|  |  |  |
| --- | --- | --- |
|  | Available online at www.sciencedirect.com |  |
| **ScienceDirect** |
| AASRI Procedia 4 ( 2013 ) 50 – 56 |
| 2013 AAS SRI Confere ence on Inte elligent Syst tems and Co ontrol  Co ompariso on of Di istance M Measure s on Fuz zzyc-Me eans Alg orithm  for r Image Classific cation P roblem  Jiho H Hana, Dong g-Chul Park ka\*, Dong--Min Wooa a, and Soo--Young Mi inb  *aDept. of E Electronics Eng., M Myongji Univ.,Gy Gyeongi-do, 449-7 728, Rep. of KOR REA*  *bKo rea Electronics T Tech. Inst.,SongN Nam, 463-816, Rep p. of KOREA* | | |

**Abst tract**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| A stu udy on the use o of two differen nt distance mea sures, Euclidea an distance and | | | | divergence dist tance, for FCM M is conducted | |
| for a an image classi fication proble m in this pape er.Conventional l FCM algorith hm which uses | | | | | Euclidean dista ance measure |
| utiliz zes only mean i information fro om an image blo ock for its featu ure while FCM M algorithm with h divergence ut tilizes both of | | | | | |
| varia ance and mean | information. E Evaluationson a a set of Caltech h databaseshow w that Divergen nce-based FCM M gives higher | | | | |
| accur racy when com mpared with som me conventional l algorithms wit th Euclidean di istance. | | | | | |
| © 2013 The Authors. Published by Elsevier B.V.© 20 013.Published d by Elsevier B B.V. | | | Open access under [CC BY-NC-ND license.](http://creativecommons.org/licenses/by-nc-nd/3.0/) | | |
| Selection and/or peer review under responsibility of American Applied Science Research Institute earch Institute | | | | | |
| *Keyw workds*: Fuzzy c-M Means, Classifica | | tion, Divergence | Measure, SOM | | |
| **1.In ntroduction** | | | | | |

Rece ent increase u se of compres ssed image da ata requires for r an automatic c tool thatcan retrieve imag e data

|  |
| --- |
| effic ciently. Severa al conventiona al algorithms adopted for th his purpose inc cludesFuzzy c c-Means (FCM M)  algo orithm andSelf f OrganizingM Map(SOM) [1] ][2].FCM algo orithm is the m most widely u used one and c can be  thou ught as an imp proved version n of earlier clu ustering. The FCM algorith hm shows mor re robustnessw when |

\* Corresponding a author. Tel.: +82-3 31-330-6975; fax x: +82-31-330-69 977

*E-mail address: pa arkd@dreamwiz. com*

2212-6716 © 2013 The Authors. Published by Elsevier B.V. Open access under [CC BY-NC-ND license.](http://creativecommons.org/licenses/by-nc-nd/3.0/)

Selection and/or peer review under responsibility of American Applied Science Research Institute

doi: 10.1016/j.aasri.2013.10.009

*Jiho Han et al. / AASRI Procedia 4 ( 2013 ) 50 – 56*  51

compared with SOM and k-Means algorithm. By combining the ideas of FCM and divergence measure for the problem of image classification, we expect a better feature extraction procedure for more accurate   
classification scheme. The method adopts the FCM algorithm with Divergence Measure [3]-[5]for acquiring texture information from image data.

Section 2 summarizes FCM algorithms with two different distance measures: Euclidean distance and divergence distance. Section 3 reports the evaluations and a comparison of different schemes through experiments. Section 4 presents conclusions.

**2.Adopted Clustering Algorithms**

*2.1.Fuzzy c-MeansAlgorithm*

The following equation is used as the objective function for FCM[2] :

 (1)

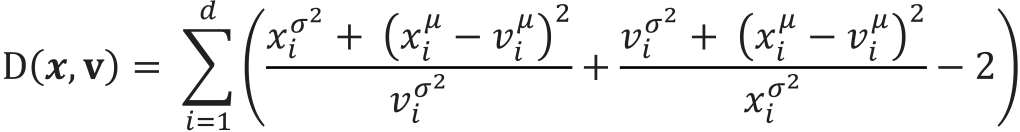
where*di*(*xk*) ,*ki,m, n,* and *c* follow the definitions in [2].   
From Eq. (1), Bezdekfinds the following equations [2]:

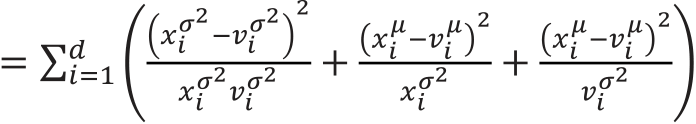
|  |  |
| --- | --- |
|  | (2)  (3) |

*2.2.FCM with Divergence Measure*

The choice of different distance measure can affect the performance of clustering results [1]. This leads to

|  |  |  |  |
| --- | --- | --- | --- |
| the idea of divergence measure for image classification problem. | and | *, i* = 1*,* | *, d* , the |
| Given two Gaussian Probability Density Functions, |
| divergence distance can be defined as follows[6,8]: |



 (4)



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 52 | *Jiho Han et al. / AASRI Procedia 4 ( 2013 ) 50 – 56* | | | | |
| where | *,* | *,* | and | follow the definition in [6,8]. Note that the divergence measure is also called as |
| *Kullback-Leibler Divergence.* | | | | |

FCM calculates the parameters of center and membership values by applying all the data at once in batch mode. However, the D-FCM used in the proposed classifier calculates and updates their parameters at each application of each data vector as the Gradient-based FCM [3]. The advantage of this iterative application and updating the center parameters was reported in [3]. When each data vector is presented to the network, the following can be also found :

|  |  |
| --- | --- |
|  | (5) |

Subsequently, the following update equation for the mean and variance was derived[6]:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | |  | follow the definitions in [6]. | (6) |
|  | | | | | (7) |
| where | , | , | , | , and |

**3.Experiments and Results**

For evaluation and comparison purpose for different classification schemes, we performed some experiments on a set of the Caltech image database[7]. The Caltech image database has been widely used as a benchmark data for various image classification problems. Examples of the four-category data used in experiments are given in Fig. 1.200 data are used for each category. Preparation for experiments follows the procedure used in [4]: 150 randomly selected training data and 50 remaining data for each class are used for the evaluation of different classifier schemes. Note that all the image data are converted to grey images with identical resolution.

The first step for designing a data classifier is the feature extraction process. The local feature extraction adopted in our experiments allows us to find localized feature information from entire image space.By collecting thefeature information from local points, each image data can be expressed in terms of features. In our experiments, 8 8 blocks are used for extracting localized information of an image data. The local information in each image window can be extracted by using conventional feature extraction methods [9]-[11]. Among them, Gabor filters and wavelet filters show reasonable performance except requiring extensive computational burden. This computational burden keeps these methods from being used for feature extraction tools for several applications. On the other hand, the Discrete Cosine Transform (DCT) is more suitable to acquire frequency information in images and DCT is used for our experiments because DCT does not require so much computational effort. Even though DCT coefficients out of each block image produces 64-dimensional one, the 32 lower frequency part of the 64 coefficients are used [4] and this yields a total 32-dimensional vector value for each local block image.

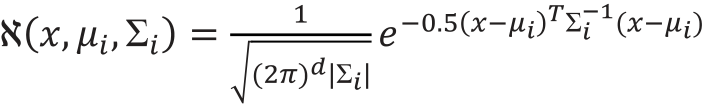
The model for a data category is formed by deciding the code vectors using clustering algorithms on its feature vectors. The modeling or training process consists of finding representativecode vector for

*Jiho Han et al. / AASRI Procedia 4 ( 2013 ) 50 – 56*  53

eachcategory with its feature vectors.Bayesian classifier is utilized for the performance evaluation on various classification schemes. Itallocates a class for a given testdata according to the following probability calculation [4]:

 (8)

 (9)

 (10)

whereM,*ci*,d (=32), and *i* follow the definitions in [4].





(a) Car (b) Motorbike



(c) Bike (d) Airplane

Fig. 1 Examples of data



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 54 | *Jiho Han et al. / AASRI Procedia 4 ( 2013 ) 50 – 56*  The same train ning and test p procedures for r the same dat ta sets are per rformed on tw wo conventiona al clustering | | | | | | | | |
| algo orithms, SOM | and FCM, an nd compared w with the accur racy of the pro oposed algorit thm. During e experiments, | | | | | | | |
| seve eral numbers o of code vecto rs up to 19 ar re used for ea ach algorithm. . Fig. 2 summ marizes their c classification | | | | | | | | |
| accu uracies. The av verage accura acies over diff ferent number rs of code vec ctors (3 to 19) ) for SOM, F | | | | | | | CM, and D- | |
| FCM M are observed d as 42.22%, 4 44.61%,and 7 73.33%, respec ctively. The c classification a accuracy is inc creased with | | | | | | | | |
| code e vector numb bers. Note th hat their class sification accu uracies are so omewhat satu urated when | | | | | | | code vector | |
| num mbers reach 1 5. Note also | | | that SOM a and FCM don n’t use the v variance while e D-FCM do oes. Table 1 | | | | | |
| summ marizes the c classification p performance f for each categ gory of image e data when 1 15 code vecto ors are used. | | | | | | | | |
| From m Fig. 2 and d Table 1, w we can notice | | | | that the cova ariance value | | utilized in D D-FCM along g with mean | | |
| infor rmation plays s an important t role in impr roving the cla ssification acc curacy for im mage data. The e divergence | | | | | | | | |
| dista ance measure | allows the im mage classifie er to have an | | | | advantage ov ver the Euclid dean distance | | | and similar |
| resu lts are reporte ed in [4]. Tabl le 2 shows a c confusion mat trix of D-FCM M based image e classifier for r the case of | | | | | | | | |
| 15 c ode vectors. C Car images ar | | e classified w well enough wh hile motorbike e data are con nfused with ai irplane data. | | | | | | |
| Som me other featur re extraction | | methods that | | can discrimin nate car data | | from airplane e data need to o be done in | | |
| futur re research. F Future resear rch also inclu ude the comp parison of F CM and Gra adient-based F FCM under  dive ergence measu ure environme ent. | | | | | | | | |

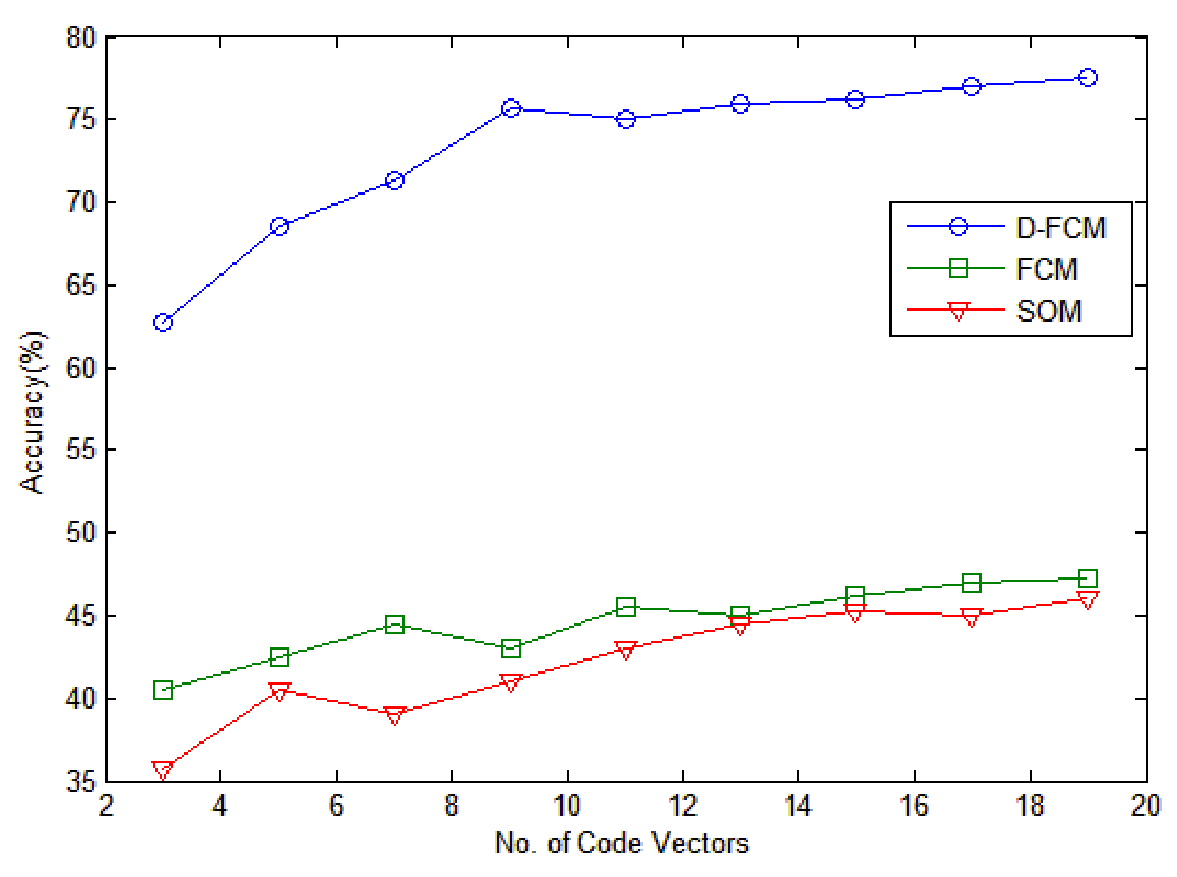


Fig. 2 2. Comparison of f classification acc curacies ofvariou us classifier schem mes over number rs of code vectors s

Table e 1. Classification n performance of SOM, FCM, and d D-FCM for the c case of 15 code v vectors

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Car | Bike | Mo otorbike | Airpl lane | Avg g. |
| Accurac cy(%) |
| SOM | 98 | 63 | 7 | 13 3 | 45.2 25 |
| FCM | 97 | 66 | 5 | 17 7 | 46.2 25 |
| D-FCM | 100 | 78 | 51 | 76 6 | 76.2 25 |

*Jiho Han et al. / AASRI Procedia 4 ( 2013 ) 50 – 56*  55

Table 2. Detailed classification results of D-FCM

*Classified*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Car | Bike | Motorbike | Airplane |
| *Inputs* | Car | 0 | 100 | 0 | 0 |
| Bike | 2 | 78 | 12 | 8 |
| Motorbike | 0 | 6 | 51 | 43 |
| Airplane | 7 | 2 | 15 | 76 |

**4.Conclusions**

A comparative study Euclidean distance and divergence distance for FCM is conducted in this paper.When used for an image classification problem, conventional FCM algorithm which uses Euclidean distance measure utilizes only mean information from an image block for its feature. On the other hand, FCM algorithm with divergence utilizes both of variance and mean information from an image block for its features. Self-Organizing Map is also used in our experiments as a baseline method for the performance comparison purpose. Experiments on Caltech database are performed for a four category problem. Classification accuracies from different classification schemes imply that divergence-based FCM is a better fit for image classification problem than the conventional FCM or SOM that use only the mean information as their features for image data. We can conclude that the divergence information of image data plays an important role in image classification problem and the resultant classification scheme that uses the divergence information as well as the mean information from image data is a better choice for image classification problem. However, a further effort should be given to find a feature extraction method for overcoming the confusion between Motorbike data and Airplane data.

**Acknowledgements**

This work was supported byIT R&D program of The MKE/KEIT (10040191) and by National Research Foundation of Korea Grantfunded by the Korean Government (2010-0009655).

**References**

[1]Bezdek, J.C. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum;1981.

[2]Kohonen,T.The Self-Organizing Map. Proceedings of IEEE, 1990;78 :1464-1480.

[3]Park D.C.et. al. Classification of Audio Signals Using Gradient-Based Fuzzy c-Means Algorithm with Divergence Measure. Proc. Pacific Rim Conference on Multimedia.2005; 698-708.

[4]Park D.C., Woo D.M. Image classification using Gradient-Based Fuzzy c-Means with Divergence Measure. Proc. Int. Joint Conference on Neural Networks.2008; 2520-2524.

[5]Park D.C. Satellite Image Classification Using a Divergence-Based Fuzzy c-Means Algorithm.Proc.Int. Conf. on Image and Signal Processing. 2012;555-561.

[6]Park D.C. et. al.Centroid Neural NetworkWith a Divergence Measurefor GPDF Data Clustering.IEEE

56  *Jiho Han et al. / AASRI Procedia 4 ( 2013 ) 50 – 56*

Transactions on Neural Networks. 2008;19(6): 948-957.

[7]http://www.vision.caltech.edu/html-files/archive.html   
[8]FukunagaK. Introduction to Statistical Pattern Recognition, Academic Press Inc.;1990.

[9]Daugman J.G. Complete Discrete 2D Gabor Transform by Neural Networks forImage Analysis and Compression. IEEE Transactions on Acoustics, Speech, and SignalProcessing.1988;36:1169-11179.

[10]Pun C.M., LeeM.C. Extraction of Shift Invariant Wavelet Features for Classification of Images with Different Sizes. IEEE Transactions on Pattern Analysis andMachine Intelligence. 2004;26(9) :1228-1233. [11]Huang Y.L., Chang R.F. Texture Features for DCT-Coded Image Retrieval andClassification. Proc. Int. Conf. on Acoustics, Speech, and SignalProcessing.1999;6:3013-3016.