Monograph:
Zheng, G.L. and Billings, S.A. (1997) An Enhanced Sequential Fuzzy Clustering Algorithm. Research Report. ACSE Research Report 674 . Department of Automatic Control and Systems Engineering

An Enhanced Sequential Fuzzy Clustering Algorithm

G. L. Zheng and S. A. Billings
Department of Automatic Control and Systems Engineering,
University of Sheffield, Mappin Street, Sheffield S1 3JD

Abstract — A sequential fuzzy clustering algorithm is proposed based on a modification to the objective function used in the fuzzy competitive learning algorithm. The new learning algorithm can be used to enhance the excitation on the non-winning centroids and to reduce the excitation on the winning centroid when the fuzziness parameter is close to unity. The excitation on the winning centroid can be further reduced when the input pattern is far away from the winning centroid. An excitation-inhibition mechanism can also be introduced into the learning such that the non-winning centroids move towards the input pattern while the winning centroid moves away from the input pattern when the winning centroid is far away from the input pattern. The new algorithm overcomes the problem of under utilization of centroids found in the k-means or related clustering algorithms and in the fuzzy competitive learning algorithm when the fuzziness parameter is close to unity. The performance of the new algorithm is demonstrated on the IRIS data set.

Keywords — Fuzzy Clustering, K-means, Sequential Clustering

Research Report No. 674

April 1997

An Enhanced Sequential Fuzzy Clustering Algorithm

G. L. Zheng and S. A. Billings
Department of Automatic Control and Systems Engineering,
University of Sheffield, Mappin Street, Sheffield S1 3JD

February 1997

Abstract — A sequential fuzzy clustering algorithm is proposed based on a modification to the objective function used in the fuzzy competitive learning algorithm. The new learning algorithm can be used to enhance the excitation on the non-winning centroids and to reduce the excitation on the winning centroid when the fuzziness parameter is close to unity. The excitation on the winning centroid can be further reduced when the input pattern is far away from the winning centroid. An excitation-inhibition mechanism can also be introduced into the learning such that the non-winning centroids move towards the input pattern while the winning centroid moves away from the input pattern when the winning centroid is far away from the input pattern. The new algorithm overcomes the problem of under utilization of centroids found in the k-means or related clustering algorithms and in the fuzzy competitive learning algorithm when the fuzziness parameter is close to unity. The performance of the new algorithm is demonstrated on the IRIS data set.

Keywords — Fuzzy Clustering, K-means, Sequential Clustering

1 Introduction

Clustering techniques are often used to organize unlabeled data samples into clusters such that the similarity among samples within a cluster and the dissimilarity among samples belonging to different clusters are maximized. Many classical clustering methods are presented in the texts by Kohonen (Kohonen, 1989), Everitt (Everitt, 1993), Anderberg (Anderberg, 1973) and others. Among these methods, the sequential k-means and related methods have been extensively studied in recent years due to the widespread interest in neural networks. One serious problem with the sequential k-means or related algorithms is that some initial centroids may get stuck in regions with few or no input data samples. Several algorithms were proposed to deal with this problem, including the frequency sensitive competitive learning algorithm (Ahalt et al., 1990), maximum-entropy clustering (Rose et al., 1990), the neural-gas network (Martinetz et al., 1993), generalized clustering networks (Pal et al., 1993), the fuzzy competitive learning algorithm (Chung and Lee, 1994) and the optimal adaptive k-means algorithm (Chinrungrueng and Sequin, 1995). All these methods consider not only the relation between the input pattern and the winning centroid but also the structural information presented in all the centroids and their relations to the input pattern. These algorithms either move all the centroids towards the input pattern according to their distances from the input pattern or they move the winning centroid, but the winner is determined in the sence of a modified distance measure. As a result, all the centroids are updated and the problem of under utilization may be eliminated.

In the present study, a modification to the fuzzy competitive learning (FCL) algorithm (Chung and Lee, 1994) is proposed. We call this algorithm enhanced sequential fuzzy clustering (ESFC). A new objective function is proposed by introducing a non-unity weighting on the winning centroid. The new algorithm achieves better clustering performance when the fuzziness parameter is small. An excitation-inhibition behaviour can also be introduced into the learning process.

The layout of the paper is as follows. The fuzzy competitive learning algorithm is presented in Section 2 and the disadvantages of this are explained. An enhanced sequential fuzzy clustering algorithm is then derived. Comparetive studies of the two algorithms when applied to the IRIS data are reported in Section 3 and conclusions are given in Section 4.

2 Enhanced Sequential Fuzzy Clustering

The fuzzy competitive learning algorithm (Chung and Lee, 1994) incorporates the fuzzy membership function derived in the fuzzy k-means (Bezdek, 1981) algorithm into the sequential k-means algorithm. At each step, all the centroids are moved towards the input pattern. If the fuzziness parameter in the algorithm is selected appropriately, the problem of under utilization associated with the sequential k-means algorithm may be eliminated. The algorithm is given as follows.

1. Store unlabeled input patterns $\mathbf{X} = \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in R^p$, where n is the number of input patterns and p is the dimension of the input patterns.

2. Set the number of centroids k, $1 < k < n$. Set the fuzziness parameter m, $1 < m < \infty$. Set the learning rate $0 < \alpha_0 < 1$. Set the maximum number of iterations T and the initial centroids $\mathbf{c}_{1,0}, \mathbf{c}_{2,0}, \ldots, \mathbf{c}_{k,0} \in R^p$.

3. For t = 1 to T, set $\alpha_t = \alpha_0(1 - t/T)$

   (a) Choose randomly an input pattern $\mathbf{x}_i \in \mathbf{X}$.

   (b) For l = 1 to k

$$\begin{aligned} \mu_{l,i,t} &= \left(\sum_{j=1}^{k} \left(\frac{||\mathbf{x}_i - \mathbf{c}_{l,t-1}||^2}{||\mathbf{x}_i - \mathbf{c}_{j,t-1}||^2} \right)^{\frac{2}{(m-1)}} \right)^{-1} \\ \mathbf{c}_{l,t} &= \mathbf{c}_{l,t-1} + \alpha_t \mu_{l,i,t}^m (\mathbf{x}_i - \mathbf{c}_{l,t-1}) \end{aligned} \tag{1}$$

   (c) Next l

4. Next t.

Assume $\mathbf{x} \in R^p$ is a stochastic input vector with a time invariant probability distribution $f(\mathbf{x})$ and $\mathbf{X} = \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in R^p$ is a set of samples drawn from $f(\mathbf{x})$ at $t = 1, 2, \ldots, n$. The centroids obtained by the fuzzy competitive learning algorithm are given by the stochastic approximate minimization of the following objective function

$$\mathbf{J} = \sum_{i=1}^{n} \sum_{l=1}^{k} \mu_{l,i}^{m} ||\mathbf{x}_i - \mathbf{c}_l||^2 \tag{2}$$

where

$$\mu_{l,i} = \left(\sum_{j=1}^{k} \left(\frac{||\mathbf{x}_i - \mathbf{c}_l||^2}{||\mathbf{x}_i - \mathbf{c}_j||^2} \right)^{\frac{2}{(m-1)}} \right)^{-1} \tag{3}$$

It is well known that the fuzzy membership function $\mu_{l,i}$ tends to unity and $1/k$ when the fuzziness parameter m tends to one and ∞ respectively. When the fuzziness parameter m tends to one, the algorithm becomes the sequential k-means algorithm. Therefore, the algorithm may still under utilize the centroids when the fuzziness parameter is close to one. When the m parameter tends to ∞, all the centroids will move towards the input pattern at the same rate and all the centroids will converge to the grand mean of the data samples. The fuzziness parameter m is critical to the performance of the fuzzy competitive learning algorithm. To overcome the problem of under utilization, larger values may be used. However, the clustering results may deteriorate because several centroids may move at similar rates. To avoid this, Chung and Lee (1994) proposed that the fuzziness parameter should be monotonically decreased during learning. However, this will considerably increase the learning time.

When the fuzziness parameter m is close to unity, the fuzzy membership functions of the centroids which are far away from the input patterns become smaller. At the same time the membership functions of the centroids which are closer to the input patterns become larger as more data are presented to the inputs. The learning algorithm therefore tends to distribute too much excitation but no restriction or inhibition on the winner while putting too much restriction and too little excitation on the loser.

To overcome this problem, we propose the introduction of a weighting parameter to balance the excitation inhibition in the learning algorithm.

$$\mathbf{J_e} = \sum_{i=1}^{n} \left(\sum_{l=1, l \neq r}^{k} \mu_{l,i}^{m} ||\mathbf{x}_i - \mathbf{c}_l||^2 + \beta \mu_{r,i}^{m} ||\mathbf{x}_i - \mathbf{c}_r||^2 \right) \tag{4}$$

where $\mathbf{c}_r$ is the winning centroid. If there is a tie break, the centroid with the lowest index will be the winner

$$\mu_{l,i} = \left(\sum_{j=1}^{k} \left(\frac{||\mathbf{x}_i - \mathbf{c}_l||^2}{||\mathbf{x}_i - \mathbf{c}_j||^2} \right)^{\frac{2}{(m-1)}} \right)^{-1} \tag{5}$$

A sample function of the objective function $\mathbf{J_e}$ is

$$\mathbf{J_x} = \sum_{l=1, l \neq r}^{k} \mu_{l,i}^{m} ||\mathbf{x}_i - \mathbf{c}_l||^2 + \beta \mu_{r,i}^{m} ||\mathbf{x}_i - \mathbf{c}_r||^2 \tag{6}$$

The objective function $\mathbf{J_e}$ may be approximately minimized by moving the centroids in the direction of the negative gradient of the function $\mathbf{J_X}$. The gradient of the sample function is

$$(\mathbf{J_X})'_{\mathbf{c}_r} = -2(\mathbf{x}_i - \mathbf{c}_r)\mu_{r,i}^m \left\{ 1 + \frac{(\beta - 1)(m\mu_{r,i} - 1)}{m - 1} \right\} \tag{7}$$

$$(\mathbf{J_X})'_{\mathbf{c}_l} = -2(\mathbf{x}_i - \mathbf{c}_l)\mu_{l,i}^m \left\{ 1 + \frac{(\beta - 1)m\mu_{r,i}}{m - 1} \right\} \qquad l \neq r \tag{8}$$

A new learning algorithm can then be obtained by replacing step 3 in the fuzzy competitive learning algorithm by the following.

3. For t = 1 to T, set $\alpha_t = \alpha_0(1 - t/T)$
   (a) Choose randomly an input pattern $\mathbf{x}_i \in \mathbf{X}$.
   (b) For l = 1 to k

$$\begin{aligned}
\mu_{l,i,t} &= \left(\sum_{j=1}^{k} \left(\frac{||\mathbf{x}_i - \mathbf{c}_{l,t-1}||^2}{||\mathbf{x}_i - \mathbf{c}_{j,t-1}||^2} \right)^{\frac{2}{(m-1)}} \right)^{-1} \\
\mathbf{c}_{r,t} &= \mathbf{c}_{r,t-1} + \alpha_t \mu_{r,i,t}^m \left\{ 1 + \frac{(\beta - 1)(m\mu_{r,i,t} - 1)}{m - 1} \right\} (\mathbf{x}_i - \mathbf{c}_{r,t-1}) \\
\mathbf{c}_{l,t} &= \mathbf{c}_{l,t-1} + \alpha_t \mu_{l,i,t}^m \left\{ 1 + \frac{(\beta - 1)m\mu_{r,i,t}}{m - 1} \right\} (\mathbf{x}_i - \mathbf{c}_{l,t-1}) \qquad l \neq r
\end{aligned} \tag{9}$$

   (c) Next l

This algorithm will be refered to as the enchanced sequential fuzzy clustering (ESFC) routine. Note that the ESFC algorithm becomes the fuzzy competitive learning (FCL) algorithm when the weighting parameter β is unity. When the fuzziness parameter is close to unity, the weighting parameter β should be larger than unity. This has the effect of increasing the learning rate for both the winning and non-winning centroids when $m\mu_{r,i} > 1$. The increament on the learning rate for the winning centroids is smaller than for the non-winning centroids. Therefore, the weighting has the effect of enhancing the excitation on the non-winning centroids and reducing the excitation on the winning centroid. If $m\mu_{r,i} < 1$, the excitation on the winning centroid is further reduced. For larger weighting, the winning centroid may even move away from the input pattern while the non-winning centroids always move towards the input pattern. An excitation-inhibition mechanism is then introduced into the learning. Because of the weighting, the clusters tend to be more compact. When the fuzziness parameter is larger, the weighting may also be used to distribute comparatively more excitation on the winning centroid and less excitation on the non-winning centroids by using a weighting smaller than unity. In this case, all the learning rates are reduced and this may increase the learning time.

Table 1: **Subsample Means of the three IRIS Subspecies and Initial Centroids**

	Subsample Means				Initial Centroids			
subspecies 1	5.0060	3.4280	1.4620	0.2460	0.9563	0.9570	0.1796	0.8507
subspecies 2	5.9360	2.7700	4.2600	1.3260	0.7169	0.4361	0.7920	0.5934
subspecies 3	6.5880	2.9740	5.5520	2.0260	0.1419	0.6295	0.1439	0.5399

3 Simulation Results

In this section the FCL and the newly proposed ESFC clusterinig algorithm are applied to the **IRIS** data (Fisher, 1939). The data set contains 150 physically classified patterns in R^4 with 50 for each of the 3 **IRIS** subspecies. The **IRIS** data has been used in many papers to demonstrate the properties of various clustering and classifier algorithms. In the present paper, the final centroids obtained by the algorithms are compared to the physically classified subsample means. The subsample means of the three subspecies are given in Table 1.

In all the simulations, the initial learning rate α_0 was set to 0.3 except when stated otherwise and the initial centroids are given in Table 1. The physical classes of the IRIS data were used only to produce the confusion matrix and were not used in the learning. The element $c_{i,j}$ in the confusion matrix represents the number of samples in the ith subspecies being classified as the jth subspecies (Pal et al., 1993). The confusion matrices were computed by applying the nearest prototype classifier (Cover, 1967) to each of the 150 patterns in the IRIS data. Each input pattern is classified as the same class to the centroid which is the closest to it.

In Table 2, the fuzziness parameter m is 1.2 which is close to unity. When the FCL algorithm ($\beta = 1.0$) was applied to the data set, two of the centroids move rapidly while the second centroid moves very slowly. When β is increased to 1.6, the second centroid moves slightly faster. Increasing the weighting parameter β further induces all the three centroids to move to the correct positions. Similar results were obtained until the weighting was increased to 3.0. For larger weighting, the algorithm applies too much inhibition onto one of the centroids so that the centroid is pushed away from the data.

In Table 3, the fuzziness parameter was increased to 1.5. The FCL ($\beta = 1.0$) algorithm still had similar problems to those in Table 2. If the weighting β is increased to 1.2, the second centroid moves slightly faster but is still not fast enough to reach the correct position. Further increases in β are sufficient to move all the centroids to the correct positions. However when the weighting was 2.6, two of the centroids moved at more or less similar rates initially. Further learning seperated the two centroids and they all reached the correct positions. Increasing the weighting parameter further produced good results until the weighting reached 3.5.

In Table 4, 5 and 6, the fuzziness parameter m was set to 2.0, 2.5 and 3.0 respectively. These are the settings at which the FCL algorithm achieves good clustering results. When the weighting parameter β was small (< 1.0), one of the centroids moved slowly. When the

Table 2: **Results of FCL** ($\beta = 1.0$) **and ESFC** ($\beta \neq 1.0$) **on the IRIS Data** ($\alpha_0 = 0.3$, T = 30)

Parameters	Final Centroids	Confusion Matrix
m=1.2 $\beta = 1.0$	5.0012 3.3441 1.5936 0.3054	50 0 0
(FCL)	0.3831 0.7656 0.2057 0.5233	0 4 46
	6.3878 2.9230 5.1394 1.8089	0 0 50
m=1.2 $\beta = 1.6$	4.9998 3.3609 1.5682 0.2943	50 0 0
(ESFC)	0.8951 1.0506 0.3385 0.4886	0 4 46
	6.4140 2.9343 5.1926 1.8448	0 0 50
m=1.2 $\beta = 1.8$	4.9966 3.4034 1.4891 0.2614	50 0 0
(ESFC)	5.8829 2.7405 4.4923 1.4815	0 48 2
	6.8347 3.0942 5.6987 2.1128	0 15 35
m=1.2 $\beta = 3.0$	4.9912 3.3833 1.5166 0.2745	50 0 0
(ESFC)	5.8981 2.7478 4.5504 1.5223	0 48 2
	6.8346 3.1018 5.7107 2.1335	0 15 35
m=1.2 $\beta = 3.5$	4.9942 3.3582 1.5654 0.3017	50 0 0
(ESFC)	-0.5468 0.0320 -0.2607 1.0875	4 0 46
	6.4809 2.9679 5.3265 1.9405	0 0 50

Table 3: **Results of FCL** ($\beta = 1.0$) **and ESFC** ($\beta \neq 1.0$) **on the IRIS Data** ($\alpha_0 = 0.3$,)

Parameters	Final Centroids	Confusion Matrix
m=1.5 $\beta = 1.0$	5.0119 3.3509 1.6002 0.3064	50 0 0
T = 30	1.0829 1.1099 0.4889 0.5204	4 0 46
(FCL)	6.3971 2.9278 5.1541 1.8171	0 0 50
m=1.5 $\beta = 1.2$	5.0150 3.3529 1.6031 0.3073	50 0 0
T = 30	1.2511 1.1823 0.6000 0.5408	0 4 46
(ESFC)	6.4077 2.9326 5.1753 1.8312	0 0 50
m=1.5 $\beta = 1.4$	4.9987 3.4009 1.4966 0.2636	50 0 0
T = 30	5.8817 2.7531 4.4428 1.4550	0 47 3
(ESFC)	6.8105 3.0752 5.6847 2.0946	0 14 36
m=1.5 $\beta = 2.6$	5.0000 3.3926 1.5116 0.2686	50 0 0
T = 30	6.3622 2.9141 5.0976 1.7875	0 47 3
(ESFC)	6.3645 2.9140 5.0920 1.7813	0 7 43
m=1.5 $\beta = 2.6$	4.9993 3.3970 1.5039 0.2667	50 0 0
T=60	5.8755 2.7574 4.4159 1.4348	0 43 7
(ESFC)	6.7949 3.0675 5.6689 2.0887	0 14 36
m=1.5 $\beta = 3.5$	4.9914 3.3696 1.5394 0.2836	50 0 0
T = 30	5.8860 2.7664 4.5060 1.4891	0 48 2
(ESFC)	6.8068 3.0824 5.6643 2.1114	0 14 36

Table 4: **Results of FCL** ($\beta = 1.0$) **and ESFC** ($\beta \neq 1.0$) **on the IRIS Data** ($\alpha_0 = 0.3$, T = 30)

Parameters	Final Centroids	Confusion Matrix
m=2.0 $\beta = 1.0$ (FCL)	5.0010 3.4038 1.4964 0.2612 5.8944 2.7662 4.4230 1.4340 6.7773 3.0633 5.6499 2.0787	50 0 0 0 47 3 0 14 36
m=2.0 $\beta = 0.7$ (ESFC)	5.0333 3.3904 1.5812 0.2946 3.1922 1.8989 1.8267 0.7360 6.3999 2.9317 5.1493 1.8148	50 0 0 4 0 46 0 0 50
m=2.0 $\beta = 0.8$ (ESFC)	5.0018 3.4109 1.4860 0.2568 5.8343 2.7415 4.3340 1.3918 6.7327 3.0461 5.5918 2.0475	50 0 0 0 46 4 0 11 39
m=2.0 $\beta = 1.2$ (ESFC)	5.0004 3.4014 1.4999 0.2626 5.8961 2.7702 4.4241 1.4325 6.7667 3.0613 5.6386 2.0787	50 0 0 0 47 3 0 13 37
m=2.0 $\beta = 1.4$ (ESFC)	5.0050 3.3926 1.5225 0.2715 6.3487 2.9101 5.0593 1.7622 6.3491 2.9105 5.0598 1.7624	50 0 0 3 44 3 0 15 35

weighting parameter was large (> 1.0), two of the centroids moved at more or less similar rates. These results are very similar to those observed in the FCL algorithm when the fuzziness parameter was either close to unity or large. The weighting parameter was introduced to improve the clustering performance of the FCL algorithm. When the fuzziness parameter is around 2.0, the FCL algorithm usually achieves good clustering results. Therefore it is resonable to expect that the weighting parameter should be close to unity as shown in these simulations. The weighting parameter can however still be used to fine tune the balance between excitation and inhibition during learning.

When the fuzziness parameter is large, the FCL algorithm tends to move all the centroids at similar rates and the final centroids may converge to the same position which resembles the grand mean of the whole data set. The weighting parameter β can then be used to enhance the excitation on the winning centroid and reduce the excitation on the non-winning centroids. However the enhancement is limited since the fuzzy membersihp function $\mu_{r,i,t}$ is generally small and a large fuzziness parameter m will further reduce this enhancement. Since a small weighting (< 1.0) reduces the updating rate for the centroids, the learning process converges very slowly. As shown in Table 7, when the learning rate in the FCL algorithm is small, two of the centroids move at more or less similar rates. The second and third centroids are close to each other. A weighting parameter of less than unity can make these two centroids move away from each other. The effect of the weighting is more clear when the initial learning rate is larger and the number of iterations are higher. Note that the square distances from the centroids $\mathbf{c}_l, l = 1,2,3$ to their corresponding physical subsample means $\bar{\mathbf{c}}_l, l = 1,2,3$ decrease with the increase in the initial learning rate and the number of iterations. In real applications however a high fuzziness parameter should

Table 5: **Results of FCL** ($\beta = 1.0$) **and ESFC** ($\beta \neq 1.0$) **on the IRIS Data** ($\alpha_0 = 0.3$, T = 30)

Parameters	Final Centroids				Confusion Matrix
m=2.5 $\beta = 1.0$	5.0021	3.4000	1.5016	0.2614	50 0 0
(FCL)	5.9109	2.7835	4.4239	1.4262	0 47 3
	6.7362	3.0543	5.6043	2.0684	0 13 37
m=2.5 $\beta = 0.5$	5.0114	3.4187	1.4910	0.2567	50 0 0
(ESFC)	5.4325	2.5769	3.7855	1.1847	0 32 18
	6.5638	2.9894	5.3658	1.9317	0 1 49
m=2.5 $\beta = 1.2$	5.0018	3.3981	1.5040	0.2622	50 0 0
(ESFC)	5.9155	2.7885	4.4284	1.4267	0 47 3
	6.7245	3.0523	5.5917	2.0675	0 13 37
m=2.5 $\beta = 1.4$	5.0049	3.3924	1.5189	0.2680	50 0 0
(ESFC)	6.3411	2.9110	5.0403	1.7536	3 46 1
	6.3412	2.9111	5.0405	1.7538	0 14 36

Table 6: **Results of FCL** ($\beta = 1.0$) **and ESFC** ($\beta \neq 1.0$) **on the IRIS Data** ($\alpha_0 = 0.3$, T = 30)

Parameters	Final Centroids				Confusion Matrix
m=3.0 $\beta = 1.0$	5.0038	3.3976	1.5032	0.2604	50 0 0
(FCL)	5.9262	2.7971	4.4288	1.4246	0 47 3
	6.7034	3.0463	5.5646	2.0563	0 13 37
m=3.0 $\beta = 0.5$	5.0041	3.4040	1.4970	0.2586	50 0 0
(ESFC)	5.8877	2.7745	4.3834	1.4062	0 47 3
	6.7230	3.0467	5.5852	2.0532	0 12 38
m=3.0 $\beta = 1.2$	5.0038	3.3961	1.5048	0.2609	50 0 0
(ESFC)	5.9330	2.8024	4.4356	1.4267	0 47 3
	6.6919	3.0438	5.5517	2.0539	0 13 37
m=3.0 $\beta = 1.4$	5.0056	3.3922	1.5144	0.2646	50 0 0
(ESFC)	6.3311	2.9112	5.0170	1.7425	3 42 5
	6.3311	2.9112	5.0170	1.7426	0 13 37

Table 7: **Results of FCL ($\beta = 1.0$) and ESFC ($\beta \neq 1.0$) on the IRIS Data** (m=7.0)

$\alpha_0 = 0.3$, T=200, $\beta = 1.0$ (FCL)	$\|\|c_l - \hat{c}_l\|\|^2$	$\alpha_0 = 0.3$ T=200, $\beta = 0.5$	$\|\|c_l - \hat{c}_l\|\|^2$
4.1978 2.5230 2.3042 0.8104	2.4999	3.6113 2.1915 2.1440 0.8477	4.3012
5.5552 2.7787 3.8627 1.2819	0.3049	4.8699 2.6076 3.0595 1.0199	2.6978
4.1978 2.5230 2.3042 0.8104	4.3162	5.8650 2.8641 4.2261 1.4038	2.6799
$\alpha_0 = 0.3$, T=400, $\beta = 1.0$ (FCL)	$\|\|c_l - \hat{c}_l\|\|^2$	$\alpha_0 = 0.3$, T=400, $\beta = 0.5$	$\|\|c_l - \hat{c}_l\|\|^2$
5.0159 3.3996 1.5036 0.2525	0.0027	4.9373 3.2923 1.6032 0.3035	0.0464
6.1548 2.8900 4.6991 1.5789	0.3191	5.5059 2.8234 3.5110 1.0972	0.8013
6.2707 2.9184 4.8645 1.6744	0.7001	6.4485 2.9567 5.1914 1.8470	0.1819
$\alpha_0 = 0.6$, T=200, $\beta = 1.0$ (FCL)	$\|\|c_l - \hat{c}_l\|\|^2$	$\alpha_0 = 0.6$ T=200, $\beta = 0.5$	$\|\|c_l - \hat{c}_l\|\|^2$
5.0162 3.3995 1.5040 0.2528	0.0027	4.9376 3.2911 1.6064 0.3051	0.0478
6.1442 2.8879 4.6828 1.5691	0.2951	5.5001 2.8264 3.4957 1.0915	0.8323
6.2838 2.9205 4.8889 1.6878	0.6495	6.4498 2.9572 5.1947 1.8494	0.1782
$\alpha_0 = 0.6$, T=400, $\beta = 1.0$ (FCL)	$\|\|c_l - \hat{c}_l\|\|^2$	$\alpha_0 = 0.6$ T=400, $\beta = 0.5$	$\|\|c_l - \hat{c}_l\|\|^2$
5.0160 3.4008 1.5008 0.2513	0.0024	5.0159 3.4024 1.5001 0.2520	0.0022
6.0249 2.8587 4.5004 1.4584	0.0911	5.9138 2.8172 4.3619 1.3923	0.0175
6.5323 2.9942 5.3250 1.9318	0.0639	6.5864 3.0102 5.4086 1.9717	0.0248
$\alpha_0 = 0.8$, T=200, $\beta = 1.0$ (FCL)	$\|\|c_l - \hat{c}_l\|\|^2$	$\alpha_0 = 0.8$ T=200, $\beta = 0.5$	$\|\|c_l - \hat{c}_l\|\|^2$
5.0165 3.4004 1.5023 0.2521	0.0025	5.0188 3.4036 1.5037 0.2538	0.0026
6.1332 2.8810 4.6598 1.5423	0.2578	5.7632 2.7642 4.1003 1.2859	0.0570
6.4063 2.9458 5.1098 1.8058	0.2779	6.5388 2.9923 5.3429 1.9359	0.0546
$\alpha_0 = 0.8$, T=400, $\beta = 1.0$ (FCL)	$\|\|c_l - \hat{c}_l\|\|^2$	$\alpha_0 = 0.8$ T=400, $\beta = 0.5$	$\|\|c_l - \hat{c}_l\|\|^2$
5.0162 3.4010 1.5007 0.2513	0.0024	5.0153 3.4016 1.5003 0.2520	0.0023
5.9943 2.8480 4.4592 1.4393	0.0620	5.9522 2.8287 4.4111 1.4156	0.0346
6.5583 3.0033 5.3681 1.9525	0.0410	6.5922 3.0123 5.4145 1.9754	0.0229

be avoided if possible.

4 Conclusions

In this paper an enhanced sequential fuzzy clustering algorithm has been proposed. The algorithm overcomes the under utilization problem usually seen in the k-means and related clustering algorithms. The weighting parameter introduced in the new algorithm can be used to balance the excitation on the winning and non-winning centroids. An excitation inhibition mechanism can also be introduced for certain parameter settings. The algorithm achieved better clustering results on the IRIS data set than fuzzy competitive learning when the fuzziness parameter is close to unity.

Acknowledgements

The authors gratefully acknowledge that this work was supported by the EPSRC under contract GR/K71370.

References

Ahalt, S. C., Krishnamurthy, A. K., Chen, P., and Melton, D. E. (1990). Competitive learning algorithms for vector quantization. *Neural Networks*, 3:277–290.

Anderberg, M. R. (1973). *Clustering Analysis for Applications*. Academic Press.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Plenum, New York.

Chinrungrueng, C. and Sequin, C. H. (1995). Optimal adaptive k-means algorithm with dynamic adjustment of learning rate. *IEEE Transactions on Neural Networks*, 6(1):157–169.

Chung, F. L. and Lee, T. (1994). Fuzzy competitive learning. *Neural Networks*, 7:539–551.

Cover, T. M.and Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theroy*, 13(1):21–27.

Everitt, B. (1993). *Clustering Analysis*. 3rd Edn, Edward Arnold.

Fisher, R. A. (1939). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, (7):179–188.

Kohonen, T. (1989). *Self-Organization and Associative Memory*. 3rd Edn, Springer, Berlin.

Martinetz, T. M., Berkovich, S. G., and Schulten, K. J. (1993). Neural-gas network for vector quantization and its application to time-series prediction. *IEEE Transactions on Neural Networks*, 4(4):558–569.

Pal, N. R., Bezdek, J. C., and Tsao, E. C.-K. (1993). Generalized clustering networks and kohonen's self-organizing scheme. *IEEE Transactions on Neural Networks*, 4:549–557.

Rose, K., Gurewitz, E., and Fox, G. C. (1990). Statistical mechanics and phase transitions in clustering. *Physical Review Letters*, 65(8):945–948.