A KERNEL METRIC AND SPATIAL INFORMATION DRIVEN SIMILARITY MEASURE FOR FUZZY C-MEANS CLUSTERING

Md Tabish Raza[1]
[1]Department of Electronics and Communication Engineering,
Saharsa College of Engineering Saharsa, Bihar, India
tabishtch826@gmail.com

Abstract

In this paper we present a new idea of fuzzy c-means based image segmentation algorithm with a novel similarity measure. Unlikely the Euclidean metric based conventional similarity measure here proposed method hired kernel induced distance measure together with spatial information influenced Euclidean norm by exploiting the significant of both, and computation of this proposed similarity measure is totally free from any adjusting parameter that overcomes trade off ambiguity. Simultaneously, the proposed method could preserve image details as well as suppress image noises via the use of local label information by the addition of the modification of similarity function which is determined by the relationship of the neighborhood pixels. Experiment has been carried out on both synthetic and different kind of real images, and result of proposed algorithm show marked segmentation performance, resistance to noisy images with complex distribution both quantitatively and qualitatively.

Keywords : Fuzzy clustering, Similarity measure, special information, kernel metric.

1. INTRODUCTION

Image segmentation is the major problem in the image analysis and computer vision. The ultimate job of any image segmentation task is to partitioning an image into multiple segments or set of pixels having homogeneity, similar in color, texture, tone etc which makes it more informative. Several image segmentation methods have been proposed in literature and that can be divided in five major groups named as thresholding [1], edge detection [2], region growing method [3], neural network [4], and clustering [5],[6],[7]. Here in this paper a clustering-based image segmentation method is discussed. Among clustering-based image segmentation methods unsupervised fuzzy c-means (FCM) clustering [8] scheme is widely accepted method. The goodness of fuzzy based image segmentation is that it employs the idea of partial membership [9] with each pixel i.e., a pixel can belong to different class with certain amount of membership. This clustering approach shows marked performance especially at difficult situation of image segmentation as intensity inhomogeneity, overlapping intensities, uncertainty, limited spatial resolution, dull contrast etc. Clustering basically approaches to decompose the given dataset into some cluster based on minimizing the objective function iteratively, and fuzzy based approach is most frequently used in different research area as pattern recognition, computer vision, image processing, compute forensics, data mining, and biomedical data analysis. But the traditional FCM image segmentation algorithm work well only with noise free images and go bad to image segmentation task while images are corrupted by different kind of noise and also sensitive to other imaging artifacts, because it only based on gray information distribution and doesn't consider any spatial information circumstances. This traditional FCM are less sensitive to noise to some extent and bring out sensible segmentation outcomes for the deliberated principality. The concepts spatial model employed for segmentation results are still not as accurate as envisioned, because only pixel-level information engrossed for a fixed window element. To overcome such a problem with traditional FCM, many ideas have been proposed in the literature where the gray level information incorporated with local special information such as preprocessing of original image by smoothing filters [10] or post processing of clustered image [11].

2. LITERATURE SURVEY

Different authors have tried to modify the FCM objective function or similarity measure incorporating local spatial information and gray information both [12],[13].[14]. Ahmed et al proposed FCM_S which modified the objective function of FCM by allowing the spatial neighborhood term. Shen et al introduces a new similarity measure with optimized degree of neighborhood attraction using neural network [15]. Enhanced FCM (EnFCM) proposed by Szilagayi et al [16] which constitute a linearly -weighted sum image from both the original image and locally averaged gray level of each pixel which is further employed in clustering with the help of gray level histogram of summed image. Cai et al [00] presented a fast generalized FCM (FGFCM) algorithm. FGFCM reckon to a segmented image output, formed with grey-level information and spatial local information by utilizing factor of similarity measure. Algorithms FGFCM together with EnFCM explained with decisive parameters a (or λ), which has influential pay-off for the performance of those techniques. Selection of these parameters is more often strenuous, because it required to hold equipoise between robustness to noise and efficacy for image details perseverance, indeed it has noise dependency. A larger value chosen to endure the noise, and on the contrary a smaller value chosen to keep the image edge details, in practice the selection of a (or λ) is done experimentally with trial-and-error experiments. Pal et al introduced the concepts of possibilistic FCM (PFCM), which sum-up membership as well as possibilities also. PFCM actually an coalesce of FCM and possibilistic FCM. These algorithms required some parameter to tune trade-off between effectiveness of details preserving and robustness of noise. Park [17] proposed the intuition FCM (IFCM) method. IFCM consider the institution level to redesign the membership function structure to attenuate the effect of noise and it is able to pick-out the inliers comparatively. However, IFCM having even now subtlety with noise due to non-consideration of spatial information. Krinidis and Chatzis [18] proposed a fuzzy c-means algorithm with local information (FLICM). FLICM introduces a new fuzzy factor employing the gray-level information. FLICM preserves image details and shown improved performance with noisy images. A refinement made by Gong et al [19] with FLICM and proposed the KWFLICM algorithm. KWFLICM yields some good enough results but with unsteady and high computational complexity. Liu et al [20] proposed FCM algorithm with adaptive local information, combining region-based and distance-based function. Deng et al presented transfer prototype-based fuzzy clustering that convoy the center of clustering towards the optimized position. Lei et al evinced a variant of FCM as fast and robust FCM (FRFCM) by employing the fitness function with spatial information. This is to lessen the computational complexity and enhance the robustness of traditional FCM. Bai et al presented possibilistic FCM with label information accomplished on improved dice overlap but restraint to maneuver non linearly separable data for computation of Euclidean distance. Here in this paper kernel distance measure has been introduced with the spatial neighborhood influenced Euclidean metric and a novel similarity measure formulated. The rest of the paper is organized as follows; the detail of standard FCM is given in section 3. Section 4 elaborates our proposed algorithm. Experimental result and performance analysis on synthetic noise images, medical images, other real images discussed in section 5. Final conclusion has been drawn in section 6.

3. TECHNICAL BACKGROUND

The Fuzzy C-Means (FCM) algorithm is an unsupervised data` clustering method was first introduced by Dunn [30] and further was extended by Bezdek [12]. Unlike the hard classification where each pixel information is grouped or clustered to any one of the clusters completely (e.g., k-means), FCM is a soft classification method allows pixels of image to different cluster with certain amount of membership based on its probability score or likelihood. It iterates for an optimal C partition cluster by minimizing the objective function J_{FCM} defined as

$$J_{FCM} = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^{m} D_{ij} \tag{1}$$

where X = $\{x_1, x_2, \ldots\ldots x_N\}$ representing the data set of total N pixels that has to be clustered into C classes with given constrain as $2 \leq C < N$, u_{ij} denotes the membership degree of x_j for the cluster i, m ε [1, ∞] is weighting exponent that controls the fuzziness of resultant clustering (usually m =2 is taken). Using the Euclidean norm $D_{ij} = d^2(x_j, v_i)$ is the distance measure in the feature space between the object x_j and the i^{th} cluster center v_i is given in equation (2) and the membership function u_{ij} must satisfy the following constraint condition mentioned in equation (3)

$$d^2(x_j, v_i) = \|x_j - v_i\|^2 \tag{2}$$

$$\sum_{i=1}^{C} u_{ij} = 1 \ , \ \forall_{1 \leq i \leq N} \ j \ ;$$

$$u_{ij} \in [0,1], \ \forall_{1 \leq i \leq c} \ \forall_{1 \leq j \leq N} \ j \ ; \tag{3}$$

$$0 < \sum_{j=1}^{N} u_{ij} < N \ , \ \forall_{1 \leq i \leq c} \ ,$$

From the constraint in equation (3) and using Lagrange multiplier method for local minimizing we have cluster centers.

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m} \tag{4}$$

and membership matrix is

$$u_{ij} = \frac{1}{\sum_{i=1}^{C} \left(\frac{D_{ij}}{D_{kj}} \right)^{\frac{2}{m-1}}} \tag{5}$$

4. METHODOLOGY

In this section the proposed algorithm is discussed which capable to finesse the problem mentioned. From equation (1) the given membership value of FCM clustering decides the segmentation results and simultaneously the membership value is determined by the similarity measurement in equation (5). Therefore, this similarity measurement value has central role toward the successful segmentation. Here in this our proposed method a novel similarity measure is defined employing the kernel metric and spatial information from the data space. The kernel induced distance deduced from the most common one Gaussian Radial Basis Function (GRBF) [22] kernel, and the given similarity measure is as follow

$$D_{ij} = \left(1 - \exp^{-\frac{d^2(x_j, v_i)}{\sigma}} \right) + \psi_{ij} d^2(x_j, v_i) \tag{6}$$

where the term where the term Ψ_{ij} induced by the local special information and is given by equation (8) as induced by the local special information and is given by equation (7) as

$$\psi_{ij} = \frac{\sum_{r=1}^{N_r} u_{ir} d_{jr}^{-1}}{\sum_{r=1}^{N_r} d_{jr}^{-1}} \tag{7}$$

Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11 , Issue 11 , Nov 2022

where d_{jr} is the squared Euclidean distance between the co-ordinate of pixels x(p_j , p_r) and x(q_j , q_r)

$$d_{jr} = (p_j - p_i)^2 + (q_j - q_i)^2 \tag{8}$$

and u_{ir} is the membership of neighboring pixel x_r to the element i and suppose that S is the nxn square window neighborhood of j^{th} sample x_j and its neighbors x_r , N_r is the number of neighboring pixels, the parameter σ estimated based on distance variance of all data points described here.

Finally, the proposed algorithm sketched below:

1. Determine the number of clusters C as 2 ≤ C ≤ N, fuzzy exponent value m (in our study m=2) and stopping criterion ε (some very small positive value).
2. Initialize the fuzzy cluster centers vector randomly V= [$v_1, v_2 \ldots\ldots v_c$].
3. Calculate the similarity measure using Eq (6).
4. Compute the u_{ij} with similarity measure from step-3.
5. Update v_i using u_{ij}.
6. If max {$v_i^b - v_i^{b+1}$}≤ε, then stop otherwise set b=b+1, and go to step-3.

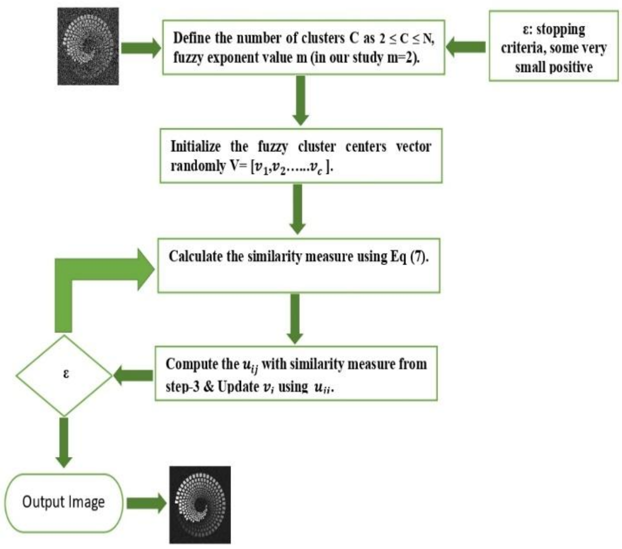
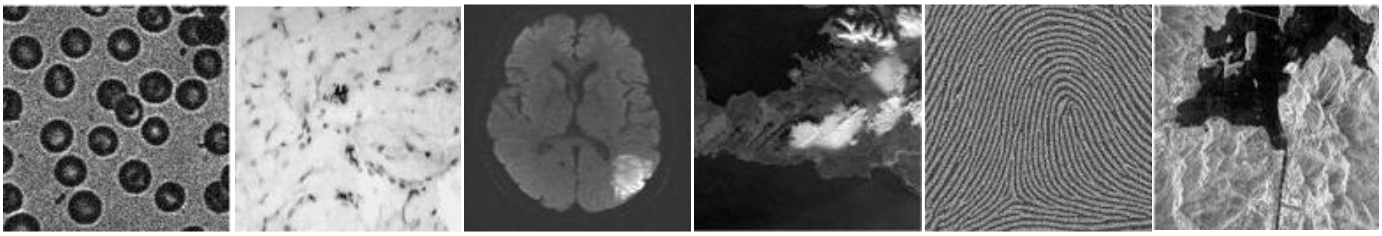


Fig. 1. Block representation for proposed algorithm applied to every image.

5. RESULTS

We have performed simulation on MATLAB R2021a with a system of Intel core™ i7 CPU @ 2.90GHz and installed RAM 16GB having 64-bit window operating system. We experiment with cases of real images in figure 1, with varying segmentation objective, the upper row shows the real images of different application while the lower row gives the segmented performance. Figure 2(a)-2(b) upper row is blood cell image and blood vessel image respectively, lower row images are experimental results of proposed techniques agrees the usefulness for pathological application. Figure 2(c) gives segmentation performance for diffusion weighted MRI brain images.



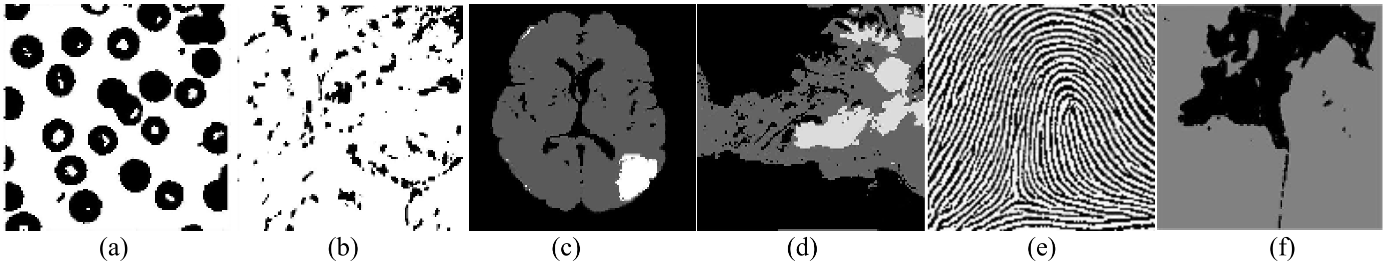


Fig. 2. Segmentation example of six different real images by proposed methods, (a) Blood cell image (b)Blood vessel image (c) DW-MRI image of ischemia (d) Iceland image (e) Fingerprint image (f) Hong Kong image. First row depicts the original images and second row gives segmented image.

Figure 2(c) gives segmentation performance for diffusion weighted MRI brain images. Figure 2(d)-2(f) are satellite images of an Iceland and Hong Kong demonstrate the accomplishment of performance on satellite images. Figure 1 (e) shows performance on fingerprint image and output of proposed algorithms improves the visual quality of finger print pattern for biometric application. The computed value of similarity index (ρ) for the synthetic image is plotted in figure 2, the value is plotted are averaged over different class as $\boldsymbol{\rho_{avr}} = \ (\sum_i^c \boldsymbol{\rho_i})/\boldsymbol{i}$ **.**

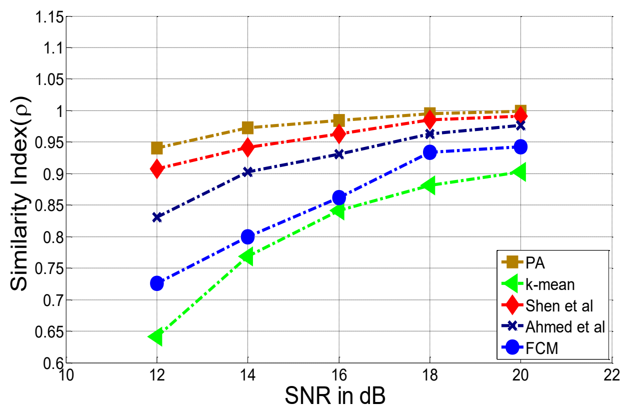


Fig. 3. Quantitative comparison results on different noise levels for the synthetic image with spatial information measure window size 3x3.

Conclusion

In this work, a novel Fuzzy C-Means (FCM) clustering-based unsupervised image segmentation algorithm is proposed. Image segmentation remains a fundamental task in image analysis and computer vision, where the objective is to partition an image into meaningful and homogeneous regions. Traditional FCM-based segmentation methods often suffer from sensitivity to noise and intensity variations, which can significantly degrade segmentation accuracy, especially in noisy environments. To address these limitations, the proposed method integrates kernel-based clustering with spatial information, resulting in a more robust and effective segmentation framework. Recently, kernel methods have gained considerable attention in unsupervised clustering due to their ability to map data into a higher-dimensional feature space, thereby capturing complex and nonlinear relationships among data points. In the proposed approach, a new distance measure is formulated by combining kernel-induced similarity with a spatially influenced Euclidean norm. This hybrid distance metric incorporates both pixel intensity characteristics and neighbourhood information, enabling the algorithm to exploit local contextual information while preserving important image structures.

The inclusion of spatial information significantly improves the robustness of the clustering process against noise and outliers. As a result, the proposed algorithm demonstrates enhanced capability in handling images containing different levels of noise as well as clusters with varying sizes and shapes. Furthermore, by utilizing local neighbourhood information, the method effectively preserves image details and boundaries while suppressing unwanted noise, thereby maintaining the overall integrity of the segmented image.

Experimental results obtained on both synthetic and real-world images demonstrate the effectiveness of the proposed approach. Comparative analysis indicates that the algorithm achieves superior segmentation performance in noisy environments and produces more accurate and visually consistent segmentation results. Unlike several existing robust FCM variants that rely on an additional balancing parameter, such as α or λ, to control the trade-off between noise suppression and segmentation accuracy, the proposed method does not require any such tuning parameter. This eliminates the complexity associated with parameter selection and enhances the practicality and adaptability of the algorithm across different image datasets. Another notable advantage of the proposed method is that it can be applied directly to the original image without requiring any pre-processing or computationally expensive pre-computation steps. Consequently, the algorithm offers a simple yet effective solution for unsupervised image segmentation. Overall, the proposed framework represents a promising compromise between segmentation accuracy, noise robustness, and computational simplicity, particularly in scenarios involving high levels of noise. Nevertheless, there remains scope for further improvement in terms of segmentation precision, computational efficiency, and adaptability to more complex imaging conditions, which can be explored in future research.

Reference

1. N. Pal and S. Pal “A review on image segmentation techniques” Pattern Recognition, vol. 26, pp. 1277–1294, 1993.
2. D. L. Pham, C. Xu, J. L. Prince, A Survey of Current Methods in Medical Image Segmentation in Annual Review of Biomedical Engineering, Vol. 2 (2000), pp. 315-338.

3. Yi-Wei Yu; Jung-Hua Wang, "Image segmentation based on region growing and edge detection," Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on, vol.6, no., pp.798, 803, 1999.

4. Kuo-Sheng Cheng; Jzau-Sheng Lin; Chi-Wu Mao, "The application of competitive Hopfield neural network to medical image segmentation," Medical Imaging, IEEE Transactions on, vol.15, no.4, pp.560, 567, Aug 1996.

5. T. C. Havens, J. C. Bezdek, C. Leckie, L. O. Hall and M. Palaniswami, "Fuzzy c-means algorithms for very large data", IEEE Trans. Fuzzy Syst., vol. 20, no. 6, pp. 1130-1146, Dec. 2012

6. Z. Jiang, T. Li, W. Min, Z. Qi, and Y. Rao, "Fuzzy c-means clustering based on weights and gene expression programming," Pattern Recognit. Lett., vol. 90, pp. 1–7, 2017

7. J. de Oliveira and W. Pedrycz, Advances in Fuzzy Clustering and its Applications. Hoboken, NJ, USA: Wiley, 2007.

8. L. Zadeh, "Fuzzy sets," Inf. Control, vol. 8, pp. 338–353, 1965.

9. M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," IEEE Trans. Med. Imag., vol. 21, pp. 193–199, 2002.

10. Y. Tolias and S. Panas, "Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions," IEEE Trans. Syst., Man, Cybern., vol. 28, no. 3, pp. 359–369, Mar. 1998.

11. Shen, S.; Sandham, W.; Granat, M.; Sterr, A., "MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization," Information Technology in Biomedicine, IEEE Transactions on, vol.9, no.3, pp.459, 467, Sept. 2005

12. F. Zhao, L. Jiao, H. Liu, and X. Gao, "A novel fuzzy clustering algorithm with non-local adaptive spatial constraint for image segmentation," Signal Process., vol. 91, no. 4, pp. 988–999, 2011

13. T. Przybyla, J. Jezewski, K. Horoba, and D. Roj, "Hybrid fuzzy clustering using LP norms," in Proc. 3rd Int. Conf. Intell. Inf. Database Syst., 2011, pp. 187–196

14. Dzung L. Pham, "spatial models for fuzzy clustering," Computer Vision and Image Understanding, vol. 84, 2001, pp. 85-97.

15. Shen S, Sandham W, Granat M, Sterr A, "MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization," IEEE Transaction on Information

Technology in Biomedicine, Sept. 2005, pp. 459-467.

16. Qinghua Zhang, Deyou Xia, Kaixuan Liu, Guoyin Wang A general model of decision-theoretic three-way approximations of fuzzy sets based on a heuristic algorithm, Information Sciences, Volume 507, 2020, Pages 522-539,ISSN 0020-0255, https://doi.org/10.1016/j.ins.20

17. Szilágyi, L., Szilágyi, S.M., Benyó, Z.: Automated medical image processing methods for virtual endoscopy. In: Proc. World Congr. Med. Phys. Biomed. Eng., pp. 2267–2270 (2006)

18. Z. Deng et al., "Transfer prototype-based fuzzy clustering," IEEE Trans. Fuzzy Syst., vol. 24, no. 5, pp. 1210–1232, Oct. 2016

19. L. Szilagyi, Z. Benyo, S. Szilagyii, and H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm," in Proc. 25th Annu. Int. Conf. IEEE EMBS, Nov. 2003, pp. 17–21

20. D.-C. Park, "Intuitive fuzzy c-means algorithm for MRI segmentation," in Proc. IEEE Int. Conf. Fuzzy Syst., Seoul, South Korea, 2010, pp. 1–7.

21. S. Krinidis and V. Chatzis, "A robust fuzzy local information c-means clustering algorithm," IEEE Trans. Image Process., vol. 19, no. 5, pp. 1328–1337, May 2010.

22. M. Gong, Y. Liang, J. Shi, W. Ma, and J. Ma, "Fuzzy c-means clustering with local information and kernel metric for image segmentation," IEEE Trans. Image Process., vol. 22, no. 2, pp. 573–584, Feb. 2013.

23. M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in Proc. Int. Conf. Knowl. Disc. Data Min., Portland, OR, USA, 1996, pp. 226–231.

24. A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," Science, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014

25. J. Hou, H. Gao, and X. Li, "DSets-DBSCAN: A parameter free clustering algorithm," IEEE Trans. Image Process., vol. 25, no. 7, pp. 3182–3193, Jul. 2016.

26. J. Shen, X. Hao, Z. Liang, Y. Liu, W. Wang and L. Shao, "Real-Time Superpixel Segmentation by DBSCAN Clustering Algorithm," in IEEE Transactions on Image Processing, vol. 25, no. 12, pp. 5933-5942, Dec. 2016

[27. Zijdenbos A, Dawant B. Brain segmentation and white matter lesion detection in MR images. Critical Reviews in Biomedical Engineering 1994; 22(5–6):401–65