International Journal of Bio-Science and Bio-Technology
Vol. 3, No. 2, June, 2011

Effectiveness Analysis of Fuzzy Unsupervised Clustering Algorithms for Brain Tissue Segmentation in Single Channel MR Image

Ghanshyam D. Parmar, Suman K. Mitra
DA-IICT
India

Abstract

Segmentation of brain tissues is one important process prior to many analyses and visualization tasks for magnetic resonance (MR) images. Clustering is one of the unsupervised techniques for doing the segmentation. Fuzzy clustering techniques have not been applied for single-channel MR images although they have shown promise in segmentation of multichannel MR images. Unfortunately, MR images always contain significant quantity of noise caused by operator performance, equipment and the environment. This noise could lead to serious inaccuracies in the segmentation result. We conduct the research in measuring the performance of fuzzy clustering algorithms over crisp clustering algorithms in different noise level for single-channel MR image. To validate the accuracy and robustness of the result of clustering algorithms we carried out experiments on simulated MR brain scans. The performance of algorithms is analyzed form three measures namely: number of iterations required, misclassification error and per class (tissue) misclassification error in different noise level present in the single-channel MR image. As, clustering is done based on some distance measure, we also compare the performance of clustering algorithms based on distance norm used for it.

Keywords: *Single-channel MR image, segmentation, fuzzy unsupervised clustering algorithm, brain tissue classification.*

1. Introduction

Magnetic resonance imaging (MRI) or nuclear magnetic resonance imaging (NMRI) [1, 2] is primarily medical imaging technique used in radiology to visualize internal structure of the body. MRI provides much greater contrast among different soft tissues of body. This ability makes it useful for neurological, musculoskeletal, cardiovascular and oncological imaging [3]. Brain matter could be generally categorized as White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF) [4, 5]. Most of brain structures are anatomically defined by the boundaries of these tissue classes [4–6]. So we need a method of segmenting tissues in classes. It is an important step for quantitative analysis of the brain and its anatomical structures. Brain tissue classification is also an important step for detection of various pathological conditions affecting brains parenchyma [7–9]. It is also used for surgical planning and simulation [10] and three dimensional visualization for diagnosis and detection of abnormalities [11, 12]. It is also useful in the study of brain development [13, 14] and human aging [15, 16].

In MR imaging, images are produced based on intensities achieved by three tissue characteristics namely: T1 relaxation time, T2 relaxation time and proton density (PD). The images obtained by these properties are known as T1- weighted MR images, T2-weighted MR images and proton density MR images respectively. The effect of these parameters on image can be varied based on the adjusting the parameters like time to echo (TE) and time to repeat of the pulse sequence [17]. By using different parameters or number of echoes in the pulse sequence, a multitude of nearly registered images with different characteristics of same object can be achieved. If only a single MR image of the object is available such an image is referred to as single-channel (single-echo) image, and in case when number of MR images of the same object at same section are obtained, they are referred as multi-channel (multi-spectral or multi-echo) images [18]. For a given scanning time, the voxel sizes achieved in multi-spectral images are larger than those achieved with single-channel images. This ability of finer voxel sizes makes single-channel image more suitable for precise and accurate quantitative measurements of anatomical structures and tissues. Nevertheless multichannel image provides more information at given voxel size than single-channel image [17, 18].

Most of segmentation techniques have relied on multi-spectral characteristics of MR images while a few studies have reported segmentation form single-channel MR images [19]. According to [19], Fuzzy segmentation techniques have not been applied for single-channel image segmentation although they have shown promise in segmentation of multi-channel images [20–22].

Clustering is the process of classification of objects, based on similarities among them. The backbone of clustering is to expose the hidden structure for the purpose of classification or data modeling. The partition to make the cluster can be crisp (hard) or fuzzy (soft). Macqueen [23] introduced the K-means algorithm to cluster numerical data in which each cluster has center called the mean [24]. The distance norm used in this algorithm is standard euclidean distance norm. K-medoid algorithm uses most centrally located data point (medoid) in the cluster to be the cluster center. The distance norm used in this algorithm is also standard euclidean distance norm. Both algorithms are hard partitioning clustering algorithms. Dunn [25] defined an objective function called C-means functional; the clustering algorithm based on its minimization is known as fuzzy C-means. The distance norm used in this algorithm is also standard euclidean distance norm. Gustafson-Kessel [26] extended the standard fuzzy C-means algorithm by employing an adaptive distance norm. The distance norm used in this algorithm is generalized squared mahalanobis distance norm. The fuzzy maxi-mum likelihood estimate (FMLE) clustering employs the distance norm based on fuzzy maximum likelihood estimate proposed by [27], the clustering algorithm based on this distance norm is described by Gath-Geva [28].

In the Section 2, we present the clustering algorithms used in this work. In section 3 the techniques for validation of the clustering results used in this work are described. Then the results of these algorithms are described in the Section 3. Finally, the conclusion of our research is described in the Section 4.

2. Material and Method

Clustering is the process of classification of objects. This is done based on the similarities among the data points or objects. It is an unsupervised method of classification, as we do not have any prior information about classes. Clustering is meant to expose the hidden structure for the purpose of classification or data modeling. Clustering is done by making the partition among the data points. The aim of clustering on the data is to make partitioning among the data points which have dissimilar characteristics and data points in the same group have similar characteristics. The partition can be crisp (hard) or fuzzy (soft).

Given the data set of N points $X = \{X_k \mid k = 1, 2, ..., N\}$, where X_k may be one dimensional or multi dimensional data point. The rows of X are called as patterns or objects, the columns are called as features or attributes and X is called pattern matrix or simply data matrix. In single-channel MR image each X_k will be intensity either of T1 weighted or T2 weighted or PD weighted image and in case of multi channel MR image it will be combination of more than one from above three images. Each X_k can also be features derived from these images.

The following two are the hard (crisp) partition clustering algorithms. Macqueen [23] introduced the K-means algorithm to cluster numerical data in which each cluster has center called the mean [24]. The distance norm used in this algorithm is standard euclidean distance norm. It makes the cluster of data points by minimizing the within cluster sum of square distance [23]. The conventional K-means algorithm is described by [29, 30]. K-medoid algorithm uses most centrally located data point (medoid) in the cluster to be the cluster center. An early K-medoid algorithm called Partitioning Around Medoids (PAM) was proposed by Kauffman [31]. The distance norm used in this algorithm is also standard Euclidean distance norm.

For a data set of **N** points, each having **n** dimensional, K-means and K-medoid classification algorithms assign each data point to one of the c clusters, by minimizing the within sum of square distance norm given by

$$\sum_{i=1}^{c}\sum_{k\in A_i}\| X_k - V_i \|^2, \quad (1)$$

where A_i is the set of data points in the i-th cluster and V_i is the center or prototype of the points in i-th cluster. Eq. (1) denotes euclidean distance norm. In K-means clustering V_i is the called cluster prototype or cluster center is mean of the cluster points, given by

$$V_i = \frac{\sum_{k=1}^{N_i} X_k}{N_i},$$

where N_i is number of points in i-th cluster. In K-medoid clustering algorithm the cluster centers are the most centrally located data points (medoids) rather than the means like in K-means algorithm. Here, the distance norm is same as Eq. (1) but the cluster prototype or centers V_i are the nearest data point to the mean in the data instead of simply the mean in case of K-means clustering algorithm.

Since the concept of Fuzzy sets was introduced by Zadek [32], fuzzy clustering has been widely discussed, studied and applied in various areas [24]. Early work on applying fuzzy set theory in clustering analysis was proposed by [33]. The following three are the soft (fuzzy) partition clustering algorithms. The Fuzzy C-means (FCM) clustering algorithm is based on the minimization of an objective functional defined by Dunn [25]. This is known as C-means functional and given by

$$J(X;U,V) = \sum_{i=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^m \| X_k - V_i \|_A^2, \quad (2)$$

where V_i is the cluster center or prototype, (μ_{ik}) is the fuzzy partitioning matrix, m is the weighting exponent controls the 'fuzziness' of the resulting cluster and $D_{ikA}^2 = \| X_k - V_i \|_A^2 = (X_k - V_i)^T A (X_k - V_i)$, given $1 \le i \le c, 1 \le k \le N$ is squared inner product distance norm. The clustering algorithm based on minimization of Fuzzy C-means functional is known as the Fuzzy C-means clustering algorithm. In fuzzy C-means clustering each V_i gives weighted mean of the cluster, and the weight is given by the membership degree of the fuzzy membership function, so the algorithm is called the fuzzy C-means algorithm. The distance norm used in this algorithm is standard euclidean distance norm. So FCM makes hyper spherical clusters. Depending on the choice of the norm inducing matrix A, as Identity or diagonal, the clusters can be of same shape and orientation or different variance in the direction of the coordinate axis.

As the disability of FCM clustering algorithm to detect the clusters of different geometrical shapes, Gustafson and Kessel [26] extended FCM algorithm to compute clusters of different shapes by using an adaptive distance norm. The objective functional for the Gustafson-Kessel algorithm is given by

$$J(X;U,V,A)=\sum_{i=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^m \| X_k - V_i \|_{A_i}^2, \tag{3}$$

where $D_{ikA_i}^2$ is the adaptive inner-product distance norm and given by $D_{ikA_i}^2 = \| X_k - V_i \|_A^2 = (X_k - V_i)^T A_i (X_k - V_i)$, given $1 \le i \le c, 1 \le k \le N$. (μ_{ik}) is the fuzzy partitioning matrix and m is the weighting exponent controls the 'fuzziness' of the resulting cluster. As the algorithm assigns A_i, the norm inducing matrix, to each cluster it allows the cluster to adapt the distance norm from local topographical structure. Hence it has the ability to detect the cluster of different geometrical shapes. The distance norm used in this algorithm is generalized mahalanobis distance norm. Here the covariance is weighted by the membership degree in the fuzzy membership function. The numerically robust Gustafson-Kessel algorithm is described by [34].

For hyper ellipsoidal clusters and clusters with variable shape and sizes Gath and Geva [28] presented an exponential distance norm based on maximum likelihood estimation. The fuzzy maximum likelihood estimate (FMLE) clustering employs the distance norm based on fuzzy maximum likelihood estimate proposed by Bezdek and Dunn [27]. This distance norm is given by

$$D_{ik}(X_k, V_i) = \frac{\sqrt{\det(F_{mi})}}{\alpha_i} \exp \frac{(X_k - V_i)^T F_{mi}^{-1} (X_k - V_i)}{2}, \tag{4}$$

$$F_{mi} = \frac{\sum_{k=1}^{N} (\mu_{ik})^m (X_k - V_i) A_i (X_k - V_i)^T}{\sum_{k=1}^{N} (\mu_{ik})^m}, \tag{5}$$

where F_{mi} is the fuzzy covariance matrix of the i-th cluster, (μ_{ik}) is the fuzzy partitioning matrix, m is the weighting exponent controls the 'fuzziness' of the resulting cluster and α_i is aprior probability of selecting the i-th cluster. The distance in Eq. (4) is used in the calculation of $P(i/X_k)$, the probability of selecting the i-th cluster given the k-th data point, is given by

$$P(i/X_k) = \frac{\dfrac{1}{D_{ik}(X_k, V_i)}}{\sum_{i=1}^{c} D_{ik}(X_k, V_i)} \tag{6}$$

By using Eq. (4, 6) the clustering is performed based on the distance norm of fuzzy maximum likelihood estimate. A nice toolbox containing these algorithms is made by [35] also in the book [36].

3. Result Validation and Discussion

Even though different scalar validity measures like Partition Coefficient (PC) [37], Classification Entropy (CE), Partition Index (SC), Separation Index(S) [38], Xie and Beni's Index (XB) [39,40], Dunn's Index and many more [41,42] for the validation of the result of the clustering algorithms are available we use the following: As the interest in computer-aided, quantitative analysis of medical image data is growing, the need for validation of such techniques is also increased. For the solution of validation problem, Simulated Brain Database (SDB) is available [43]. The Simulated Brain Database contains a set of realistic MRI data volumes [44] produced by a MRI simulator [45]. This data set is used in our work to evaluate the performance of the tissue classification algorithms in a setting where the truth is known [46]. The detail about the noise used in our work for analysis is described in [43–46].

Table 1: Number of Iterations Required for Clustering in Different Noise Level

	Noise Level in (%)					
Algorithm	0	1	3	5	7	9
K-means	12	10	11	12	11	23
K-medoid	8	9	12	12	11	31
Fuzzy C-means	13	13	13	16	23	52
Gustafson-Kessel	13	13	13	16	23	52
Gath-Geva	30	30	42	46	73	50

Table 2: Misclassification Error in Different Noise Level in (%)

	Noise Level in (%)					
Algorithm	0	1	3	5	7	9
K-means	1.3983	1.6441	3.0023	5.0820	8.4217	15.1202
K-medoid	1.3927	1.6441	2.9796	5.0877	8.4217	15.1653
Fuzzy C-means	1.3077	1.5024	3.0136	5.0766	9.0209	17.9696
Gustafson-Kessel	1.3077	1.5024	3.0136	5.0766	9.0209	17.9696
Gath-Geva	19.4520	10.6072	4.4149	5.2194	8.4258	39.7725

The results of these algorithms in different noise level for the single-channel MR image data set are described in this section. As these algorithms are iterative algorithms, the number of iterations required for the algorithm in different noise level is described in Table 1. In K-means algorithm as the noise level in the MR image is increasing from 0% to 9%; the number of iterations required to do the clustering are also increased from 12 to 23, causing increase in processing time to make the clusters. In case of K-medoid algorithm, in lower noise level comparatively fewer number of iterations are required than that of K-means algorithm, but in higher noise level they increase than that of K-means algorithm. Also the numbers of iterations required for clustering are increased as the noise level in the MR image increase. Fuzzy C-means and Gustafson-Kessel algorithms have shown same number of iterations in 0% to 9% of noise level. In both cases the required iterations are increasing as the noise level in MR image increase. The required iterations are higher than that required by K-means and K-medoid algorithms. Also in very high noise level the iterations required by these two algorithms have increased to very high number. Finally, in Gath-Geva algorithm the required iterations are higher than that of all the previous algorithms described. Also, the required iterations are increasing as the noise level in the MR image increase. The reason for such behavior of this algorithm, Gath-Geva [28] pointed that this algorithm does not perform well as it seeks an optimum in a narrow local region due to the exponential distance norm [24].

The misclassification error in % for different noise level for above discussed algorithms has shown in Table 2. In K-means algorithm the misclassification error increases as the noise level in the MR image increase. In lower noise levels, the value of misclassification error is nearer to the noise present in the MR image. However in higher noise level, the misclassification error is greater than the noise present in the MR image. In K-medoid algorithm the misclassification error increases as the noise level in the MR image increase. In lower noise levels, the value of misclassification error is nearer to the noise present in the MR image. However in higher noise level, the misclassification error is greater than the noise present in the MR image. However, the misclassification error in all noises level is less than that of K-means algorithm. In Fuzzy C-means and Gustafson-Kessel algorithms the misclassification error increase as the noise level present in the MR image increase. Both algorithms are showing almost same misclassification error in the noise level from 0% to 9%. Also, the misclassification error in both algorithms is lesser than that of K-means and K-medoid algorithms. In lower noise levels, the value of misclassification error is nearer to the noise present in the MR image. However in higher noise level, the misclassification error is greater than the noise present in the MR image. Finally, in Gath-Geva algorithm the misclassification error changes dramatically with different noise level. In lower and higher noise level the misclassification error is higher than the noise level and in moderate noise level the error is nearer to the noise level present in the MR image. Other than moderate noise level the misclassification error is higher than all previous algorithms. The reason for such behavior of this algorithm is same as explained before.

The most interesting part of the result is the misclassification error per class (tissue). As we are classifying the MR image into Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF), we compute the misclassification error per class for single-channel MR image after the clustering done by the discussed clustering algorithms. The per class misclassification error is shown in Table 3. As for the clustering algorithms in noise level of 0% to 9%, Table 3 shows the misclassification error in Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF). As we increase the noise level the per tissue class misclassification error for K-means, K-medoid, Fuzzy C-means and Gustafson-Kessel algorithms increase. In these algorithms Gray Matter tissue has the highest per class misclassification error than that of for White Matter (WM) and Cerebrospinal Fluid (CSF) in all noise levels. In lower noise levels, the per class misclassification error for Gray Matter is higher than the noise level, for Cerebrospinal Fluid is nearer to the noise level and for White Matter it is below than the noise level present in the MR image. However in higher noise level, the per class misclassification error for Gray Matter is much higher than the noise level, for Cerebrospinal Fluid it is still below to the noise level and for White Matter it is higher than the noise level present in the MR image. Finally for Gath-Geva algorithm the per class misclassification error changes dramatically with different noise level. In lower and higher noise level the per class misclassification error is higher than the noise level and in moderate noise level the error is nearer to the noise level present in the MR image. Other than moderate noise level the per class misclassification error is higher than all previous algorithms. In some noise levels the per class misclassification error is zero for Gray Matter, but same time it has higher value for Cerebrospinal fluid and White Matter.

Table 3: Per Class (Tissue) Misclassification Error in Different Noise Level in (%)

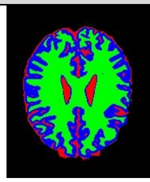
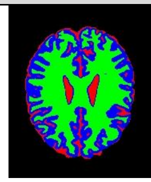
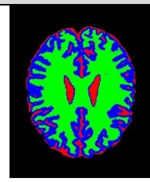
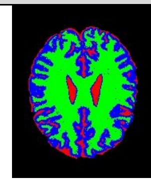
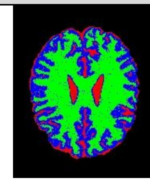
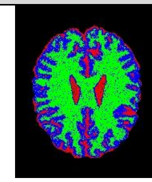
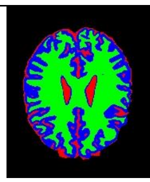
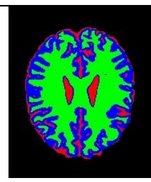
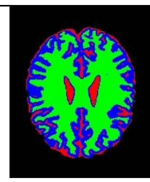
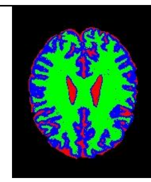
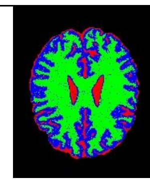
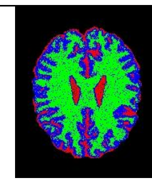
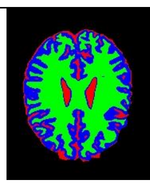
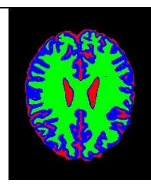
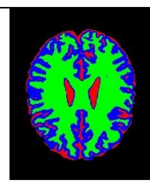
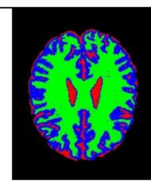
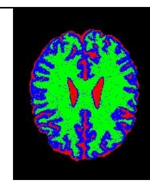
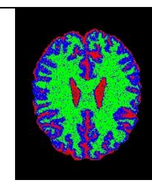
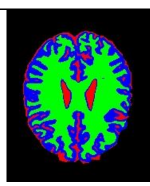
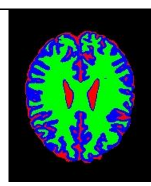
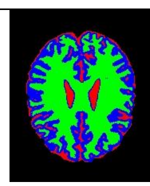
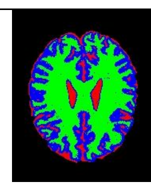
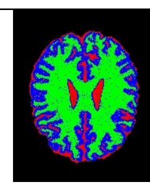
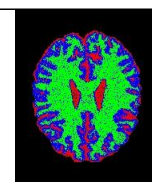
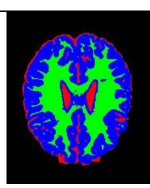
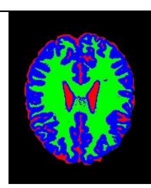
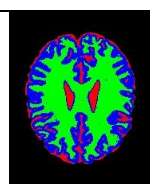
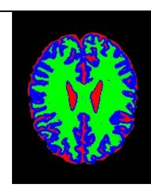
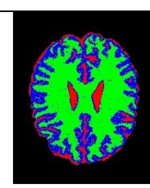
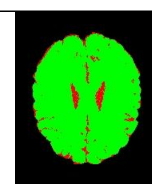
Noise (%)	Tissue Class	Algorithm K-means	K-medoid	Fuzzy C-means	Gustafson Kessel	Gath-Geva
0	CSF	1.1609	1.0491	1.4975	1.4975	18.3700
	GM	2.9360	2.9525	2.3122	2.3122	0
	WM	0.4587	0.4587	0.6304	0.6304	50.3032
1	CSF	1.1905	1.1905	1.4781	1.4781	8.7751
	GM	3.2237	3.2237	2.5331	2.5331	0.0777
	WM	0.7273	0.7273	0.8567	0.8567	22.3753
3	CSF	2.3563	2.3563	2.4152	2.4152	2.1191
	GM	5.8174	5.6955	5.4351	5.4351	2.8842
	WM	1.4543	1.4871	1.7063	1.7063	6.3340
5	CSF	5.1313	5.0031	4.4957	4.4957	5.2058
	GM	9.1017	9.1574	8.6765	8.6765	7.2049
	WM	2.9852	2.9852	3.3469	3.3469	4.4356
7	CSF	6.1996	6.1996	4.6125	4.6125	6.1175
	GM	14.6564	14.6564	14.6975	14.6975	19.3332
	WM	6.2980	6.2980	7.8438	7.8438	3.7682
9	CSF	6.4626	6.4582	4.1181	4.1181	48.9286
	GM	24.9951	24.9951	29.6461	29.6461	all
	WM	15.3460	15.4591	20.3271	20.3271	0

Also, in very high noise level the per class misclassification error is zero for White matter, very high for Cerebrospinal Fluid and all misclassified in case of Gray matter. The images of the classification result are shown in Table 4, where blue color denotes Gray Matter, green color denotes White Matter and red color denotes Cerebrospinal Fluid.

4. Conclusion

This paper presented a comparison of hard and fuzzy clustering algorithms for tissue classification of single-channel MR image in different noise level. The performance of these algorithms has measured in different noise level using three measures namely: number of iterations required, the misclassification error and the per class (tissue) misclassification error. We successfully classify the brain tissues in single-channel MR image in different noise level using hard and fuzzy clustering algorithms. The effect of noise present in the single-channel MR image is measured on the number of iterations required to do the clustering, the misclassification error and the per class (tissue) misclassification error. The proper selection of distance norm used in the clustering algorithm can give better classification result by reducing the misclassification error, with cost of number of iterations required to do the clustering.

Table 4: Segmentation Result of Brian Tissues in the Single-Channel MR Image

	Noise Level in (%)					
	0	1	3	5	7	9
K-means						
K-medoid						
Fuzzy C-means						
Gustafson-Kessel						
Gath- Geva						

Acknowledgement

The authors would like to thank Balazs Balasko, Janos Abonyi and Balazs Feil for making available the Fuzzy Clustering and Data Analysis Toolbox [35], and also for giving the help in the error debugging while practicable implementation.

References

[1] P. C. Lauterbur. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. Nature, 242(5394):190–191, March 1973.

[2] P. C. Lauterbur. Magnetic resonance zeugmatography. Pure Appl. Chem., 40(1-2):149–157, 1974.

[3] Lucy Frank Squire. Fundamentals of radiology. Harvard University Press, 4th ed edition, 1988.

[4] Hanna Damasio. Human brain anatomy in computerized images. Oxford University Press US, 2 editions, 2005.

[5] P. Bourgouin Henri M. Duvernoy. The human brain: surface, three-dimensional sectional anatomy with MRI, and blood supply. Springer, 2 editions, 1999.

[6] John Nolte. The human brain: an introduction to its functional anatomy. Mosby, 1981.

[7] C.R.Jack. Brain and cerebrospinal fluid volume: Measurement with MR imaging. Radiol., (178):22–24, 1991.

[8] T. E. Schlaepfer, G. J. Harris, A. Y. Tien, L. W. Peng, S. Lee, E. B. Federman, G. A. Chase, P. E. Barta, and G. D. Pearlson. Decreased regional cortical gray matter volume in schizophrenia. American Journal of Psychia-try, 151(6):842–848, 1994.

[9] W.L. Oslen, F.M. Longo, D.M. Mills, and D. Norman. White matter disease in aids: Finding at MR imaging. Neuroradiol, 169:445–448, 1988.

[10] K. Fitzgerald. Medical electronics. IEEE Spectrum, 28(1):76–78, 1991.

[11] R.A Robb. Three-Dimensional Biomedical Imaging. New York: VCH, 1995.

[12] M.E Brummer, R.M. Mersereau, R.L. Eisner, and R.R.J. Lewine. Automatic detection of brain contours in mri data sets. IEEE Trans. Med. Imag., 12(2):153–166, 1993.

[13] A. K. H. Miller, R. L. Alston, and J. A. N. Corsellis. Variation with age in the volumes of gray and white matter in the cerebral hemispheres of man: Measurement with an image analyzer. Neuropathol., Appl. Neurobiol., 6:119– 132, 1980.

[14] T. Autti, R. Raininko, S.L. Vanhanen, M. Kallio, and Santavuori. Mri of normal brain from early childhood to middle age. Neuroradiol, 36:644–648, 1994.

[15] D. G. M. Murphy, C. DeCarli, M. B. Schapiro, S.I. Rapoport, and B. Hor-witz. Age-related differences in volumes of subcortical nuclei, brain matter, and cerebrospinal fluid in healthy men as measured with magnetic resonance imaging. Arch. Neurol., 49:839–845, 1992.

[16] K. O. Lim, R. B. Zipursky, M. C. Watts, and A. Pfefferbaum. Decreased gray matter in normal aging: An in vivo magnetic resonance study. J. Gerontol.: Biolog. Sci., 47(1):B26–30, 1992.

[17] Ray H. Hashemi, William G Bradley, and Christopher J. Lisanti. MRI: the basics. Lippincott Williams & Wilkins, 2 edition, 2004.

[18] Donald W. McRobbie, Elizabeth A. Moore, and Martin J. Graves. MRI from picture to proton. Cambridge University Press, 2003.

[19] J.C. Rajapakse, J.N. Giedd, and J.L. Rapoport. Statistical approach to segmentation of single-channel cerebral MR images. IEEE Transactions on Medical Imaging, 16(2):176–186, 1997.

[20] M.E Brand, T.P Bohan, and L.A Kramer. Estimation of csf and gray matter volumes in hydrocephalic children using fuzzy clustering of MR images. Comput. Med. Imaging Graph, 18(1):25–34, 1994.

[21] L.O Hall, M Bensaid, A, L.P Clarke, R.P Velthuizen, M.L Silbiger, and J.C Bezdek. A comparison of neural networks and fuzzy clustering techniques in segmenting magnetic resonance images of the brain. IEEE Trans. Neural Networks, 3(5):672–682, 1992.

[22] Jzau sheng Lin, Kuo sheng Cheng, and Chi-Wu Mao. Segmentation of multi-spectral magnetic resonance image using penalized fuzzy competitive learning network. Computers And Biomedical Research, 29:314–326, 1996.

[23] J. Macqueen. Some methods for classification and analysis of multivariate observations. In University, editor, Fifth Berkeley Symposium on Mathematical Statistics and Probability, pages 281–297, 1967.

[24] Guojun Gan, Chaoqun Ma, and Jianhong Wu. Data Clustering: Theory, Algorithms, and Applications (ASA-SIAM Series on Statistics and Applied Probability). SIAM, 2007.

[25] J. C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Journal of Cybernetics and Systems, 3(3):32–57, 1973.

[26] D.E. Gustafson and W.C. Kessel. Fuzzy clustering with fuzzy covariance matrix. In Proceedings of the IEEE CDC, San Diego, pages 761–766, 1979.

[27] J. C. Bezdek and J. C. Dunn. Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. IEEE Trans. Comput., 24(8):835–838, 1975.

[28] I. Gath and A.B. Geva. Unsupervised optimal fuzzy clustering. IEEE Trans-actions on Pattern Analysis and Machine Intelligence, 11(7):773–780, 1989.

[29] J. Hartigan. Clustering Algorithms. Toronto: John Wiley & Sons., 1975.

[30] J. Hartigan and M Wong. Algorithm as136: A k-means clustering algorithm. applied Statistics, 28(1):100–108, 1979.

[31] L. Kaufman and P. J. Rousseeuw. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley, 1990.

[32] L Zadeh. Fuzzy sets. Information and Control, 8:338–353, 1965.

[33] R Bellman, R Kalba, and L Zadeh. Abstraction and pattern classification. Journal of Mathematical Analysis and Applications, 2:581–586, 1966.

[34] R. Babuka, P.J. van der Veen, and U. Kaymak. Improved covariance estimation for gustafson-kessel clustering. IEEE International Conference on Fuzzy Systems, 2:1081–1085, 2002.

[35] Balasko Balazs, Janos Abonyi, and Balazs Feil. Fuzzy clustering and data analysis toolbox.

[36] Janos Abonyi and Balazs Feil. Cluster Analysis for Data Mining and System Identification. Birkhauser¨ Basel, 2007.

[37] James C. Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms (Advanced Applications in Pattern Recognition). Springer, July 1981.

[38] A.M. Bensaid, L.O. Hall, J.C. Bezdek, L.P. Clarke, M.L. Silbiger, J.A. Arrington, and R.F. Murtagh. Validity-guided (re)clustering with applications to image segmentation. Fuzzy Systems, IEEE Transactions on, 4(2):112– 123, May 1996.

[39] X Xie and G Beni. A validity measure for fuzzy clustering. IEEE Transaction on Pattern Analysis and Machine Intelligence, 15(11):1101–1113, 1991.

[40] X.L. Xie and G. Beni. A validity measure for fuzzy clustering. IEEE Trans-actions on Pattern Analysis and Machine Intelligence, 13(8):841–847, 1991.

[41] Weina Wang and Yunjie Zhang. On fuzzy cluster validity indices. Fuzzy Sets Syst., 158(19):2095–2117, 2007.

[42] F. Ho¨ppner, F. Klawonn, R. Kruse, and T. Runkler. A validity measure for fuzzy clustering. IEEE Transaction on Pattern Analysis and Machine Intelligence, 15(11):1101–1113, 1991.

[43] Chris A. Cocosco, Vasken Kollokian, Remi K.-S. Kwan, G. Bruce Pike, and Alan C. Evans. Brainweb: Online interface to a 3d mri simulated brain database. NeuroImage, 5:425, 1997.

[44] R. K. Kwan, A. C. Evans, and G. B. Pike. Mri simulation-based evaluation of image-processing and classification methods. IEEE Trans Med Imaging, 18(11):1085–1097, November 1999.

[45] Remi Kwan, Alan C. Evans, and G. Bruce Pike. An extensible mri simulator for post-processing evaluation. pages 135–140. Springer-Verlag, 1996.

[46] D. Louis Collins, Alex P. Zijdenbos, Vasken Kollokian, John G. Sled, N. J. Kabani, Colin J. Holmes, and Alan C. Evans. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imaging, 17(3):463–468, 1998.