k^*. Thus, for big n, transformed distortion J_k is almost zero when $k < k^*$, then the value jumps abruptly and increases linearly when $k \geq k^*$. The Sugar and James' "jump" algorithm exploits such behavior in order to determine the most likely value of k as the true number of clusters. The estimated number of clusters corresponds to the index of maximal value of transformed distortion function J_k:

$$k^* = \arg\max_k J_k.$$

Thus, the problem of finding the true number of clusters can be interpreted as a particular case of more general problems, namely, the fault detection or the problem of locating the discontinuity points of an implicitly defined function. Generally, the problem can be formulated as follows. Let us take a real-valued function f on the interval $[0; 1]$ having no more than one jump discontinuity $x^* \in (0; 1)$. The problem concerned in [11] consists in finding confidence interval for x^* subject to:

1) The function $f(\cdot)$ is Lipschitz continuous with a Lipschitz's constant C on the intervals $(0; x^*)$ and $(x^*; 1)$;

$$\mathbf{A1}: \quad C \geq \max_{j=2,\ldots,k^*-1,k^*+1,\ldots,k_{max}} |I(j) - I(j-1)|,$$

2) If jump discontinuity exists at the point x^*, then the jump size at this point is above a certain constant value $B > 0$.

$$\mathbf{A2}: \quad B \leq |I(k^*) - I(k^* - 1)|.$$

To implement the above methodology in the framework of the clustering concept, consider the transformed "distortions" $I(\cdot)$, proposed by Sugar and James [5]. Without loss of generality, assume that $I(0) = I(1)$ and introduce a continuous piecewise linear function f as follows:

$$f_I\left(\frac{k}{k_{max}}\right) = I(k),$$

$$f_I(x) = I(k) + \left(x - \frac{k}{k_{\max}}\right)(I(k+1) - I(k)),$$

for

$$\frac{k}{k_{\max}} \leq x \leq \frac{k+1}{k_{\max}}, k = 0, \ldots, k^* - 2, k^*, \ldots, k_{\max} - 1,$$

$$f_I(x) = I(k^* - 1),$$

for

$$\frac{k^* - 1}{k_{\max}} \leq x \leq \frac{k^*}{k_{\max}}.$$

An algorithm which implements the approach in question can be described as follows:

1) Choose the reliability parameter $\beta \in (0, 1)$.
2) Choose the parameters N, T so that:

$$T = \left[\frac{4\sqrt{2}Ck_{\max}}{\sqrt{1-\beta}BN} - \frac{1}{N}\right] + 1. \tag{1}$$

3) Choose randomly T groups of N points from interval $(0, 1)$:

$$Z_t = \{z_{tn},\ n = 1, ..., N\},\ t = 1, \ldots, T.$$

and denote

$$Z = \bigcup_t Z_t.$$

Below in the proof of Theorem 1 it will be shown that the largest distance between two sequential points belonging to Z does not exceed $\frac{B}{4C}$ with probability of $1 - \beta$.

4) Choose $D, M > 0$. For each one of the groups Z_t, $t = 1, \ldots, T$ construct uniform approximation for $f_I(x)$:

$$g_t(x) = \sum_{m=0}^{M} d_{tm} p_m(x),\ t = 1, \ldots, T, \tag{2}$$

minimizing the error

$$\gamma_t = \max_{x \in Z_t} |g_t(x) - f_I(x)|$$

subject to

$$|d_{tm}| \leq D,\ m = 0, \ldots, M,\ t = 1, \ldots, T.$$

Here a convex optimization MATLAB TOOLBOX (YALMIP, SeDuMi or cvx) can be applied.

If one of the approximation problems is not resolved then one has to return to step 1 and to start over.

5) The the set

$$\Delta = \{\tilde{x} = xk_{max} : x \in (0;1),\ \chi(x) > L(x)\} \quad (3)$$

where

$$\chi(x) = \max_{t=1,\dots,T} g_t(x) - \min_{t=1,\dots,T} g_t(x),\ x \in (0;1) \quad (4)$$

$$L(x) = B - \frac{B}{4Ck_{\max}} h(x) - 2\gamma,$$

$$h(x) = \max_{z\in[z_l(x),z_r(x)]} \max_{t=1,\dots,T} |g_t'(z)|, \quad (5)$$

is a confidence interval fo rx^* and following theorem holds:

Theorem 1: If conditions **A1** *and* **A2** *formulated above are satisfied,* **then** *the set* Δ *defined in (3) is not empty and contains the point* x^*k_{max}*, equal to the true number of clusters with probability of at least* $p = \beta$.

III. APPLICATIONS

From future control systems point of view it is important to develop algorithms which are able to work in real-time. With the ability to perform clustering in the real-time operation mode one can extend its appliance to other related fields, such as logistics distribution, various graph theory problems, multi-agent and mobile systems, networking, wireless communications, dynamic coverage control, bioinformatics and medicine. In this section we provide three possible applications of our fast randomized algorithm for determining the true number of clusters, and explain how some decision-making problems in control systems can benefit essentially.

A. Dynamic Coverage Control

In the most general sense, the Dynamic Coverage Problem main objective can be summarized as: identify clusters of mobile objects and determine the optimal locations and dynamics of resources such that these resources continuously provide adequate coverage of the mobile objects, throughout the time horizon of interest.

Dynamic coverage problems inherit the comuptational complexity of facility location problems that arise in a variety of static applications such as locational optimization, facility location, optimal coding, pattern recognition and learning, and data clustering and classification [12]. The static coverage problems are known to be NP-hard [13], where the cost functions are non-convex and are typically riddled with multiple local minima. This complexity is compounded further by the inclusion of dynamics of constituent elements. Even in the static setting, there are relatively few algorithms that develop mechanisms to inhibit getting trapped at local minima and reduce sensitivity to initial conditions.Therefore, it makes related problems a perfect opportunity to expand basic ideas of our algorithm in order to significantly reduce computational costs with guaranteed probability level of success.

Problems related to dynamic coverage are considered in [14], [15] where the emphasis is on distributed implementations, i.e., under limited information flow between individual elements. These algorithms have the advantage of distributed implementation, however they are sensitive to the initial placement of the resources and suffer from drawbacks analogous to those found in Lloyds algorithm. In contrast to this distributed approach, there is scant research that addresses the development of algorithms for problems of a non-distributed nature, or that aim simultaneously to attain global solutions and maintain low computational expense. In [16], [17], a maximum entropy principle (MEP)-based approach is discussed where dynamic coverage of mobile objects under given velocity fields is achieved by designing corresponding velocity fields for the resources.

Recent work of Yunwen Xu, S. Salapaka and C. L. Beck [18] was on the the first, which considered tracking cluster centers when site dynamics involve accelerations.

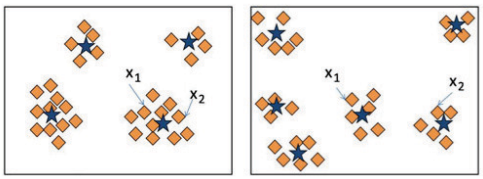

Fig. 1. Clustering moving objects in a given area. The squares and stars denote the positions of the sites and resources respectively. Sites x_1 and x_2 reside in the same cluster at the time instance shown in the left figure. A split occurs and causes them to reside in different clusters at the time instance shown in the right figure.

B. Medical Image Segmentation

As the technique used to create images of the human body, medical imaging plays an important role in medical diagnosis and surgical guidance. Medical image segmentation is very critical in the quantification process in that it can assist clinicians by extracting boundaries, surfaces or volumes of organs. However, uncertainty is widely present in medical images because of noise and other imaging artifacts, poor contrast and intensity variations. Among the various clustering methods, the unsupervised fuzzy clustering methods, particularly fuzzy C-means (FCM) algorithm, have been widely used for medical image segmentation in that they are effective in dealing with uncertainty.

Despite its popularity in medical image segmentation, the traditional FCM has many problems to be settled for accurate and robust medical image segmentation. It has been well known that the traditional FCM algorithm is sensitive to noise inherent in medical images because it does not take into account the spatial dependencies between the clustered data. To overcome the disadvantageous influence of noise, Shen et al [19] introduced the feature difference between neighboring pixels in brain magnetic resonance (MR) images and the relative locations of neighboring pixels to modify similarity measure in FCM. Yu et al [20] have proposed 2DFCM method in which the spatial constraints provided by the denoising

data and the Gabor wavelet based textural information are introduced to produce satisfactory segmentation results for corrupted images with intensity variations. To address the ignorance of the spatial neighborhood information in FCM , an improved spatial FCM algorithm has been proposed, which incorporates the spatial neighborhood information into the traditional FCM algorithm based on the membership function of the center pixel and its neighboring pixels [21]. Zhang et al [22] have proposed the kernelized fuzzy C-means (KFCM) algorithm, which is realized by modifying the objective function in the conventional FCM algorithm using a kernel-induced distance metric and a spatial penalty on the membership functions.

Additionally, the FCM has the drawback that the desired number of clusters must be specified in advance. This is a disadvantage for medical image segmentation in that the ground truth is always not available for medical imagery. To address this problem, Li et al [23] have proposed a modified fast FCM algorithm for automatic MR brain image segmentation. The advantage of this algorithm is that it can determine the number of clusters automatically by combining the Otsu algorithm [24] with fast FCM algorithm. Hung et al [25] have presented the modified suppressed fuzzy c-means method. This method performs clustering and parameter selection for the suppressed FCM algorithm simultaneously and it can easily select the parameter in the suppressed FCM based on exponential separation strength among clusters.

Owing to the iterative nature, the FCM method is computationally intensive. To increase its implementation efficiency, a high speed parallel fuzzy C means algorithm for brain tumor image segmentation has been presented by S. Murugavalli et al [26]. The algorithm is highly efficient in that it has the advantages of both the sequential FCM and parallel FCM for the clustering process in the segmentation techniques. Zhou et al [27] have introduced a mean shift based fuzzy c-means algorithm by introducing mean field term into FCM. This algorithm requires less computational time than the traditional FCM and therefore it can be effectively used for extracting skin lesion borders from dermoscopic images.

Apart from the modification of the FCM method, it has been combined with other methods to realize medical image segmentation. S. Pradhan et al [28] have combined FCM with the hidden Markov random field (HMRF) model for brain MR image segmentation. The proposed algorithm uses the HMRF model to model the image class labels and offers an FCM-type treatment of the HMRF model. This MRF-FCM algorithm provides an effective means for brain MR image segmentation in that it integrates the spatial coherency modeling capabilities of the HMRF model with the flexibility of FCM-type method. Li et al [29] combines spatial fuzzy clustering with level set methods for automated medical image segmentation. Here the initial segmentation is produced by spatial fuzzy clustering. The fuzzy level set algorithm directly evolves from it and the controlling parameters of level set evolution are also estimated from the results of fuzzy clustering.

C. Genomes Classification

The problem of bacteria classification arose long before the start of the Genomic Era. The approach that existed at that time is known as species categorization. Bacteria were divided into groups on the basis of one or more traits such as morphological, physiological and environmental. Using the Compositional Spectrum method, it is possible to investigate the connection between the frequency pattern and various natural bacteria classifications established on the basis of categorization or phylogeny. The goal of the study is in finding the answer to the question as to whether entire genomes or their large parts reflect the latter classifications.

An important part of a genomes classification is the determination of the natural number of groups where genomes belonging to same group are more similar one to another in comparison of genomes found in different subsets. This is also essential for our purposes since we are going to compare a formally obtained categorization with the biological one, which implies that the number of groups must be determined in a formal way using only the Compositional Spectrum distances, without any additional biological information in the dataset. The estimation of the number of groups presents an ill-posed problem of crucial relevance in cluster analysis. The correct number of clusters in a dataset can depend on the scale at which the elements are measured. The approaches for solving this problem apply, in particular, the following two methodologies. The first one is based on the geometrical features, such as dispersion within and between the clusters. The second one, based on the cluster stability properties, considers the closeness of repeated results of the algorithm. The closeness can be characterized by means of statistical stability.

However, in the case when the expected number of clusters is relative big, the accepted methods possess frequently very high computational complexity. Such a problem arises naturally in problems connected to the separation of bacteria mixtures ones the number of bacteria is unknown and has to be estimated using only the data itself.

IV. SIMULATION RESULTS

To check whether the proposed algorithm can be applied to a large number of clusters, a synthetic dataset was generated. It contained 3100 clusters, each composed of 8 - 19 instances. Instances in each cluster were generated according to a Gaussian distribution on the unit square $[0, 1] \times [0, 1]$ with a random center for each cluster.

We consider the interval $[1, 8000]$ which contains the real number of clusters. For each point the transformed distortion function $I(k)$ is calculated using the algorithm of Sugar and James. Note, for each $k = 1, \ldots, 8000$ we proceed the clustering algorithm (k-means) and after that computed $I(k)$. The results are presented in Fig. 2.

The scenario approach which was described above allows us to significantly reduce the number of clustering algorithm computation. Assuming that $B > 1.0$ and $C < 0.002$, we choose $\beta = 0.9$; $M = 20$; $N = 29$; $D = 0.6$. Hence, subject to (1) $T = 10$. Thus, having computed only 290 values of

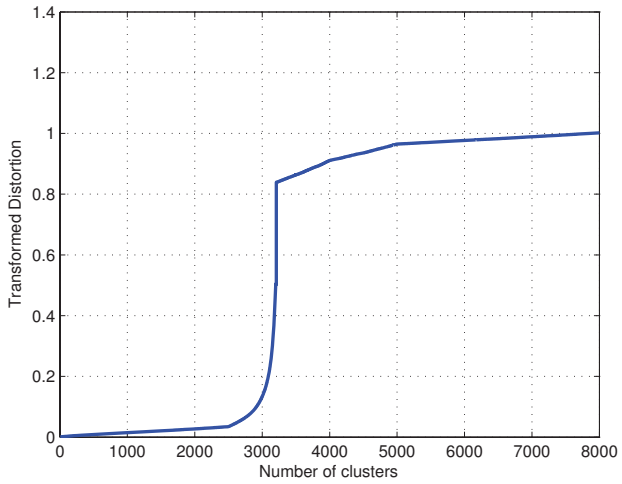

Fig. 2. Index function for initial data set.

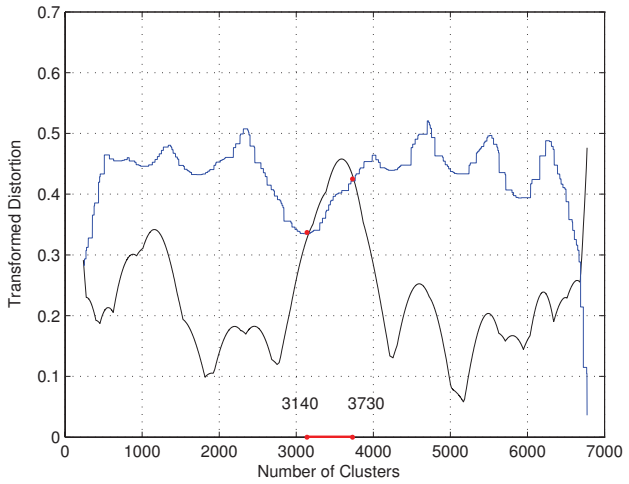

Fig. 4. Resulting function and confidence level.

$I(k)$, we reduced interval of uncertainty 14 times, instead of computing 7400 values, as traditional deterministic approach implies, in order to obtain the confidence interval Δ with probability of at least $\beta = 0.9$. Three approximation curves $g_t(\cdot)$ are shown on Fig. 3 along with the resulting function $\chi(\cdot)$.

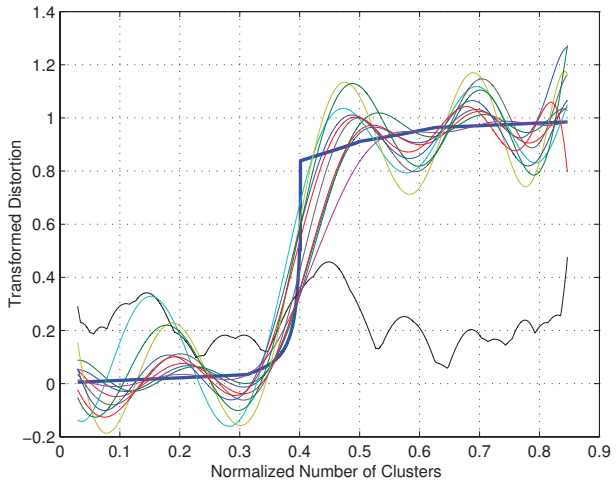

Fig. 3. Approximation curves.

With the assumption $B > 1.0$ and $C < 0.002$ we obtain the level of decision making, which is shown in Fig. 4 along with the resulting function $\chi(\cdot)$. A peak near the point $x^* = 3200$ can be observed. So if we consider segment $[2750, 3410]$ to be the confidence interval Δ, then for obtainig eventual solution one needs to perform 660 computations of index function $I(x)$. Thus the total number of index function computations equals to 950, which is considerably less than initial number of 8000.

V. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

A new randomized algorithm of stable clustering has been suggested. The algorithm is based on randomized approach and ideas of scenario approach along with well-known algorithms of cluster number determination using inner index functions. Significant decrease of computational complexity is theoretically proved for sufficiently generalized case. Given the speed optimization of the proposed algorithm allows to determine the true number of clusters in real-time, several possible applications for the algotirhm in question have been outlined.

B. Future Works

In future work, it will be of interest to further develop our randomized approach. The choice of clustering algorithm is subject to investigation. Originally we opted for k-means due to its simplicity, but there is a major drawback of all techniques similar to k-means, they often converge on local minima rather than the global minimum. For example, Rose in [30] and [31] proposed an algorithm based on a deterministic annealing optimization method, which was shown to be capable of avoiding local minima in situations where descent minimization algorithms (i.e k-means) failed to do so.

The major future challenge is to apply our algorithm to actual problems in the field of control systems and decision-making as we listed in Section III. It is necessarily specify more precisely proper metrics for each case, pick and deploy faster and more correct clustering algorithm.

REFERENCES

[1] *Boyd S., Mattingley J., Grant M. and Wang Y.* Real-Time Embedded Convex Optimization // 2011 IEEE Multi-Conf. on Systems and Control, September 2011.

[2] *Calafiore G. and Campi M.C.* The scenario approach to robust control design // IEEE Trans. Automat. Control, vol. 51, no. 5, May 2006, pp. 742–753.

[3] *Calafiore G. and Fagiano L.* Robust Model Predictive Control: The Random Convex Programming Approach (I) // In Proc. of 2011 IEEE Multi-Conf. on Systems and Control, September 2011.
[4] *Hartigan J.* Statistical theory in clustering // J. Classification, vol. 2, 1985.
[5] *Sugar C. and James G.* Finding the number of clusters in a data set: An information theoretic approach // J. of the American Statistical Association, vol. 98, 2003, pp. 750–763.
[6] *Levine E. and Domany E.* Resampling method for unsupervised estimation of cluster validity // Neural Computation, vol. 13, 2001, pp. 2573–2593.
[7] *Jain A.K. and Moreau J.V.* Bootstrap technique in cluster analysis // Pattern Recognition, vol. 20. 1987, pp. 547–568.
[8] *Cuevas A., Febrero M. and Fraiman R.* Estimating the number of clusters // The Canadian Journal of Statistics, vol. 28(2), 2000, pp. 367–382.
[9] *Cuevas A., Febrero M. and Fraiman R.* Cluster analysis: A further approach based on density estimation // Computational Statistics and Data Analysis, vol. 28, 2001, pp. 441–459.
[10] *Volkovich Z., Barzily Z. and Morozensky L.* A statistical model of cluster stability // Pattern Recognition, vol. 41, 2008, pp. 2174–2188.
[11] *Granichin O. and Khalidov V.* Randomized approach to finding function discontinuity // Stochastic Optimization in Computer Science, vol. 1(1), 2005, pp. 73–80.
[12] *Sharma P., Salapaka S., and Beck C.* A Scalable Approach to Combinatorial Library Design for Drug Discovery // Journal of Chemical Information and Modeling, vol. 48, no. 1, 2008, pp. 2741.
[13] *Aloise D., Deshpande A., Hansen P., and Popat P.* Np-hardness of euclidean sum-of-squares clustering // Mach. Learn., vol. 75, no. 2, 2009, pp. 245248.
[14] *Cortes J., Martnez S., Karatas T. and Bullo F.* Coverage Control for Mobile Sensing Networks: Variations on a Theme // Proceedings of Mediterranean Conference on Control and Automation, July 2002.
[15] *Sheu J.B.* A Fuzzy Clustering-Based Approach to Automatic Freeway Incident Detection And Characterization // Fuzzy Sets Syst., vol. 128, no. 3, 2002, pp. 377388.
[16] *Sharma P., Salapaka S., and Beck C.* Entropy Based Algorithm for Combinatorial Optimization Problems with Mobile Sites and Resources // In Proc. of American Control Conference, 2008, pp. 12551260.
[17] *Sharma P., Salapaka S., and Beck C.* Entropy-based framework for dynamic coverage and clustering problems // IEEE Transactions on Automatic Control, 2010.
[18] *Xu Y., Salapaka S. and Beck C.L.* Dynamic Maximum Entropy Algorithms for Clustering and Coverage Control // In Proc. of 49-th IEEE Conf. on Decision and Control, 2010, pp. 1836–1841.
[19] *Shen S., Sandham W.A., Grant M.H., Patterson J. and Dempsey M.F.* Fuzzy clustering based applications to medical image processing // Proc. IEEE EMBS 25th Annu. Int. Conf., pp. 747750, 2003.
[20] *Yu J. and Wang Y.* Molecular Image Segmentation Based on Improved Fuzzy Clustering // Proc.of International Journal on Biomedical Imaging, vol. 2007, no. 1, Jan. 2007.
[21] *Beevi S.Z., Sathik M.M., Kannan D.S.* A robust fuzzy clustering technique with spatial neighborhood information for effective medical image segmentation // International Journal of Computer Science and Information Security, Vol. 7, No. 3, pp. 132138, 2010.
[22] *Zhang D., Chen S.* A novel means algorithm with application in medical image segmentation // Artificial Intelligence in Medicine, vol. 32, pp. 3750, 2004
[23] *Li M., T. Huang T., Zhu G.* Improved fast fuzzy c-means algorithm for medical MR images segmentation // Proceedings of the WGEC Second International Conference on Genetic and Evolutionary Computing, pp. 285288, 2008.
[24] *Otsu N.* A Threshold Selection Method from Gray Level Histograms // IEEE Trans. Systems, Man, and Cybernatics, vol. SMC-8, pp. 6266, 1979
[25] *Hung W., Yang M., Chen D.* Parameter selection for suppressed fuzzy c-means with an application to MRI segmentation // Pattern Recognition Letters, Vol. 27, No. 5, pp. 424-438, 2006.
[26] *Murugavalli S., V. Rajamani V.* A high speed parallel fuzzy c-mean algorithm for brain tumor segmentation // ICGST International Journal on Bioinformatics and Medical Engineering, Vol. 6, No. 1, pp. 2934, 2006.
[27] *Zhou H., Schaefer G., Sadka A., Celebi M.E.* Anisotropic mean shift based fuzzy c-means segmentation of skin lesions // Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology, pp. 438443, 2008.
[28] *Pradhan S., D. Patra D.* Unsupervised brain magnetic resonance image segmentation using HMRF-FCM framework // 2009 Annual IEEE India Conference IEEE India Conference (INDICON), Dec, 2009.
[29] *Li B.N., Chui C.K., Chang S., Ong S.H.* Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation // Computers in Biology and Medicine, vol 41, pp. 110, 2011.
[30] *Rose K.* Deterministic Annealing, Clustering and Optimization. // Ph.D. Thesis, California Inst. Technol, Pasadena, 1991.
[31] *Rose K.* Deterministic annealing for clustering, compression, classication, regression and related optimization problems. // Proc. IEEE, Vol. 86, No. 11, 1998, pp. 2210–2239.