

FUZZY C-MEANS WITH VARIABLE COMPACTNESS

Snehashis Roy¹, Harsh Agarwal¹, Aaron Carass¹, Ying Bai¹, Dzung L. Pham², and Jerry L. Prince¹

¹Image Analysis and Communications Laboratory, Electrical and Computer Engineering,

²MedIC, Neuroradiology Division, Radiology and Radiological Science,

The Johns Hopkins University

{snehashis,harsh,aaron_carass,ybai,l,pham,prince}@jhu.edu

ABSTRACT

Fuzzy c-means (FCM) clustering has been extensively studied and widely applied in the tissue classification of biomedical images. Previous enhancements to FCM have accounted for intensity shading, membership smoothness, and variable cluster sizes. In this paper, we introduce a new parameter called “compactness” which captures additional information of the underlying clusters. We then propose a new classification algorithm, FCM with variable compactness (FCMVC), to classify three major tissues in brain MRIs by incorporating the compactness terms into a previously reported improvement to FCM. Experiments on both simulated phantoms and real magnetic resonance brain images show that the new method improves the repeatability of the tissue classification for the same subject with different acquisition protocols.

Index Terms— Biomedical image processing, fuzzy sets, image segmentation, magnetic resonance imaging.

1. INTRODUCTION

The fuzzy c-means (FCM) algorithm [1] has been extensively used in medical image segmentation. It has been especially successfully at classifying major tissues from magnetic resonance images of the human brain [2,3]. FCM converges readily, is scale and shift invariant, and allows for the straightforward incorporation of multichannel data. Furthermore, FCM directly yields soft segmentations (in the form of membership functions) that are typically desirable intermediate data structures (as opposed to hard classifications) for further analysis [4]. The original algorithm has also been adapted to account for image shading [5], membership smoothness [6], and variable cluster size [7,8].

Recently, Clark *et. al.* [9] observed that there exists inconsistency in tissue segmentations arising from the choice of pulse sequence in data acquisition. We have had similar observations in our experiments. The inconsistency is caused by the fact that the relative size of clusters can be considerably

different among multi-parametric scans of the same subject. This is an undesirable property as we prefer the performance of the segmentation algorithm to be independent of the data acquisition process. In order to achieve consistent segmentations, we introduce a new parameter, called *compactness*, into the FCM framework to obtain a novel variational formulation that has one compactness parameter per class predetermined for different acquisitions. We name the new algorithm *FCM with variable compactness* (FCMVC). The same idea is also incorporated into an enhanced variant of FCM — Fuzzy and Noise Tolerant Adaptive Segmentation Method (FANTASM) [6], which yields FANTASM with Variable Compactness (FVC).

In this paper, we describe an optimization process to solve the proposed FVC formulation. To validate the proposed method, we compare the performance of FVC against several existing methods, including FANTASM, FCM with a fuzzy covariance matrix [7] (FCMV), and Gaussian Mixture Model (GMM). Experiments show that FVC achieves improved consistency in segmentations when magnetic resonance brain images are acquired using two different pulse sequences (spoiled gradient (SPGR) and magnetization prepared rapid gradient echo (MP-RAGE)) from the same subject.

2. BACKGROUND

Mathematically, FCM is the solution of the following energy function:

$$J_{\text{FCM}} = \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^q (y_j - v_k)^2 \quad (1)$$

where y_j is the observed image intensity at the j^{th} pixel, C is the number of classes, v_k 's are class centroids, Ω is image domain, and u_{jk} is the membership function value of the j^{th} pixel for the k^{th} class. Membership functions must be non-negative, and satisfy the constraint $\sum_{k=1}^C u_{jk} = 1, \forall j \in \Omega$. The energy function is minimized if high membership values are assigned to observations close to centroids and low membership values to observations away from centroids. The parameter q is a weighting exponent and is constrained by

This work was supported by the NIH/NINDS under grant 5R01NS037747.

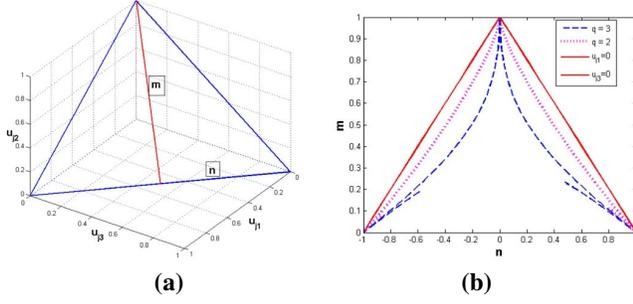


Fig. 1. (a) $u_{j1} + u_{j2} + u_{j3} = 1$ plane. (b) Plot of $u_{jk}(y_j; q)$ on the plane. The left red line is $u_{j3} = 0$ and the right red line is $u_{j1} = 0$.

$q > 1$. When $q = 1$, FCM reduces to the hard K-means algorithm with the membership functions taking binary values. As the value of q increases, the fuzziness of the membership functions also increase.

Since we are interested in segmenting three major tissues in the human brain – the cerebrospinal fluid (CSF), the gray matter (GM), and the white matter (WM), we consider a three class problem in the following manner. Let v_1, v_2 , and v_3 be the class centroids with $v_1 < v_2 < v_3$. Let y_j vary over $(-\infty, +\infty)$ and assume v_k 's are fixed. Recalling that $\forall j, \sum_{k=1}^3 u_{jk} = 1$, we can represent the membership functions $[u_{j1} \ u_{j3} \ u_{j2}]$, for a particular y_j , as a point on the plane $x + y + z = 1$ (Fig. 1(a)) by taking $u_{j1} = x, u_{j2} = z, u_{j3} = y$. We plot the membership points as a continuous function of y_j on this plane and show the 2D projection of the plane in Fig. 1(b) for the choices of $q = 2$ and $q = 3$. The three vertices in Fig. 1(b) starting from bottom left in clockwise direction are v_1, v_2 , and v_3 , respectively. When $y \ll v_1$ or $y \gg v_3$, the curve tends to $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$, i.e, the center of the plane. In our case, for $y_j \in [v_1, v_2]$, it is expected that y_j is a mixture of CSF and GM, with u_{j3} being small. Similarly u_{j1} is expected to be small for $y_j \in [v_2, v_3]$, as it is a mixture of GM and WM. Considering the effect of q on the memberships, we observe that, if q increases, WM membership u_{j3} for $y_j \in [v_1, v_2]$ increases, and so does the CSF membership u_{j1} for $y_j \in [v_2, v_3]$. This demonstrates that the parameter q captures information about the variability of clusters. Based on this interpretation, we assume that if the WM cluster has large variance, then u_{j3} for $y_j \in [v_1, v_2]$ should increase. And similarly, if CSF cluster has large variance, then u_{j1} for $y_j \in [v_2, v_3]$ should increase.

Several modifications of FCM have been proposed in the literature. A generalized L_p norm has been introduced in [10] to include variability of classes into energy function. Another way to take into account cluster variability is to incorporate a covariance matrix between classes, such as proposed with FCMV [7]. A simplified single-channel formulation is given

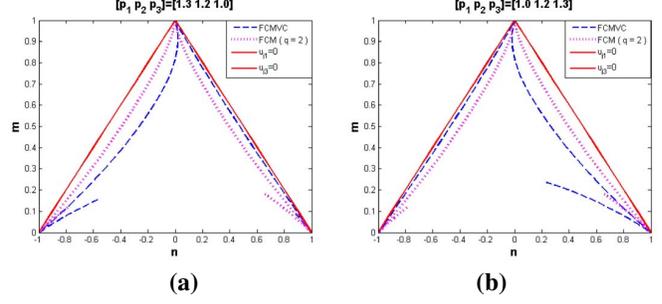


Fig. 2. Plot of $\bar{u}_{jk}(y_j; \mathbf{p})$, for different parameters p_k .

as follows:

$$J_{\text{FCMV}} = \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^q \left(\frac{y_j - v_k}{\sigma_k} \right)^2, \quad (2)$$

where σ_k is the variance of the k^{th} class. In the next section, we will describe a new approach to model the variability of clusters.

3. FCM WITH VARIABLE COMPACTNESS (FCMVC)

3.1. Problem Formulation

We introduce a fixed parameter $p_k, k = 1, \dots, C$, in the FCM framework as minimization of the following function,

$$J_{\text{FCMVC}} = \sum_{j \in \Omega} \sum_{k=1}^C \bar{u}_{jk}^2 (y_j - v_k)^{2p_k}. \quad (3)$$

Minimization of J_{FCMVC} gives the membership functions as,

$$\bar{u}_{jk} = \frac{(y_j - v_k)^{-2p_k}}{\sum_{l=1}^C (y_j - v_l)^{-2p_l}}.$$

We note that $\mathbf{p} = [p_1 \dots p_C] = [1 \dots 1]$ is the same as FCM with $q = 2$. The plot of memberships \bar{u}_{jk} for a fixed set of v_k and p_k are shown in Fig. 2. We observe that for $\forall y_j \in [v_1, v_2]$, \bar{u}_{j3} with $p_1 > p_2 > p_3$ has increased from u_{j3} with $\mathbf{p} = [1 \ 1 \ 1]$ (Fig. 2(a)). Similarly $\forall y_j \in [v_2, v_3]$, \bar{u}_{j1} with $p_1 < p_2 < p_3$, has increased from u_{j1} (Fig. 2(b)). It shows that with decreasing p_k , memberships increase for a fixed $|y_j - v_k|$. So we infer that p_k is a measure of the compactness of cluster k . It is chosen to be large for small classes and small for large classes.

Based on the proposed method FCMVC, we now modify the FANTASM energy function [6] to FVC,

$$J_{\text{FVC}} = \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^2 (y_j - g_j v_k)^{2p_k} + \lambda_1 \sum_{j \in \Omega} \sum_{r=1}^2 (D_r \star g)_j^2 + \lambda_2 \sum_{j \in \Omega} \sum_{r=1}^R \sum_{s=1}^R (D_r \star D_s \star g)_j^2$$

$$+ \frac{\beta}{2} \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^2 \sum_{l \in N_j} \sum_{m \neq k} u_{lm}^2$$

Ω , u_{jk} , p_k , y_j , v_k have been described earlier. Here, g_j is a scalar gain field to account for image inhomogeneity, D_r and D_s are first order finite difference operators, $R = 2$ denotes vertical and horizontal directions, λ_1 and λ_2 are regularizations on gain field, N_j is the set of first order neighbors of pixel j and β is a smoothing coefficient. λ_1 , λ_2 and β are determined empirically.

Minimization of J_{FVC} can be performed using the following algorithm.

1. Choose suitable compactness parameters p_k (discussed in the following section).
2. Obtain initial estimate of v_k .
3. Assume $g_j = 1, \forall j \in \Omega$.
4. Compute membership functions,

$$u_{jk} = \frac{\left((y_j - g_j v_k)^{2p_k} + \beta \sum_{l \in N_j} \sum_{m \neq k} u_{lm}^2 \right)^{-1}}{\sum_{k=1}^C \left((y_j - g_j v_k)^{2p_k} + \beta \sum_{l \in N_j} \sum_{m \neq k} u_{lm}^2 \right)^{-1}}.$$

5. Update the centroids,

$$v_k = \frac{\sum_{j \in \Omega} u_{jk}^2 g_j (y_j - g_j v_k)^{2p_k - 2} y_j}{\sum_{j \in \Omega} u_{jk}^2 g_j^2 (y_j - g_j v_k)^{2p_k - 2}}.$$

6. Update gain field coefficients by solving the spatially varying equation,

$$\begin{aligned} 2y_j \sum_{k=1}^C u_{jk}^2 p_k (y_j - g_j v_k)^{2p_k - 2} v_k = \\ 2g_j \sum_{k=1}^C u_{jk}^2 p_k (y_j - g_j v_k)^{2p_k - 2} v_k^2 \\ + \lambda_1 (H_1 \star g)_j + \lambda_2 (H_2 \star g)_j. \end{aligned}$$

7. If the algorithm converges, stop; otherwise go to step 4.

4. RESULTS

The compactness parameters are estimated based on a set of SPGR training data [4], for which we have manually selected landmarks on the cortical boundaries [11]. An exhaustive search was performed to estimate \mathbf{p} , such that the segmentation isocontours generated by FVC were close to the landmarks. The compactness parameters estimated on the training data are $\mathbf{p}_S = [0.9, 1.00, 1.06]$. Next we estimated the compactness parameters for MP-RAGE using a training set of SPGR and MP-RAGE data of the same subject. The parameters for MP-RAGE are estimated based on an exhaustive search so that the isocontours of MP-RAGE segmentation line

up with the isocontours of SPGR using \mathbf{p}_S . The estimated compactness for MP-RAGE was $\mathbf{p}_M = [1.25, 1.0, 0.95]$. Variances $\sigma_k^2, k = 1, 2, 3$, from Eq. 2, were also estimated in exactly the same way. The estimated variances were found to be $\sigma_S^2 = [3.1, 1.44, 1]$ for SPGR and $\sigma_M^2 = [1, 1.52, 2.55]$ for MP-RAGE. We used these values of $\mathbf{p}_S, \mathbf{p}_M, \sigma_S, \sigma_M$ on the test images.

We conducted two experiments to verify that FVC can improve the consistency of the segmentation of the same subject under the two acquisition protocols. Our first experiment involves two synthetic data sets which we used to simulate the same object imaged with MP-RAGE and SPGR imaging parameters. We generated the phantoms from a fuzzy classification truth model using the statistical model outlined in [12]. The phantoms possessed 5% noise and did not have any inhomogeneity. The misclassification rate, defined as the ratio of the # of correctly classified voxels against the total # of non-background voxels, is reported in Tab. 1 for classifying the MP-RAGE and SPGR phantoms using GMM (with variable priors and variances), FANTASM, FCM with Fuzzy Covariance matrix (FCMV) and FVC. Our second experiment involved a pair of real SPGR and MP-RAGE T1 images for five subjects. We computed the hard segmentations using each of the methods used in the first experiment and report the Jaccard coefficient between the hard segmentations, averaged over all three classes. We used smoothing and inhomogeneity correction while using GMM and FCMV. The results are in Tab. 2, while Fig. 3 shows one of the data sets.

Table 1. Percent Misclassification rate for SPGR/MP-RAGE with Ground Truth for two simulated phantoms. M is MP-RAGE and S is SPGR. FCMV and FVC represent FCM with Fuzzy Covariance matrix and FANTASM with Variable Compactness. The two rows correspond to two different phantoms.

GMM		FANTASM		FCMV		FVC	
S	M	S	M	S	M	S	M
5.64	12.78	3.62	5.66	4.45	5.11	3.55	3.51
7.51	9.54	6.32	8.43	5.32	6.11	4.72	4.86

Tab. 1 shows that FANTASM has nearly the same misclassification rate as FVC on SPGR images, but fails to do as well on MP-RAGE images. From Fig. 3, we observe that GMM tends to over-estimate gray matter in SPGR, while FANTASM tends to over-estimate CSF in MP-RAGE. We also observe that FANTASM produces visually good results on SPGR data. It is also observed from Fig. 3 that FCMV performs better than FANTASM on MP-RAGE, but fails to do so on SPGR. Finally we can see that FVC performs better on both SPGR and MP-RAGE.

Table 2. Jaccard Coefficients, averaged over all three classes, between SPGR and MP-RAGE hard segmentation of five subjects. See Tab. 1 for abbreviations.

#	GMM	FANTASM	FCMV	FVC
1	0.7383	0.5611	0.7781	0.7952
2	0.7047	0.4764	0.7501	0.7538
3	0.8134	0.6670	0.8391	0.8559
4	0.7707	0.5209	0.6363	0.7921
5	0.7377	0.4205	0.4173	0.7588

5. DISCUSSION AND CONCLUSION

A new compactness parameter in the FCM framework is introduced and we outlined an automatic fuzzy segmentation method based on the parameter. We have satisfied our primary goal of obtaining more similar segmentations from the same data acquired under differing protocols, SPGR and MP-RAGE. The proposed approach is currently designed for single channel data, it can be readily extended to a multichannel algorithm. In the future, we would like to develop a robust and fully automated framework for estimating the compactness parameters, as opposed to our current practice of estimation based on training.

6. REFERENCES

- [1] J. C. Bezdek, "A convergence theorem for the fuzzy ISO-DATA clustering algorithms," *IEEE Trans. on Pattern Anal. Machine Intell.*, Jan. 1980.
- [2] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition," *Medical Physics*, vol. 20, pp. 1033–1048, 1993.
- [3] T. D. Pham, "Image segmentation using probabilistic Fuzzy C-Means clustering," *Intl. Conf. on Image Proc.*, vol. 1, pp. 722–725, 2001.
- [4] X. Han, D. L. Pham, D. Tosun, M. E. Rettmann, C. Xu, and J. L. Prince, "CRUISE: Cortical Reconstruction Using Implicit Surface Evolution," *NeuroImage*, vol. 23, pp. 997–1012, November 2004.
- [5] D. L. Pham and J. L. Prince, "An adaptive Fuzzy C-Means algorithm for image segmentation in the presence of intensity inhomogeneities," *Pattern Recognition Letters*, vol. 20, pp. 57–68, January 1999.
- [6] D. L. Pham, "Robust fuzzy segmentation of magnetic resonance images," *14th IEEE Symposium on Computer-Based Medical Systems*, pp. 127–131, 2001.
- [7] D. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," *Proc. IEEE Conf. Decision Contr.*, 1979.
- [8] N. Cavalcanti and F. de A. T. de Carvalho, "An adaptive Fuzzy C-Means algorithm with the L_2 norm," *Australian Conf. on Artificial Intel.*, pp. 1138–1141, 2005.
- [9] K. A. Clark, R. P. Woods, D. A. Rottenber, A. W. Toga, and J. C. Mazziotta, "Impact of acquisition protocols and processing streams on tissue segmentation of T1 weighted MR images," *NeuroImage*, vol. 29, pp. 185–202, 2006.
- [10] R. J. Hathaway, J. C. Bezdek, and Y. Hu, "Generalized Fuzzy C-Means clustering strategies using L_p norm distances," *IEEE Trans. on Fuzzy Systems*, vol. 8, no. 5, pp. 576 – 582, October 2000.
- [11] D. Tosun, M. E. Rettmann, D. Q. Naiman, S. M. Resnick, M. A. Kraut, and J. L. Prince, "Cortical Reconstruction Using Implicit Surface Evolution: A Landmark Validation Study," *MICCAI*, vol. 3216, pp. 384–392, 2004.
- [12] D. L. Pham, J. L. Prince, C. Xu, and A. P. Dagher, "An automated technique for statistical characterization of brain tissues in magnetic resonance imaging," *Intl. Journal of Patt. Recog. and Artificial Intell.*, vol. 11, no. 8, pp. 1189–1211, 1997.

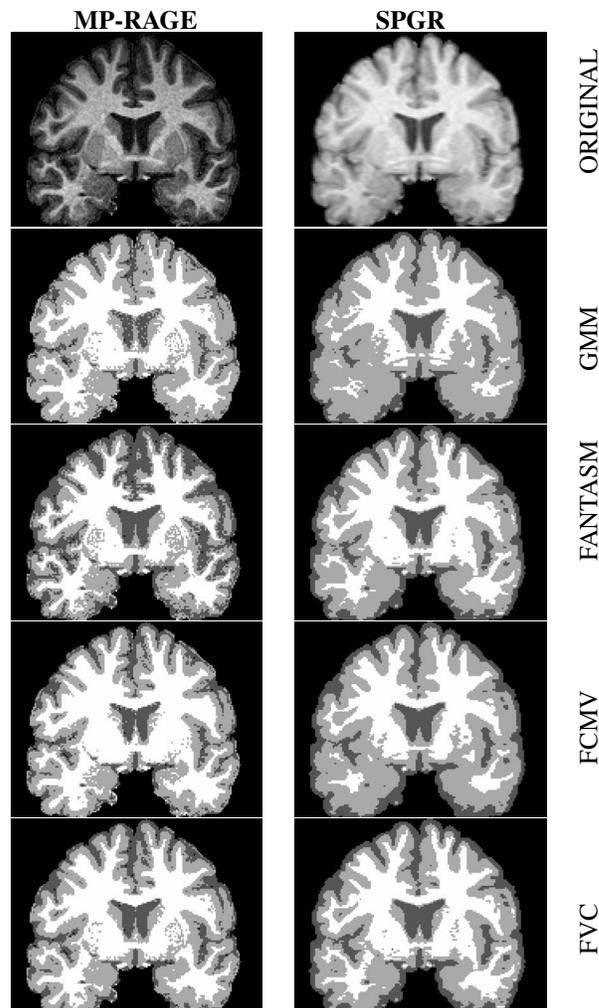


Fig. 3. Comparison between hard segmentation of real MP-RAGE and SPGR images, subject #1. See Tab. 1 for abbreviations.