Kernel Fuzzy Similarity Measure-Based Spectral Clustering for Image Segmentation

Yifang Yang[1], Yuping Wang[1], and Yiu-ming Cheung[2]

[1] School of Computer Science and Technology, Xidian University, Xi'an 710071, China
yangyifang@xsyu.edu.cn, ywang@xidian.edu.cn
[2] Department of Computer Science, Hong Kong Baptist University, Hong Kong
ymc@comp.hkbu.edu.hk

Abstract. Spectral clustering has been successfully used in the field of pattern recognition and image processing. The efficiency of spectral clustering, however, depends heavily on the similarity measure adopted. A widely used similarity measure is the Gaussian kernel function where Euclidean distance is used. Unfortunately, the Gaussian kernel function is parameter sensitive and the Euclidean distance is usually not suitable to the complex distribution data. In this paper, a novel similarity measure called kernel fuzzy similarity measure is proposed first, Then this novel measure is integrated into spectral clustering to get a new clustering method: kernel fuzzy similarity based spectral clustering (KFSC). To alleviate the computational complexity of KFSC on image segmentation, Nyström method is used in KFSC. At last, the experiments on three synthetic texture images are made, and the results demonstrate the effectiveness of the proposed algorithm.

Keywords: spectral clustering, kernel fuzzy-clustering, image segmentation, Nyström method.

1 Introduction

Image segmentation is just to segment an image into different sub-images with different characters and get some interested objects. It is an important process of image analysis and image understanding [1], and plays a fundamental role in computer vision as a requisite step in such tasks as object detection, classification, and tracking [2]. As one of the key methods in image segmentation, clustering algorithms have been widely used in image segmentation [3,4,5,6,7]. In the past few decades, spectral clustering algorithms [4,8,9,10,11] have shown great promise in image segmentation and attracted more and more interest due to its high performance. It realizes dimension reduction by transforming the original dataset into a new one in a lower-dimensional eigenspace, then performs clustering by utilizing the eigenvectors of the normalized similarity matrix derived from the original dataset in lower-dimensional eigenspace.

However, there are still some open problems in spectral clustering algorithms. First of all, the similarity measure is usually constructed by the Gaussian kernel function based on Euclidean distance, while the Gaussian kernel function is parameter sensitive and the Euclidean distance is often not suitable to the complex distribution data [12], therefore, it is crucial to design suitable similarity measures for spectral clustering algorithms. In order to overcome the influence of the scale parameter, Zelnik-Manor and

Perona [13] proposed a self-tuning spectral clustering algorithm (STSC) that utilized a local scale for each data point to replace the single scale parameter, however, the effect of the local scale parameter is similar to that in Gaussian kernel function and the distance is still Euclidean distance. Thus the algorithm cannot overcome the drawbacks of the original spectral clustering [20]. Fischer, B. et al.[21] proposed a path-based similarity. This similarity considers two points in one cluster if they are connected by a set of successive points in dense regions. This is intuitively reasonable. However, it is not robust enough against noise and outliers[22]. Feng Zhao et al.[23] proposed a fuzzy similarity measure by utilizing the prototypes and partition matrix obtained by fuzzy c-means clustering algorithm, however, This similarity is susceptible to the result of fuzzy c-means clustering algorithm. In addition, when the scale n of the data set is relatively large, the overall time complexity and space complexity of standard spectral clustering can reach $O(n^3)$ and $O(n^2)$, respectively [9], which make it difficult to store and decompose a large similarity matrix. Fowlkes et al. [4] presented the Nyström approximation technique to alleviate the computational burden of spectral clustering algorithms.

Several papers reported that the kernel fuzzy-clustering method (KFCM) had better performance than the standard fuzzy clustering method (FCM). Authors in [24] reported the higher classification rate of kernel fuzzy-clustering algorithm on a 2-dimensional non-linearly separable synthetic data set. Also, the kernel-based clustering algorithm performs better on data with non-spherical clusters such as the ring clusters [25].

In this paper, a novel kernel fuzzy similarity measure is proposed and a new spectral clustering algorithm based on this measure is used to the image segmentation. To alleviate the computational complexity, time and space complexity of the algorithm, the Nyström method is applied to the algorithm.

The rest of this paper is organized as follows. In Section 2, we present a short overview about techniques of KFCM, and construct the kernel fuzzy similarity measure. Nyström approximation technique is briefly introduced and the proposed KFSC method is described in Section 3. Section 4 contains parameter setting, computational experiments results, the analysis and discussion. Finally, conclusions are made in Section 5.

2   Kernel Fuzzy Similarity Measure and KFSC Method

2.1   Kernel Fuzzy Similarity Measure

Given a data set $X = \{x_1, x_2, \cdots, x_c\}$ in the p-dimensional space R^p. The KFCM algorithm partitions X into c fuzzy subsets by minimizing the following objective function:

$$J_m(U, V) = 2 \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^m (1 - K(x_k, v_i)) \tag{1}$$

where c is the number of clusters , n is the number of data points, $U = (u_{ik} | i = 1, 2 \cdots, c, k = 1, 2 \cdots, n)$ is partition matrix with $u_{ik}(1 \le i \le c, 1 \le k \le n)$ being the fuzzy membership of kth pixel belonging to the ith cluster and satisfying $\sum_{i=1}^{c} u_{ik} = 1,$

$V = \{v_1, v_2, \cdots, v_c\} \subset R^p$ is cluster centers with $v_i (1 \leq i \leq c)$ being the centroid of ith cluster. $K = e^{-\| x - y \|^2 / t}$ with t being a positive number called Kernel width.

By minimizing $J_m(U, V)$, one can get the partition matrix U and clustering center V [18]

$$\mu_{ik} = \frac{\{1/(1 - K(x_k, v_i))\}^{1/(m-1)}}{\sum\limits_{j=1}^{c} \{1/(1 - K(x_k, v_j))\}^{1/(m-1)}} \tag{2}$$

$$v_i = \frac{\sum\limits_{k=1}^{n} (\mu_{ik})^m K(x_k, v_i) x_k}{\sum\limits_{k=1}^{n} (\mu_{ik})^m K(x_k, v_i)} \tag{3}$$

If the element u_{ij} of matrix U is seen as the probability that the i-th data point belonging to the j-th cluster, we can reasonably assume that two data points belonging to the same cluster have higher similarity, while two data points belonging to different cluster have lower similarity. Let $\mathbf{u}_i$ be the ith column vector of matrix U. It is the probability of data point x_i belonging to c clusters. We can also assume that the greater the inner product of $\mathbf{u}_i$ and $\mathbf{u}_j$, the higher the similarity of data points x_i and x_j. Based on this idea, a new kernel fuzzy similarity measure is proposed

Algorithm 1: A New Kernel Fuzzy Similarity Measure
Input: .data set X to be clustered.
Output: Obtain the similarity matrix S of the dataset.

Step 1. Cluster the data set X into c clusters via KFCM, and obtain the partition matrix U.
Step 2. Let $U = \{\mathbf{u}_1, \cdots, \mathbf{u}_i \cdots, \mathbf{u}_n\}$, here, $\mathbf{u}_i$ is the ith column vector of matrix U, it consist of the membership value that data point x_i belonging to c clusters.
Step 3. For each x_i and x_j,
        If x_i and x_j belonging to the same cluster,
        $s_{ij} = 1$
        Else s_{ij} is equal to the inner product of $\mathbf{u}_i$ and $\mathbf{u}_j$.
Step 4. Finally, the similarity matrix S is obtained.

2.2   KFSC Method for Texture Image Segmentation

In order to apply our method to texture image segmentation, NSCT (Non-subsampled Contourlet Transform) texture feature extracted from the image are used as a suitable image representation. We perform three-level discrete wavelet decomposition on texture images to extract a ten-dimension energy feature using a 16×16 window, which can be written as

$$E = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |coef(i, j)| \tag{4}$$

where $M \times N$ is the subband size and $|coef(i, j)|$ is the coefficient in the ith row and jth column of the subband.

When the size of the dataset is large, spectral clustering encounters an expensive computation and large matrix storage in pairwise similarity construction. Moreover, the algorithm requires much time and memory to compute the eigenvectors of the similarity matrix. Suppose the number of the dataset is N. We select $n(n \lll N)$ sample points randomly, which can define an n-by-n affinity matrix A. B denotes the similarity matrix between sample points and the remaining points with $B \in R^{n \times m}$ and $m = N - n$. The original affinity matrix W can be rewritten as

$$[A \quad B; B^T \quad C]$$

To alleviate the computation and storage burden, Nyström method calculates it by

$$W = \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix}$$

Note that $B^T A^{-1} B$ is easy to obtain. Thus the method reduces the computational complexity of spectral clustering. Based on all these, an outline of KFSC for texture image Segmentation is as follows:

Algorithms 2: KFSC

Input: An $M \times M$ image I to be segmented; the number k of the image segments; kernel parameter σ in NJW (Nyström) and number l of samples in NJW (Nyström); $l > k$.

Step 1. Extract Feature by using NSCT, compute the texture features of each pixel in the image by Eq.(4), and obtain the texture feature dataset X with M^2 instances and 10 attributes.

Step 2. Construct the similarity matrix S using proposed kernel fuzzy similarity measure.

Step 3. Compute the dominant eigenvectors matrix $V \in R^{n \times n} (n = M^2)$ by adopting Nyström approximation technique.

Step 4. Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \lambda_k$ be the k largest eigenvalues of L and $p^1, p^2 \cdots, p^k$ be the corresponding eigenvectors. Form the matrix $P = [p^1, p^2 \cdots, p^k] \in R^{n \times k}$ and here p^i is the column vector.

Step 5. Form the matrix Y from P by renormalizing each of s rows to have unit length (i.e., $Y_{ij} = \{P_{ij} / \sum_j (P_{ij})^2\}$).

Step 6. Treat each row of Y as a point in R^k, and cluster them into k clusters via K-means algorithm to obtain the final segmentation result of image I.

3   Experimental Results and Analysis

3.1   Experimental Setup and Performance Measure

To study the performance of KFSC, it is tested on three synthetic texture images in this section, and compared with three other algorithms: FCM [], KMEANS [], and NJW []. The experiments are implemented in MATLAB 7.10 (R2010a) on a computer with Intel (R) Xeon (R) CPU, 2.53 GHz and Windows XP Professional.

In the experiments, kernel parameter t is set using fast bandwidth selection rule in [19]. For dataset $X = \{x_1, x_2, \cdots, x_N\}$, the data center is denoted by $\overline{x}$. The mean distance is denoted by $\overline{d}$. The bandwidth is defined by $t = \frac{1}{N-1} \sum_{i=1}^{N} (d_i - \overline{d})^2$. In FCM: the maximum number of iterations was 100 and the fuzzy exponent was 2.0. In NJW: the scale parameter varied in the interval [0.15 1] with step length 0.02. In KFSC: the centroids obtained by KMEANS are taken as the initial centroids of KFSC to reduce the instability in initialization. For all the methods, we performed 10 independent runs. The minimum average CE rates were listed in Table 1.

Clustering Error (CE) is used to evaluate clustering performance and defined by [26] $CE = 1 - \frac{\sum_{i=1}^{n} \delta(y_i, map(c_i))}{n}$, where n is the number of samples and y_i and c_i denote the true label and the clustering label of the algorithm, respectively, $\delta(y, c)$ equals one if $y = c$, and zero, if else. $map(\cdot)$ maps each cluster label to a category label. The smaller the Clustering Error, the better the performance.

3.2   Analysis of Experimental Results

In this section, the three synthetic texture images with two, three, and four categories from the Brodatz album of the University of Southern California are used in the experiments. The statistical results (average CE rates) are shown in Table 1.

Table 1. Comparison of the Minimum average CE rates (%) obtained by compared algorithms

	$Texture2$	$Texture3$	$Texture4$
FCM	10.38	8.25	5.35
KMEANS	8.93	8.10	5.12
NJW (Nyström)	5.69	7.66	4.92
KFSC	5.07	7.08	4.84

It can be seen from Table 1 that KFSC outperforms the other three algorithms. Therefore, the proposed KFSC is effective and non sensitive to the scale parameter on spectral clustering.

Fig. 1 shows the segmentation results of the synthetic texture image with two categories of the compared algorithms. It can be seen that there are some misclassified spots in black regions in Fig. 1(c) and (d); whereas, the segmented result in Fig. 1(e) is the better, and the result in Fig.1(f) is the best. This coincides with the results in Table 1. From the segmentation results for two categories texture images, KFSC method obtains the best performance.

Fig. 2 shows the segmentation results of the synthetic texture image with three categories. The similar conclusion can be obtained.

Fig. 3 shows the segmentation results of the synthetic texture image with four categories. In Figs. 3(c) and 3(d), there are also some misclassified points on the upper left of the figures. Thus, FCM and KMEANS perform the worst, NJW performs better than FCM and KMEANS, and KFSC obtain the best performance.
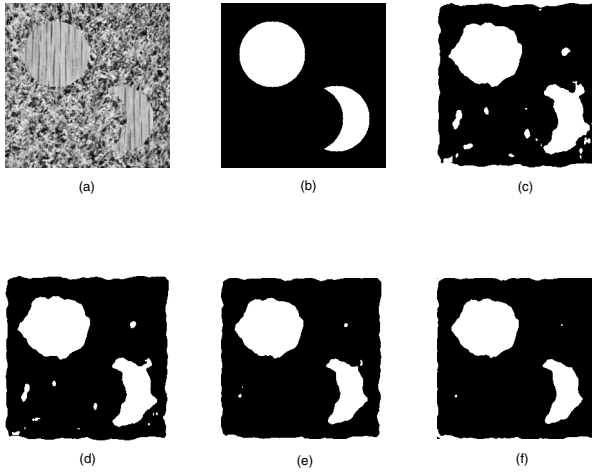
Fig. 1. The segmentation results of the synthesized texture image with two categories. (a) The original image (256 256 pixels); (b) The ideal segmentation, (c) FCM , (d) KMEANS, (e) NJW (Nyström) , (f) KFSC at t=0.23.
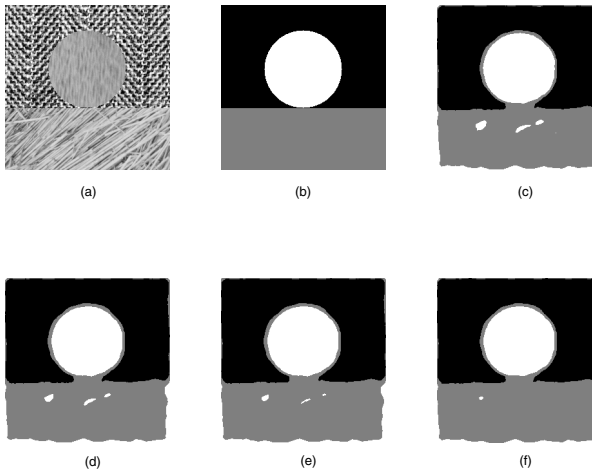

Fig. 2. The segmentation results of the synthesized texture image with three categories. (a) The original image (256 256 pixels); (b) The ideal segmentation, (c) FCM , (d) KMEANS, (e) NJW (Nyström) , (f) KFSC at t=0.25.
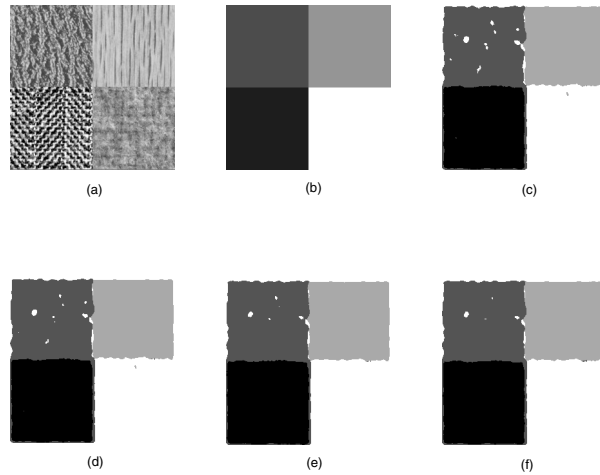
Fig. 3. The segmentation results of the synthesized texture image with four categories. (a) The original image (256 256 pixels); (b) The ideal segmentation, (c) FCM , (d) KMEANS, (e) NJW (Nyström) , (f) KFSC at t=0.28.

As seen from Fig. 1-3 and Table 1, KFSC have obtained more satisfying segmentation results than other three algorithms. Thus, we can conclude that KFSC is efficient and robust.

4   Conclusion

In order to solve the drawbacks of the similarity measure based on Euclidean distance, a novel image segmentation method (KFSC) is proposed based on a novel kernel fuzzy similarity measure. The proposed KFSC can not only avoid the influence of the scale parameter on spectral clustering effectively, but also overcome the drawback of Euclidean distance measure. In addition, the Nyström method is applied to alleviate the computational complexity. Experiments on three synthetic texture images indicate the proposed algorithm is efficient and robust.

References

1. Shi, J., Malik, J.: Normalized cuts and image segmentation. In: IEEE Conf. Computer Vision and Pattern Recognition, vol. 3(2), pp. 731–737. IEEE Computer Society (1997)
2. Yilmaz, A., Javed, O., Shah, M.: Object tracking: A survey. ACM Computing Surveys 38(4), 1–45 (2006)

3. Chen, S.C., Zhang, D.Q.: Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. IEEE Transactions on Systems Man and Cybernetics Part B: Cybernetics 34(4), 1907–1916 (2004)
4. Fowlkes, C., Belongie, S., Chung, F., Malik, J.: Spectral grouping using the Nyström method. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(2), 214–225 (2004)
5. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8), 888–905 (2000)
6. Zhao, F., Jiao, L.C., Liu, H., Gao, X.B.: A novel fuzzy clustering algorithm with nonlocal adaptive spatial constraint for image segmentation. Signal Processing 91(4), 988–999 (2011)
7. Liu, H., Zhao, F., Jiao, L.: Fuzzy spectral clustering with robust spatial information for image segmentation. Applied Soft Computing 12, 3636–3647 (2012)
8. Liu, H.Q., Jiao, L.C., Zhao, F.: Non-local spatial spectral clustering for image segmentation. Neurocomputing 74(1-3), 461–471 (2011)
9. Gou, S.P., Zhuang, X., Jiao, L.C.: Quantum Immune Fast Spectral Clustering for SAR Image Segmentation. IEEE Geoscience and Remote Sensing Letters 9(1) (January 2012)
10. Chen, W., Feng, G.: Spectral clustering with discriminant cuts. Knowledge-Based Systems 28, 27–37 (2012)
11. Rebagliati, N., Verri, A.: Spectral clustering with more than K eigenvectors. Neurocomputing 74, 1391–1401 (2011)
12. Su, M.C., Chou, C.H.: A modied version of the K-means algorithm with a distance based on cluster symmetry. IEEE Trans. Pattern Anal. Mach. Intell. 23, 674–680 (2001)
13. Zelnik-Manor, L., Perona, P.: Self-tuning spectral clustering. In: Advances in Eighteenth Neural Information Processing Systems (NIPS), Vancouver, Canada, pp. 1601–1608 (2004)
14. Kim, D.W., Lee, K.Y., Lee, D., Lee, K.H.: Evaluation of the performance of clustering algorithms in kernel-induced feature space. Pattern Recognit. 38(4), 607–611 (2005)
15. Graves, D., Pedrycz, W.: Performance of kernel-based fuzzy clustering. Electron. Lett. 43(25), 1445–1446 (2007)
16. Graves, D., Pedrycz, W.: Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study. Fuzzy Sets Syst. 161(4), 522–543 (2010)
17. Chen, L., Chen, C.L.P., Lu, M.: A Multiple-Kernel Fuzzy C-Means Algorithm for Image Segmentation. IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics 41(5), 1263–1274 (2011)
18. Zhang, D.Q., Chen, S.C.: A novel kernelized fuzzy C-means algorithm with application in medical image segmentation. Artif. Intell. Med. 32(1), 37–50 (2004)
19. Tsai, D.-M., Lin, C.-C.: Fuzzy C-means based clustering for linearly and nonlinearly separable data. Pattern Recognition 44, 1750–1760 (2011)
20. Zhang, X., Li, J., Yu, H.: Local density adaptive similarity measurement for spectral clustering. Pattern Recognition Letters 32, 352–358 (2011)
21. Fischer, B., Buhmann, J.M.: Path-based clustering for grouping of smooth curves and texture segmentation. IEEE Trans. Pattern Anal. Machine Intell. 25(4), 513–518 (2003)
22. Chang, H., Yeung, D.-Y.: Robust path-based spectral clustering. Pattern Recognit. 41(1), 191–203 (2008)
23. Zhao, F., Liu, H., Jiao, L.: Spectral clustering with fuzzy similarity measure. Digital Signal Processing 21, 701–709 (2011)
24. Zeyu, L., Shiwei, T., Jing, X., Jun, J.: Modified FCM clustering based on kernel mapping. In: Proc. of the Internat. Society for Optical Engineering, vol. 4554, pp. 241–245 (2001)
25. Graves, D., Pedrycz, W.: Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study. Fuzzy Sets and Systems 161, 522–543 (2010)
26. Bach, F., Jordan, M.: Learning spectral clustering. In: Proceedings of NIPS 2003, pp. 305–312 (2003)