

A Comparative Study of MRI Image Segmentation based on Fast Kernel Clustering Analysis

Smita Haribhau Zol
Dept of CS and IT,
Dr. B.A.M. U, Aurangabad

R. R. Deshmukh, Ph.D.
HOD, CS and IT Department,
Dr. B.A.M.U, Aurangabad

ABSTRACT

Kernel-based clustering provides a better analysis tool for pattern classification, which implicitly maps input samples to a highdimensional space for improving pattern separability. For this implicit space map, the kernel trick is believed to elegantly tackle the problem of “curse of dimensionality”, which has actually been more challenging for kernel-based clustering in terms of computational complexity and classification accuracy, which traditional kernelized algorithms cannot effectively deal with. In this paper, we have analyzed the merits and deficiencies of KFCM-I/KFCM-II, and KFMC-III and pointed out the connections of these three algorithms.

General Terms

Image processing technique, segmentation.

Keywords

Kernel-based clustering, dimensionality reduction, speeding-up scheme, magnetic resonance imaging (MRI) image segmentation, intensity inhomogeneity correction

1. INTRODUCTION

Among the MRI analyses, precise measurement of tissue distributions is critical for further treatment. But the precise measurement by means of MRI image segmentation is still challenging because of the task itself and a variety of noise corruptions. Though it is not quite a problem to qualitative clinic diagnosis, the multiplicative field could severely tamper with accurate machine-based image processing and analysis. So, the quantitative MRI image analysis needs to rely on bias field.

The prospective methods are basically to reduce the effect of bias fields before imaging process. The prospective methods use image processing techniques, such as filtering, surface interpolation, segmentation, histogram, etc., for intensity inhomogeneity correction after imaging process.

In this paper, we discuss the retrospective correction using segmentation results. Due to the assistance of intensity correction to image segmentation accuracy, the two parts could be considered as the reciprocal ones in one process. For segmentation purposes, a rich variety of algorithms could be used. Here we are interested in segmentation based on fuzzy kernel-clustering due to the adequacy of fuzziness and kernel trick. Using the kernel trick computing the inner product of two mapped data points in original space, kernel-based pattern analysis is assumed to elegantly tackle the problem of “curse of dimensionality” and could therefore be used for effectively segmenting MRI medical images. In this paper, we introduce several kernel-clustering based image segmentation approaches and discuss their merits and drawbacks.

2. FRAMEWORK OF CLUSTERING-BASED IMAGE SEGMENTATION

Conventionally, fuzzy C-means (FCM) clustering is used for image segmentation. To overcome its effectiveness, kernel trick and spatial constraints could be used. For example, the directly kernelized version of FCM, known as KFCM-I could be used to improve the segmentation accuracy by implicitly mapping data to a higher dimensional kernel space and performing FCM clustering in kernel space, transforming the linear method to a nonlinear one. Using a variety of schemes, KFCM-I could be improved. For example, a kernel-induced distance and spatial penalty term are introduced in the objective function of the algorithm called KFCM-II, which aims to preserve the efficacy of kernel tricks and incorporating the spatial continuity in clustering process. Theoretically, both KFCM-I, KFCM-II and KFMC-III could be used for image segmentation. Before the discussion of their applications, it is necessary to briefly introduce some of important backgrounds.

2.1 FMC

KFCM-I and KFCM-II are all based on FCM algorithm. FCM aims to extract C clusters from D dimensional data set containing N data points $x = (x_1, x_2, \dots, x_N)^T \in R^{N \times D}$ by minimizing the following clustering objective function:

$$J_{FCM} = \sum_{i=1}^C \sum_{k=1}^N \mu_{i,k}^m \|x_k - U_i\|^2, \quad (1)$$

where U_i is the i th cluster center, represented as a linear sum of input data points. $\mu_{i,k}$ (satisfying $\mu_{i,k} \in [0,1]$ & $\sum_{i=1}^C \mu_{i,k} = 1$) denotes the membership of x_k belonging to the i th cluster. And m is the clustering constant and usually set as 2 for some convenience.

After the initialization, FCM alternately iterates until the termination condition is reached and then uses the obtained $\mu_k^i \in R^{C \times N}$ to transform the fuzzy clusters to hard ones.

2.2 KFMC-I

The direct kernelized version of FCM, known as KFCM-I, is actually quite straightforward. In KFCM-I, $\Phi(\bullet)$ denotes the implicit Mercer map from original data space to higher dimensional kernel space (where the mapped data points are believed more separable), and equation of FCM could be rewritten as follows:

$$J_{KFCM-I} = \sum_{i=1}^C \sum_{k=1}^N \mu_{i,k}^m \|\Phi(x_k) - v_i^\Phi\|^2$$

where v_i^Φ is the i th cluster obtained in kernel space.

Due to the implicitness of $\Phi(\cdot)$, The iterative formulae of this equation could not be directly computed. This dilemma could be avoided by simplification of the formula and more specifically, equation is rewritten as:

$$v_i^\Phi = \alpha_i^T \Phi(x) = \sum_{k=1}^N \alpha_{i,k} \Phi(x_k)$$

Due to the kernel trick satisfying $k(x, y) = \Phi(x)^T \Phi(y)$ for the Gaussian kernel $K(x, y) = \exp(-\|x - y\|^2 / (2\sigma^2))$ the simplified equation becomes

$$\|\Phi(x_k) - v_i^\Phi\|^2 = 1 - 2 \sum_{k_1=1}^N \alpha_{i,k_1} K(x_k, x_{k_1}) + \sum_{k_1=1}^N \sum_{k_2=2}^N \alpha_{i,k_1} \alpha_{i,k_2} K(x_{k_1}, x_{k_2})$$

Now this equation could be iteratively computed. Although the clustering becomes feasible, the computational complexity has been significantly increased. Even without considering the calculation and storage of kernel matrix, this equation expands the original one multiplication to at least N^2 multiplications, so the complexity increases from $O(1)$ to $O(N^2)$, which is not favourable.

2.3 KFMC-II

The KFMC-II, is designed to solve the complexity problem. In KFMC-II, the v_i^Φ has a pre-image v_i existing in the original space. Compared with KFMC-I, the computational complexity has been significantly reduced.

KFMC-II is effective under following considerations:

- 1) $K(x_k, v)$ is proved a distance defined in the original space.
- 2) Compared with the cluster center in KFMC-I, KFMC-II is a kernel weighted one, but also obtained in the original space. So KFMC-II is actually a degenerated one from KFMC-I in kernel space to input space. This degeneration disqualifies KFMC-II as a kernelized clustering. So, KFMC-II is called “fuzzy clustering with kernel-induced distance”.

Another important issue is that the assumption of the existence of pre-image does not actually hold. The linear sum containing at least two mapped data points does not have an exact pre-image in data space. However, Equation of KFMC-II is still effective as it is actually the approximate pre-image.

Although it guarantees the clustering effectiveness, the aforementioned approximate pre-image formulation inevitably brings the imperfection to KFMC-II, making it less robust and of accuracy limitation.

2.4 KFMC – III

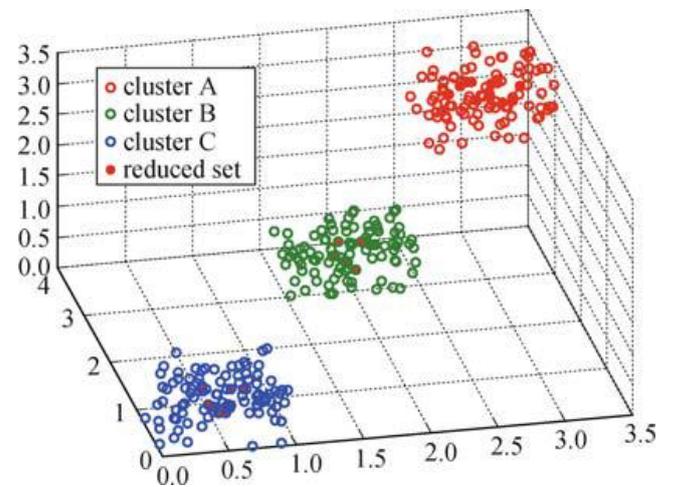
KFMC-III adopts a reduced-set representation for kernel clustering in order to tackle the problems of KFMC-I and KFMC-II. KFMC-III does not believe to use unnecessary and less accurate all N data points to represent the C cluster centers. Instead of this, it uses the data sets which contain

some typical data points with higher memberships belonging to specific cluster, to represent the corresponding cluster centers. Here, the collections of these data points are called as reduced sets.

Each clustering iteration is computed only once because this term is actually a constant when S_i has been determined.

So, the most complex term is computationally distributed over each k . Using this trick, the calculation of equation is accelerated after a relatively slow iteration initialization. KFMC-III gives the reduced-set representation of kernel cluster center.

So the clustering iterates until the termination condition is reached.



In Fig. 1, The illustration of three clusters and their reduced sets is given. Because of their high memberships in reduced sets, the data points are generally some with the nearest distances to their cluster centers while the cluster centers are believed appropriately represented by those data points.

The reduced sets contain typical data points with highest memberships belonging to corresponding clusters. Using this principle, the reduced sets could be determined.

3. CONNECTION BETWEEN KFMC-I, KFMC-II AND KFMC-III (VIA REDUCED-SET REPRESENTATION)

KFMC-III (via reduced-set representation) could be regarded as a generalized version of KFMC-I and KFMC-II. KFMC-I and KFMC-II respectively corresponds to KFMC-III with the largest and smallest reduced sets. Apparently, if all reduced sets are chosen as the whole input set $\{x_1, x_2, \dots, x_N\}$,

and the diagonal elements in $\sum_i = \text{diag}(\sigma_{i,1}^2, \sigma_{i,2}^2, \dots, \sigma_{i,D}^2)$ are identical,

i.e., $\sum_i^\Delta = \text{diag}(\sigma^2, \sigma^2, \dots, \sigma^2)(\forall_i)$, KFMC-III

therefore becomes KFMC-I. On the other hand, $\sum_i^\Delta = (\text{diag}(\sigma^2, \sigma^2, \dots, \sigma^2))$, but setting all reduced sets respectively have only one element, so, KFMC-III is reduced to KFMC-II. So the idea of KFMC-III is to adjust the sizes of each reduced set to well balance the accuracy and efficiency of clustering.

4. SEGMENTATION PERFORMANCES OF KERNEL CLUSTERING ALGORITHMS OVER σ

For studying segmentation, the performances of different kernel clustering algorithms over isotropic Gaussian kernel parameter σ were compared. This setting was mean for the KFCM-III with Σ_i parameterization, it satisfies $\Sigma_i = \sigma I$ (where I is an identity matrix).

By using different σ , this experiment gets a series of accuracy curves of FCM, KFCM-I, KFCM-II and KFCM-III over σ for segmenting a coronal slice respectively corrupted by 0%, 1% and 3% noise. Here, the clustering accuracy of FCM does not depend on σ . So, its segmentation curve is represented as a horizontal line. Fig.2 shows segmentation accuracy curves. The statistics are obtained in the experiments using different algorithms over different values of σ , ranging from 10 to 150 with the step of 5. It is easy to find that the segmentation accuracy of all kernel-based clustering (KFCM-I, KFCM-II and KFCM-III) are all dependent on σ and there exists relative stable range that the tuned σ is located in. Outside the range, the accuracy dramatically deteriorates. The experimental results show that, the discussed KFCM-III outperforms its counterparts when using tuned σ .

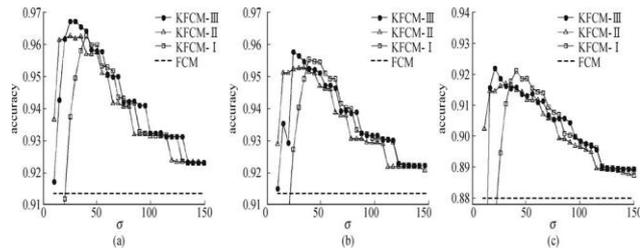


Fig. 2 Typical segmentation accuracy curves of FCM, KFCM-I, KFCM-II and KFCM-III (via reduced-set representation) over different values of σ on a phantom slice respectively corrupted by different noise levels. (a) Results on image slice corrupted by 0% noise; (b) results on image slice corrupted by 1% noise; (c) results on image slice corrupted by 3% noise

5. SEGMENTATION COMPARISON

In an experiment on a coronal slice, the visual segmentation comparison of FCM, KFCM-I, KFCM-II and KFCM-III is given in Fig. 3. The segmentation results are pseudo-coloured, with the red region corresponding to GM, blue region corresponding to WM and the green region corresponding to CSF. In Fig. 3, the optimal segmentations obtained by different algorithms using tuned Gaussian kernel parameter σ are given. It is easy to find that the result obtained by KFCM-III is the closest to the ground-truth.

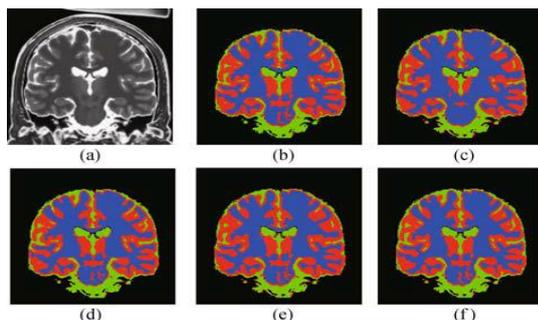


Fig.3 Pseudo-coloured segmentation results of FCM, KFCM-I, KFCM-II and KFCM-III on a phantom slice with red region corresponding to GM, green region corresponding to CSF and blue region corresponding to WM. (a) Original image; (b) ground-truth segmentation; (c) result of FCM; (d) optimal result of KFCMI with tuned parameter $\sigma=40$, segmentation accuracy is 96.28%; (e) optimal result of KFCM-II with tuned parameter $\sigma=25$, segmentation accuracy is 96.28%; (f) optimal result of KFCM-III via reduced-set representation, segmentation accuracy is 96.73%

Reduced sets and bias field estimation for the segmentation experiment corresponding to the results presented in Fig. 3, the pseudo-coloured reduced-set representation of KFCM-III in Fig. 4.

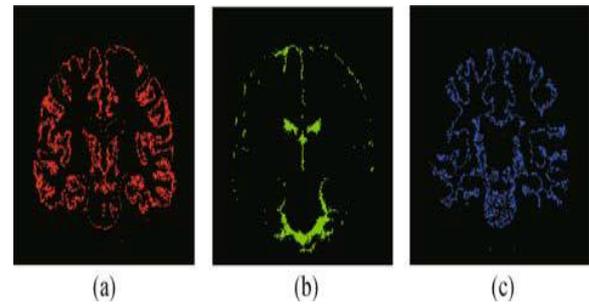


Fig. 4 Pseudo-coloured reduced sets of KFCM-III on a phantom slice with $\sigma=25$. (a) Reduced set for GM; (b) reduced set for CSF; (c) reduced set for WM

Figure 4 shows the pseudo-coloured pixels of reduced sets for the three clusters, with red region, green region and blue region corresponding to GM, CSF and WM, respectively.

The elements of reduced set are basically some pixels with the highest memberships belonging to specific cluster. It is not difficult to find from Fig. 4 that the reduced sets are apparently the subsets of corresponding hard clusters. Figure 5 shows the original target image, corrected image and the bias field estimation. It is apparent that the corrected image is more piece-wise smooth than uncorrected one. The estimated bias field is shown in Fig. 5(c). Besides the low frequencies reflecting the true bias field, the obtained bias field also contains some high frequencies, spatially occurring on the borders of different tissues. High frequencies could further be removed for more accurate bias field estimation. There are a number of high-frequency filtering methods available.

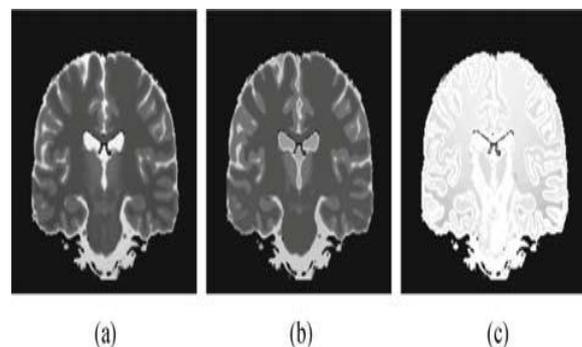


Fig.5 Bias field estimation and image correction (a) Uncorrected image, (b) corrected image, (c) bias field estimation

The discussed image correction and the preclassification scheme could benefit each other, making a fast segmenting convergence. Using the discussed intensity correction scheme, the number of effective (unique) input data points M has been an approximate exponentially decreasing function of clustering iterations.

Figure 6 gives the curve of M over the clustering iterations, where the legend symbol pnXrfY (e.g., pn1rf40) means the experimental slice is corrupted by X% noise and Y % bias field. For clarity purpose, only four curves have been given in Fig. 6 although the experiments with other image corruption combinations had given the similar curves like the four ones shown in Fig. 6. Because of the decrease of M in clustering iteration, the computational complexity has also iteratively reduced. By this way, the introduction of intensity correction does not bring additional computational complexity while improving the segmentation accuracy if correction process is well harnessed.

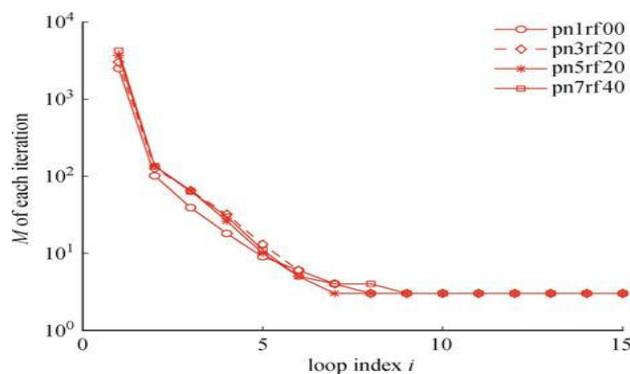


Fig. 6 Curves of number of effective input data points over clustering iterations with image intensity correction

6. CONCLUSION

In this paper, we have studied the merits and deficiencies of FCM, KFCM-I, KFCM-II, and KFMC-III with finding out the connections of algorithms. In our study, KFCM-II could be regarded as the approximate pre-image version of KFCM-I. As it is a direct kernelized version of classic FCM algorithm, needs the unnecessary computational complexity. Due to their deficiencies, KFMC-III gives better results. KFMC-III a novel fast kernel clustering algorithm, via reduced-set representation could be considered as a generalized version of KFCM-I and KFCM-II, and combines the strength of both when using appropriate parameterization. In this work, we have contended that the parsimony for representing cluster centers could be exploited. Using this idea, the use of some data points with highest memberships, included in a subset of input data called reduced set for specific cluster, to represent the corresponding cluster center is used. KFMC-III could be favourably used for dealing with large scale clustering problems. The application on MRI image segmentation with a bias field estimation formulated by clustering residuals shows its effectiveness and superiority to its kernelized counterparts.

7. REFERENCES:

[1] Candes E J, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 2006, 52(2): 489–509

[2] Girolami M. Mercer kernel-based clustering in feature space. *IEEE Transactions on Neural Networks*, 2002, 13(3): 780–784

[3] Zhang D Q, Chen S C, Zhou Z H. Learning the kernel parameters in kernel minimum distance classifier. *Pattern Recognition*, 2006, 39(1): 133–135

[4] Liang Liao, Yanning Zhang MRI image segmentation based on fast kernel clustering Analysis, *Front. Electr. Electron. Eng. China* 2011, 6(2): 363–373 In: Higher Education Press and Springer-Verlag Berlin Heidelberg 2011

[5] Candes E J, Tao T. Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Transactions on Information Theory*, 2006, 52(12): 5406–5425

[6] Donoho D L. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, 52(4):1289–1306

[7] Szilagy L, Benyo Z, Szilagy S M, Adam H S. MR brain image segmentation using an enhanced fuzzy C-means algorithm. In: *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2003, 1: 724–726

[8] Cai W L, Chen S C, Zhang D Q. Fast and robust fuzzy cmeans clustering algorithms incorporating local information for image segmentation. *Pattern Recognition*, 2007, 40(3): 825–838

[9] M`uller K R, Mika S, R`atsch G, Tsuda K. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 2001, 12(2): 181–202

[10] Zhang D Q, Chen S C. A novel kernelized fuzzy C-means algorithm with application in medical image segmentation. *Artificial Intelligence in Medicine*, 2004, 32(1): 37–40

[11] Chen S C, Zhang D Q. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. *IEEE Transactions on System, Man and Cybernetics, Part B*, 2004, 34(4): 1907–1916

[12] Bi L P, Huang H, Zheng Z Y, Song H T. New heuristic for determination Gaussian kernel's parameter. In: *Proceedings of 2005 International Conference on Machine Learning and Cybernetics*. 2005, 7: 4299 – 4304

[13] Wang W J, Xu Z B, Lu W Z. Determination of the spread parameter in the Gaussian kernel for classification and regression. *Neurocomputing*, 2003, 3(4): 643–663

[14] Ahmed M, Yamany S, Mohamed N, Farag A A, Moriarty T. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 2002, 21(3): 193–199

[15] Aubert-Broche B, Evans A C, Collins L. A new improved version of the realistic digital brain phantom. *Neuroimage*, 2006, 32(1): 138–145

[16] Collins D L, Zijdenbos A P, Kollokian V, Sled J G, Kabani N J, Holmes C J, Evans A C. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 1998, 17(3): 463–468

[17] Cocosco C A, Kollokian V, Kwan R K S, Evans A C. BrainWeb: online interface to a 3-D MRI simulated brain.