

Hybrid Objective Function of Fuzzy c -Varieties and Cross-shape Fuzzy Cluster Extraction

Daisuke Yoshida, Katsuhiro Honda, Akira Notsu, Hidetomo Ichihashi

Department of Computer Science and Intelligent Systems

Osaka Prefecture University

1-1 Gakuen-cho, Nakaku, Sakai, Osaka 599-8531, Japan

Email: {honda, notsu, ichi}@cs.osakafu-u.ac.jp

Abstract—This paper proposes an FCM-type clustering method for capturing cross-shape fuzzy clusters in multi-dimensional data spaces. The proposed objective function is a combination of Fuzzy c -Varieties (FCV) and cross-shape cluster extraction in 2-D space, which is an extended linear fuzzy clustering model with local coordinate rotation. FCV is responsible for finding 2-D planes, on which cross-shape prototypes exist. Each prototypical cross is estimated on the FCV prototypes. Fuzzy memberships are updated using a combined clustering criterion of distances between data samples and FCV prototypes and measures for linear clustering on the FCV prototypes with local coordinate rotation.

Keywords—co-clustering, collaborative filtering, principal component analysis.

I. INTRODUCTION

Many variants of FCM-type clustering models [1], [2], [3] have been proposed by extending the original Fuzzy c -Means (FCM) algorithm [4], which is a fuzzy counterpart of k -Means clustering [5]. If we have some structural information on cluster prototypes, we can modify the point-type prototypes of FCM by replacing with other shapes. Fuzzy c -Lines (FCL) [6] is a linear fuzzy clustering model, in which cluster prototypes are defined by lines having infinite length. Fuzzy c -Varieties (FCV) [7] extended FCL, in which cluster prototypes are defined by linear varieties such as planes and hyper-plane. These linear clustering models have close connection with local principal component analysis (local PCA) [8], [9], in which fuzzy PCA [10] is performed in each cluster based on the FCM-type alternate optimization scheme [11], [12].

Using other non-linear-shape prototypes, many FCM variants such as Fuzzy c -Spherical Shells (FCSS) [13] with circular-shape prototypes, Fuzzy c -Quadric Shells (FCQS) [14] with quadric surfaces and Fuzzy c -Rectangular Shells (FCRS) [15] with rectangular-shape prototypes have been proposed. Yoshida *et al.* [16] proposed a technique for extracting cross-shape clusters in 2-D data space by introducing local coordinate rotation in linear fuzzy clustering. Each cross-shape cluster is composed of two linear sub-clusters and is captured in a single local minor component extraction context [17] by considering coordinate rotation of one sub-cluster around intersection of two linear sub-prototypes.

This work was funded in part by the Ministry of Education, Culture, Sports, Science and Technology, Japan, under Grant-in-Aid for Scientific Research (20700214).

In this paper, the cross-shape extraction model is extended to multi-dimensional cases. In the Gaussian mixture models (GMMs) field, a combined likelihood function was proposed for capturing 2-D non-linear manifolds [18]. In the same manner, this paper proposes a combined objective function of FCV and cross-shape cluster extraction. The objective function for FCV composed of distances between data samples and 2-D plane-like prototypes is responsible for capturing 2-D planes, on which each cross-shape exists. Each cross-shape prototype is estimated in the prototypical 2-D planes. The clustering criteria for updating fuzzy memberships are given by the two criteria.

The remaining parts of this paper are organized as follows. Section II briefly reviews the FCM-type linear fuzzy clustering models and the cross-shape prototype model. Section III proposes a combined objective function. Numerical experimental result is presented in Section IV. Finally, Section V gives conclusion.

II. FCM-TYPE LINEAR FUZZY CLUSTERING AND CROSS-SHAPE CLUSTER EXTRACTION

A. Fuzzy c -Lines and Fuzzy c -Varieties

Assume that we have n samples with m dimensional observation $\mathbf{x}_i = (x_{i1}, \dots, x_{im})$, $i = 1, \dots, n$ and the goal is to partition the samples into C clusters. The FCM algorithm uses point-type prototypes given by mean vectors $\mathbf{b}_c = (b_{c1}, \dots, b_{cm})$, $c = 1, \dots, C$ and the clustering criterion for cluster c is the squared Euclidean distance d_{ci} between sample $\mathbf{x}_i$ and prototype $\mathbf{b}_c$. The objective function to be minimized is the weighted within-sum-of-distances.

$$\begin{aligned} L_{fcm} &= \sum_{c=1}^C \sum_{i=1}^n u_{ci}^\theta d_{ci} \\ &= \sum_{c=1}^C \sum_{i=1}^n u_{ci}^\theta \|\mathbf{x}_i - \mathbf{b}_c\|^2. \end{aligned} \quad (1)$$

$u_{ci} \in [0, 1]$ is the fuzzy membership degree of sample i in cluster c and follows the probabilistic constraint of $\sum_{c=1}^C u_{ci} = 1$. θ is the fuzzification parameter called ‘fuzzifier’. Membership assignment becomes fuzzy as θ becomes large, while $\theta = 1$ reduces a crisp partition of k -Means.

The optimal cluster center for each cluster with fixed memberships u_{ci} is derived as the membership-weighted mean

vector.

$$\mathbf{b}_c = \frac{\sum_{i=1}^n u_{ci}^\theta \mathbf{x}_i}{\sum_{i=1}^n u_{ci}^\theta}. \quad (2)$$

With fixed cluster prototypes $\mathbf{b}_c$, the optimal memberships satisfying the probabilistic constraint are given as

$$u_{ci} = \left[\sum_{\ell=1}^C \left(\frac{d_{ci}}{d_{\ell i}} \right)^{1/(\theta-1)} \right]^{-1}. \quad (3)$$

In the singular situation where some data samples are on prototypes (zero-distances), the data samples are crisply assigned to the prototypes and have zero-memberships to other clusters. These two phases of cluster prototype estimation and membership estimation are repeated until converge based on the alternate optimization scheme.

FCL replaces the point-type prototype with 1-dimensional lines defined as follows:

$$\text{Line}_c(\mathbf{b}_c, \mathbf{a}_c) = \{x | x = \mathbf{b}_c + \mathbf{a}_c t; t \subseteq R\}. \quad (4)$$

Each line is defined by two vectors of cluster center $\mathbf{b}_c$ and basis vector $\mathbf{a}_c$. The prototypical line of cluster c is to pass through its cluster center and to be spanned by basis vector $\mathbf{a}_c$. The clustering criteria are squared Euclidean distances between sample points and linear prototypes, and the objective function is defined as follows:

$$\begin{aligned} L_{fcl} &= \sum_{c=1}^C \sum_{i=1}^n u_{ci}^\theta D_{ci} \\ &= \sum_{c=1}^C \sum_{i=1}^n u_{ci}^\theta (||\mathbf{x}_i - \mathbf{b}_c||^2 - |\mathbf{a}_c^\top (\mathbf{x}_i - \mathbf{b}_c)|^2). \end{aligned} \quad (5)$$

The optimal memberships u_{ci} with fixed cluster prototypes are calculated by Eq.(3) replacing the clustering criteria with $D_{ci} = ||\mathbf{x}_i - \mathbf{b}_c||^2 - |\mathbf{a}_c^\top (\mathbf{x}_i - \mathbf{b}_c)|^2$. With fixed memberships, optimal prototype estimation is achieved in a 2 step procedure. $\mathbf{b}_c$ is often updated in a same formula with Eq.(2). Next, basis vector estimation is reduced to fuzzy PCA [10], in which the conventional PCA procedure is performed in a fuzzy set where the fuzzy membership degree of each sample is given by u_{ci}^θ . The optimal basis vector of cluster c is the eigen vector corresponding to the largest eigenvalue of the following generalized fuzzy scatter matrix S_{fc} :

$$S_{fc} = \sum_{i=1}^n u_{ci}^\theta (\mathbf{x}_i - \mathbf{b}_c)(\mathbf{x}_i - \mathbf{b}_c)^\top. \quad (6)$$

FCV extended FCL into multi-dimensional case. Using p independent basis vectors $\mathbf{a}_{ck}, k = 1, \dots, p$, p -dimensional linear varieties to be used as the prototype of cluster c is given as

$$\begin{aligned} V_c^p(\mathbf{b}_c; \mathbf{a}_{c1}, \mathbf{a}_{c2}, \dots, \mathbf{a}_{cp}) = \\ \{x | x = \mathbf{b}_c + \sum_{k=1}^p t_k \mathbf{a}_{ck}; t_k \in R\}, \end{aligned} \quad (7)$$

and the clustering criteria are replaced with the distances between data samples and the linear varieties. The membership

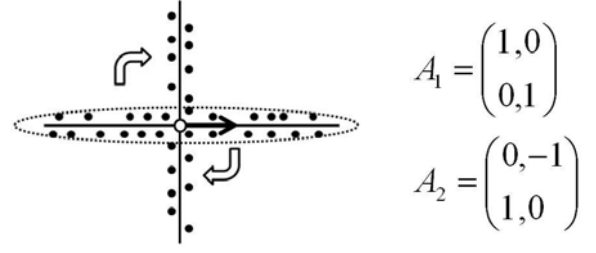


Fig. 1. Definition of prototypical cross-shape and rotation matrix

updating rule is Eq.(3) with the modified criteria. If cluster center $\mathbf{b}_c$ is also updated by Eq.(2), the basis vectors are given by the eigenvectors corresponding to the p largest eigenvalues of S_{fc} .

B. Cross-shape Cluster Extraction in 2-D Data Space

Yoshida *et al.* [16] proposed a technique for capturing the cross-shape substructure by using a single linear prototype. Assume that samples were drawn from a 2-D data space ($m = 2$). The cross-shape prototype of a cluster can be defined by two elements of center point and first normal vector as shown in Fig. 1 where the two lines cross at the center point.

In the model, each cross-shape cluster is further divided into two linear sub-clusters corresponding to the two component lines. The first sub-cluster memberships u_{c1i} represent the degree of belongingness to the first prototypical line whose normal vector is $\mathbf{a}_c$. The second ones u_{c2i} correspond to the second line whose basis vector is $\mathbf{a}_c$. The conventional membership value for cluster c is recovered by $u_{ci} = u_{c1i} + u_{c2i}$. In order to jointly estimate the two memberships of u_{c1i} and u_{c2i} in a single linear clustering-like procedure, vertical rotation of the second sub-cluster around the center point is considered. The local coordinate rotation matrix for the second sub-cluster is given as A_2 in Fig. 1. The distances between sample i and sub-cluster j in cluster c is given as follows:

$$D_{cji} = ||\mathbf{a}_c^\top \mathbf{A}_j (\mathbf{x}_i - \mathbf{b}_c)||^2 \quad (8)$$

The objective function is defined based on the minor component analysis concept [17], in which linear manifolds are formulated by normal vectors.

$$\begin{aligned} L_{cross} &= \sum_{c=1}^C \sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta D_{cji} \\ &= \sum_{c=1}^C \sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta |\mathbf{a}_c^\top \mathbf{A}_j (\mathbf{x}_i - \mathbf{b}_c)|^2, \end{aligned} \quad (9)$$

where the probabilistic constraint for memberships is modified as:

$$\sum_{c=1}^C \sum_{j=1}^2 \sum_{i=1}^n u_{cji} = 1. \quad (10)$$

Fuzzy memberships for sub-clusters are given by

$$u_{cji} = \left[\sum_{\ell=1}^C \left(\frac{\|\mathbf{a}_c^\top \mathbf{A}_j(\mathbf{x}_i - \mathbf{b}_c)\|^2}{\|\mathbf{a}_\ell^\top \mathbf{A}_j(\mathbf{x}_i - \mathbf{b}_\ell)\|^2} \right)^{\frac{1}{\theta-1}} \right]^{-1}. \quad (11)$$

From the necessary condition for the optimality, center points of prototypical cross-shape are updated as

$$\mathbf{b}_c = \frac{\sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta (\mathbf{A}_j^\top \mathbf{a}_c \mathbf{a}_c^\top \mathbf{A}_j \mathbf{x}_i)}{\sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta \mathbf{A}_j^\top \mathbf{a}_c \mathbf{a}_c^\top \mathbf{A}_j}. \quad (12)$$

Normal vectors $\mathbf{a}_c$ is derived in a similar way to minor component analysis so that $\mathbf{a}_c$ corresponds to the smallest eigenvalue of the following modified scatter matrix:

$$\mathbf{S}_{fc} = \sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta (\mathbf{x}_{ci}^j - \mathbf{v}_c)(\mathbf{x}_{ci}^j - \mathbf{v}_c)^\top, \quad (13)$$

where $\mathbf{x}_{ci}^j$ and $\mathbf{v}_c$ are sample vector and cluster center after local coordinate rotation such that

$$\mathbf{x}_{ci}^j = \mathbf{A}_j(\mathbf{x}_i - \mathbf{b}_c), \quad (14)$$

$$\mathbf{v}_c = \frac{\sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta \mathbf{A}_j(\mathbf{x}_i - \mathbf{b}_c)}{\sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta}. \quad (15)$$

Because of coordinate rotation, samples are virtually duplicated and the scatter matrix is calculated using $2n$ objects.

III. HYBRID OF FCV AND CROSS-SHAPE FUZZY CLUSTER EXTRACTION

A. Active Curve Axis Gaussian Mixture Models

Zhang *et al.* [18] proposed a non-linear extension of Gaussian mixture models (GMMs) in order to deal with curve manifolds using a non-linear probability model called active curve axis Gaussian model. The non-linear Gaussian model is generated by bending the normal Gaussian model at the first principal axis in 2-D data space. Then, they modified the mixture models into multi-dimensional data space by assuming that the principal axis curves are located in 2-D planes determined by the first two eigenvectors. The modified likelihood function was defined by combining the two measures: deviations from 2-D planes and distances from principal axis on the 2-D planes. The model parameters are estimated by a generalized Expectation-Maximization (EM) algorithm.

In this paper, a similar idea of combining two criteria is performed in order to capture cross-shape clusters in multiple-dimensional data space.

B. Combination of Two Criteria and Generalized Algorithm

In this paper, FCV is combined with cross-shape cluster extraction in 2-D plane. Assume that we have n samples with m dimensional observation $\mathbf{x}_i = (x_{i1}, \dots, x_{im})$, $i = 1, \dots, n$ and the cluster prototypes are given by cross-shape composed of two lines on 2-D plane. In the FCV part, the prototypical 2-D plane in cluster c is defined by cluster center $\mathbf{q}_c$ and

basis vectors $\mathbf{p}_{c1}$ and $\mathbf{p}_{c2}$. Then, the FCV criterion of distance between samples $\mathbf{x}_i$ and the 2-D plane is calculated as:

$$d_{ci}^* = \|\mathbf{x}_i - \mathbf{q}_c\|^2 - \sum_{k=1}^2 |(\mathbf{x}_i - \mathbf{q}_c)^\top \mathbf{p}_{ck}|^2. \quad (16)$$

Next, in order to consider the cross-shape cluster extraction in the 2-D plane, samples are projected on the 2-D plane as $\mathbf{y}_i = (y_{i1}, y_{i2})^\top$:

$$\mathbf{y}_{ij} = \mathbf{p}_{cj}^\top (\mathbf{x}_i - \mathbf{q}_c), \quad j = 1, 2. \quad (17)$$

The cross-shape prototype on the 2-D plane is defined by cluster center (intersection of two lines) $\mathbf{b}_c$ and first normal vector $\mathbf{a}_c$. Then, the criterion of distance between the projected sample vector $\mathbf{y}_i$ and the prototypical cross is calculated as:

$$d_{ci}^{**} = \|\mathbf{a}_c^\top \mathbf{A}_j(\mathbf{y}_i - \mathbf{b}_c)\|^2. \quad (18)$$

using the two criteria, the combined objective function is defined as:

$$\begin{aligned} L &= \sum_{c=1}^C \sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta (d_{ci}^* + d_{ci}^{**}) \\ &= \sum_{c=1}^C \sum_{j=1}^2 \sum_{i=1}^n u_{cji}^\theta \left(\|\mathbf{x}_i - \mathbf{q}_c\|^2 - \sum_{k=1}^2 |(\mathbf{x}_i - \mathbf{q}_c)^\top \mathbf{p}_{ck}|^2 + \|\mathbf{a}_c^\top \mathbf{A}_j(\mathbf{y}_i - \mathbf{b}_c)\|^2 \right). \end{aligned} \quad (19)$$

The iterative algorithm is composed of three steps. In the first step, prototypical 2-D planes are estimated. Cluster centers $\mathbf{q}_c$ are calculated by Eq.(2) where $u_{ci}^\theta = u_{c1i}^\theta + u_{c2i}^\theta$. Basis vectors $\mathbf{p}_{ck}$ are given by the k -th principal eigenvector of Eq.(6) replacing $\mathbf{b}_c$ with $\mathbf{q}_c$.

In the second step, cross-shape prototypes are estimated on 2-D planes. Intersection vectors $\mathbf{b}_c$ are calculated by Eq.(12) replacing $\mathbf{x}_i$ with $\mathbf{y}_i$. Normal vectors $\mathbf{a}_c$ are given by the eigenvectors corresponding to the smallest eigenvalues of Eq.(13) replacing $\mathbf{x}_i$ with $\mathbf{y}_i$.

Then, in the third step, fuzzy memberships for sub-clusters u_{cji} are calculated as:

$$u_{cji} = \left[\sum_{\ell=1}^C \left(\frac{d_{ci}^* + d_{ci}^{**}}{d_{\ell i}^* + d_{\ell i}^{**}} \right)^{\frac{1}{\theta-1}} \right]^{-1}. \quad (20)$$

IV. NUMERICAL EXAMPLE

A numerical experiment was performed with an artificially generated data set composed of 80 samples. Figure 2 shows the 3-D plots of the data set including two separate cross-shape clusters. The two cross-shape clusters exist on two 2-D planes, which are mutually orthogonal as shown in Fig. 3, and the intersection areas of the two crosses have small distances for both two planes. The goal of the experiment is to capture the two cross-shape sub-structures by partitioning samples into two clusters ($C = 2$). The fuzzification parameter was set as $\theta = 2$. Figure 4 shows the clustering result derived by FCV

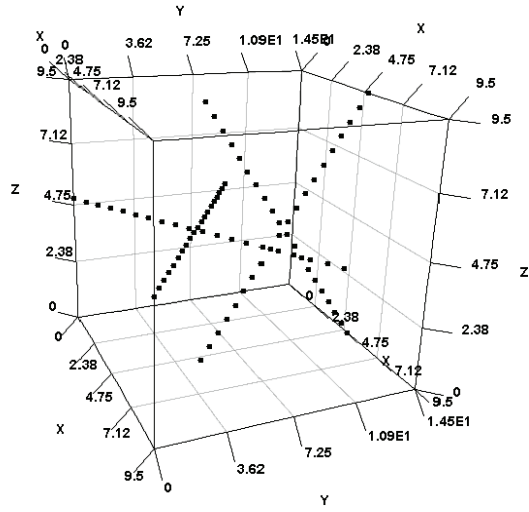


Fig. 2. 3-D plots of the data sets

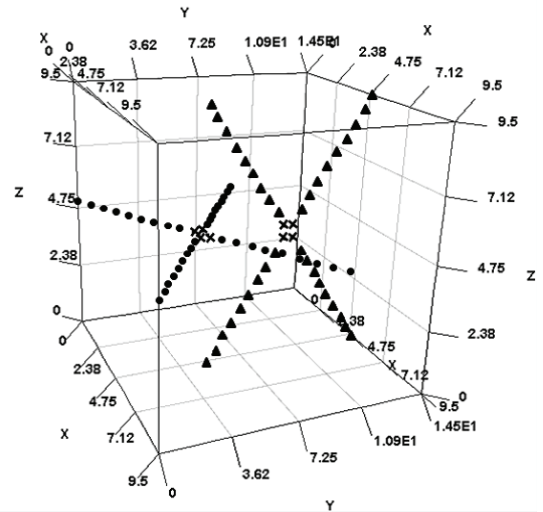


Fig. 4. Clustering result of FCV

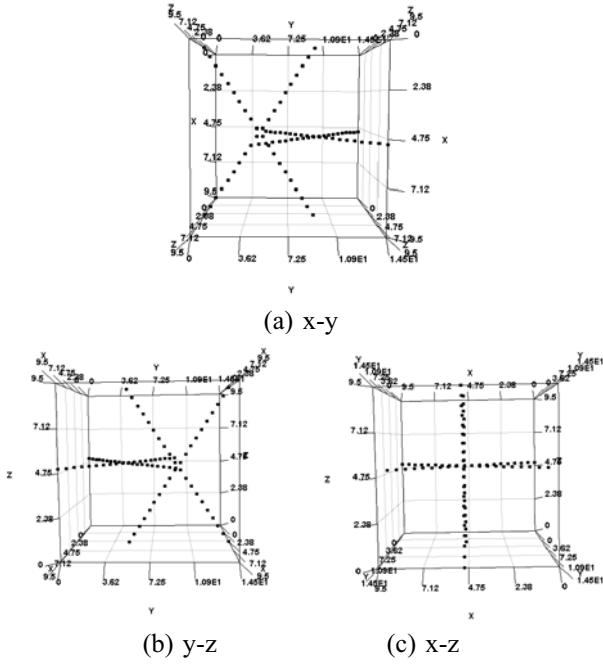


Fig. 3. Other views of Fig. 2

where circles and triangles indicate the maximum membership classification and times indicate the samples having very fuzzy memberships of $u_{ci} \in [0.4, 0.6]$. FCV is useful for capturing 2-D planes, on which cross-shape clusters exist, while it cannot reveal cross-shape substructure. Figure 5 shows the result by the proposed method where two cross-shape clusters were properly extracted. The intersection areas were also properly assigned to clusters by considering both two criteria.

Table I compares the membership degrees of the eight data points, which are described by $\times$ in the intersection of two cross-shapes in Fig. 4. The table indicates that the proposed hybrid approach made it possible to clarify the cross-shape

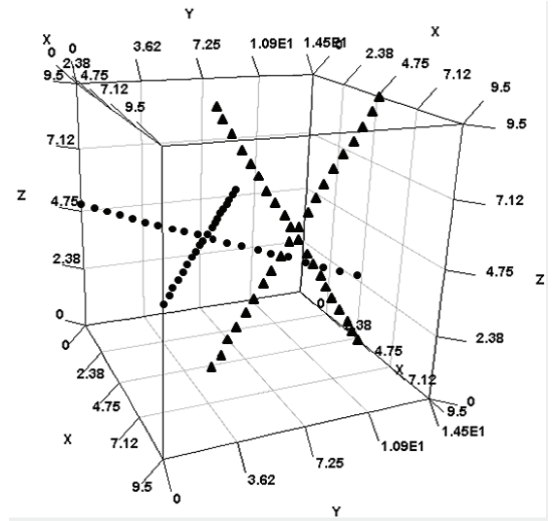


Fig. 5. Clustering result of proposed method

cluster partitions by assigning more crisp memberships. In this way, the proposed hybrid objective function contributes to extraction of cross-shape substructures in multi-dimensional data space.

V. CONCLUSION

This paper proposed a hybrid objective function of FCV and cross-shape cluster extraction for extracting cross-shape substructures in multi-dimensional data space. The FCV part is responsible for capturing 2-D planes, on which cross-shape substructures exist, while cross-shape cluster extraction part performs linear sub-space learning with local coordinate rotation. Fuzzy memberships are estimated considering both criteria.

A possible future work is hybridization with FCM in such a way with Fuzzy *c*-Elliptotypes (FCE) [7] for avoiding the size problem of FCL-type linear clustering models. Another

data point		FCV		proposed method	
		Cluster1	Cluster2	Cluster1	Cluster2
Cross1	1	0.535	0.465	0.943	0.057
	2	0.550	0.450	0.947	0.053
	3	0.574	0.426	0.955	0.045
	4	0.625	0.375	0.977	0.023
Cross2	1	0.474	0.526	0.023	0.977
	2	0.465	0.535	0.045	0.955
	3	0.450	0.550	0.053	0.947
	4	0.426	0.574	0.057	0.943

TABLE I
COMPARISON OF MEMBERSHIP DEGREES IN INTERSECTION OF TWO
CROSSES

potential future work is related to other shape recognition tasks. By changing rotation angles, the proposed model has ability for handling other shape prototypes.

REFERENCES

- [1] F. Höppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis*, Jhon Wiley & Sons, 1999.
- [2] J. C. Bezdek, J. Keller, R. Krishnapuram and N. R. Pal, *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Springer, 2005.
- [3] S. Miyamoto, H. Ichihashi and K. Honda, *Algorithms for Fuzzy Clustering*, Springer, 2008.
- [4] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, 1981.
- [5] J. B. MacQueen, "Some methods of classification and analysis of multivariate observations," *Proc. of 5th Berkeley Symposium on Math. Stat. and Prob.*, pp. 281-297, 1967.
- [6] J. C. Bezdek, C. Coray, R. Gunderson and J. Watson, "Detection and characterization of cluster substructure I. Linear structure fuzzy c -lines," *SIAM Jour. of Appl. Math.*, vol.40, no.2, pp.339-357, 1981.
- [7] J. C. Bezdek, C. Coray, R. Gunderson, and J. Watson, "Detection and characterization of cluster substructure 2. Fuzzy c -Varieties and convex combinations thereof," *SIAM J. Appl. Math.*, vol.40, no.2, pp.358-372, 1981.
- [8] N. Kambhatla and T. K. Leen, "Dimension reduction by local principal component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1493-1516, 1997.
- [9] G. E. Hinton, P. Dayan, and M. Revow, "Modeling the manifolds of images of handwritten digits," *IEEE Trans. on Neural Networks*, vol. 8 no. 1, pp. 65-74, 1997.
- [10] Y. Yabuuchi and J. Watada, "Fuzzy principal component analysis and its application," *Biomedical Fuzzy and Human Sciences*, vol.3, pp.83-92, 1997.
- [11] K. Honda and H. Ichihashi, "Linear fuzzy clustering techniques with missing values and their application to local principal component analysis," *IEEE Trans. on Fuzzy Systems*, vol.12, no.2, pp.183-193, 2004.
- [12] K. Honda and H. Ichihashi, "Regularized linear fuzzy clustering and probabilistic PCA mixture models," *IEEE Trans. on Fuzzy Systems*, vol. 13, no. 4, pp. 508-516, 2005.
- [13] R. Krishnapuram, O. Nasraoui and H. Frigui, "The fuzzy c spherical shells algorithms: a new approach," *IEEE Trans. on Neural Networks*, vol.3, no.5, pp.663-671, 1992.
- [14] R. Krishnapuram, H. Frigui and O. Nasraoui, "Fuzzy and possibilistic shell clustering algorithms and their application to boundary detection and surface approximation - Part I," *IEEE Trans. on Fuzzy systems*, vol.3, no.1, pp.29-43, 1995.
- [15] F. Höppner, "Fuzzy shell clustering algorithms in image processing: fuzzy c -rectangular and 2-rectangular shells," *IEEE Trans. on Fuzzy Systems*, vol.5, no.4, pp. 599-613, 1997.
- [16] D. Yoshida, K. Honda, A. Notsu and H. Ichihashi, "Local coordinate rotation for cross-shape cluster extraction in linear fuzzy clustering," *Proc. of Joint 5th International Conference on Soft Computing and Intelligent Systems and 11th International Symposium on Advanced Intelligent Systems*, pp.239-242, 2010.
- [17] E. Oja, "Principal components, minor components and linear neural networks," *Neural Networks*, vol.5, 1992, pp.927-935.
- [18] B. Zhang, C. Zhang and X.Yi, "Active curve axis gaussian mixture models," *Pattern Recognition*, vol.38, pp.2351-2362, 2005.