Neurocomputing 144 (2014) 346–356

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Image segmentation using spectral clustering of Gaussian mixture models

Shan Zeng [a], Rui Huang [b,*], Zhen Kang [a], Nong Sang [b]

[a] College of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan, Hubei 430023 China
[b] School of Automation, Huazhong University of Science and Technology, Wuhan, Hubei 430074 China

ARTICLE INFO

Article history:
Received 5 December 2013
Received in revised form
10 March 2014
Accepted 25 April 2014
Communicated by Qingshan Liu
Available online 28 May 2014

Keywords:
Image segmentation
GMMs
EM algorithm
KL divergence
Floyd's algorithm
Spectral clustering

ABSTRACT

A novel image segmentation method that combines spectral clustering and Gaussian mixture models is presented in this paper. The new method contains three phases. First, the image is partitioned into small regions modeled by a Gaussian Mixture Model (GMM), and the GMM is solved by an Expectation–Maximization (EM) algorithm with a newly proposed Image Reconstruction Criterion, named EM-IRC. Second, the distances among the GMM components are measured using Kullback–Leibler (KL) divergence, and a revised Floyd's algorithm developed from Zadeh's operations is used to build the similarity matrix based on those distances. Finally, spectral clustering is applied to this improved similarity matrix to merge the GMM components, i.e., the corresponding small image regions, to obtain the final segmentation result. Our contributions include the new EM-IRC algorithm, the revised Floyd's algorithm, and the novel overall framework. The experimental evaluation on the IRIS dataset and the real-world image segmentation problem demonstrates the effectiveness of our proposed approach.

1. Introduction

Image segmentation is an important component of image understanding. It aims at grouping image pixels based on their properties such as gray level, color, and texture, and segmenting an image allows us to access the image content at the level of objects. Image segmentation has been an active research area for a long time. There are many segmentation algorithms but none is generic enough for automatic segmentation of all kinds of images. Existing image segmentation approaches can be broadly divided into four categories: thresholding, clustering, edge detection and region extraction [1]. In this study, clustering-based image segmentation approaches will be considered [2,3].

In clustering-based image segmentation approaches, spectral clustering and soft clustering are two typical type of methods. Spectral clustering concerns the optimization of the graph cut problem in spectral graph theory, and has shown promising performance in image segmentation [3]. In spite of the success of various spectral clustering algorithms, however, when the images are heavily corrupted by noise, outliers and other imaging artifacts, these methods often fail [4]. Soft clustering algorithms also have proved to be able to obtain good results in image segmentation. In soft clustering, each image pixel is allowed to belong to to more than one cluster, associated with a set of membership levels, which indicate the strength of the associations between the pixel and the clusters. Fuzzy C-Means (FCM) clustering and Expectation–Maximization (EM) for Gaussian Mixture Models (GMMs) are important representatives of soft clustering algorithms, and there are many improvements and variants for FCM and EM, such as SMEM [5], SEM [6], and FFCM_SNLS [7]. For these algorithms, the noise problem is still left unsolved. Many strategies, e.g., incorporating spatial information [8–10], are introduced to make FCM and EM robust against noise to some extent. In this paper, we choose the EM algorithm for GMMs as a building block of our framework.

The EM algorithm has been traditionally used to find the maximum likelihood estimates of the parameters in GMMs. A key issue in such a problem is the model complexity [11], and a crucial problem in GMM based segmentation is how to choose a proper number of Gaussian components. Generally, the larger the number is, the more closely the image recreated using the Gaussian components resembles the original image. However, the computational burden and the data overfitting will also be higher.

The number of components can be estimated using the Bayesian Information Criterion (BIC) [12]. This methodology, however, often leads to solutions with an unreasonably high number of poorly separated components. In order to solve this

* Corresponding author.
E-mail address: ruihuang@hust.edu.cn (R. Huang).

http://dx.doi.org/10.1016/j.neucom.2014.04.037
0925-2312/

problem, some authors proposed some alternatives to the EM-BIC combination for determination of the number of components and parameter estimation. Although such methods might seem suitable for clustering purposes, they often lead to parameter estimates that are not based on the maximum likelihood criterion [13]. In practice, we also found the number of components given by BIC is too large for the image segmentation problem. For example, for the image shown in Fig. 1(a), the number of components estimated by BIC is on the order of hundreds.

In this paper, a fast and accurate method for finding an optimal number of Gaussian components based on the Image Reconstruction Criterion (IRC) is proposed, which is named EM-IRC. In the EM-IRC algorithm, the reconstructed image obtained from the GMM fitting is used to compute the reconstruction error and the optimal number of components is the one at which the reconstruction error reaches the threshold. The EM-IRC algorithm usually obtains a much smaller number of components than EM-BIC.

However, in the real-world image segmentation problem, the number of regions perceived by humans is often still much smaller than the number found by the algorithm. For instance, a smooth but slowly-varing region will be poorly fitted by Gaussian distributions, and in that case GMM-based clustering tends to represent this one non-Gaussian cluster by a mixture of two or more Gaussian components. This is because EM-IRC selects the number of Gaussian components to provide a good approximation to the density, rather than to fit the human perception. Therefore, the number of Gaussian components for the purpose of image reconstruction is generally greater than the number of image segments for the purpose of image segmentation. If we interpret the number of mixture components as the number of image segments, this can still lead to over-segmentation. Although the number of Gaussian components selected by the EM-IRC algorithm is greatly reduced compared to the EM-BIC method, we still need to merge the Gaussian components generated by EM-IRC, for the purpose of image segmentation in practice.

Assuming that the number of clusters in the data is smaller than the number of Gaussian components in the model, and a cluster could be associated with multiple components, Jean–Patrick Baudry proposed to first select the total number of Gaussian mixture components using BIC and then combine them hierarchically according to an entropy criterion [14]. Hennig proposed a similar approach, where some aggregation criteria are proposed, based either on modality or on misclassification probabilities [15]. Pastore and Tonellato proposed a hierarchical aggregation algorithm based on a generalization of the definition of silhouette-width taking into account the Mahalanobis distances induced by the precison matrices of the components of the fitted GMM [16]. Tung et al. proposed the CURLER algorithm [17]. This algorithm works in two steps. First, the EM algorithm is adopted to form the subclusters (smaller clusters). Second, subclusters that are close both in proximity and orientation are merged by

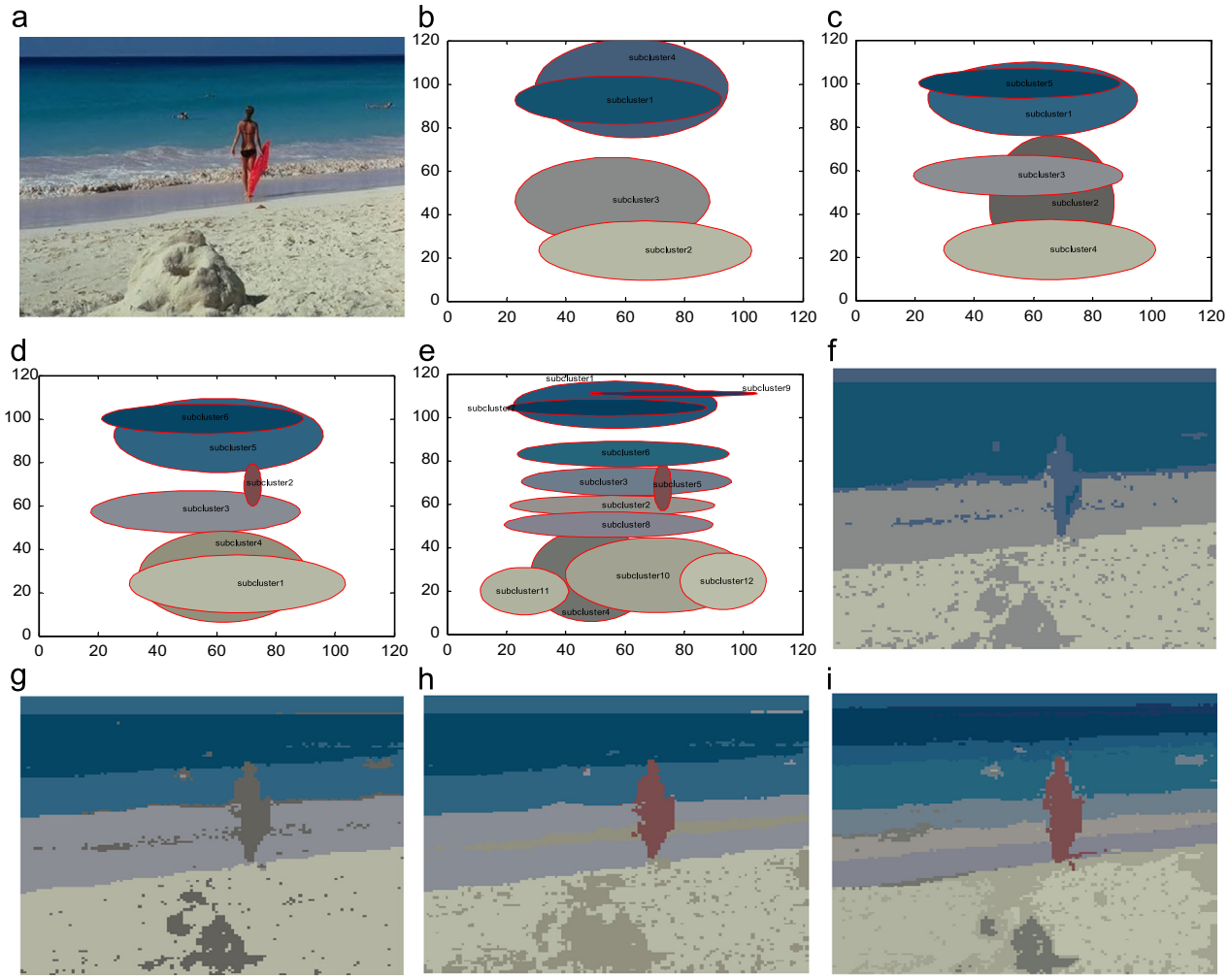


Fig. 1. The visualization of the GMMs. (a) Original image; (b)–(e) the GMMs with 4, 5, 6, 12 components; and (f)–(i) the corresponding image reconstruction results of (b)–(e).

co-sharing level to form bigger clusters. Our work follows the general framework in [17], however, we merge subclusters by spectral clustering to improve the merged results.

In spectral clustering, two issues need to be considered: the measurement of the similarity between subclusters and the construction of the similarity matrix. Kullback–Leibler (KL) divergence is a widly used non-symmetric measure of the difference between two probability distributions. Given those Gaussian components obrained by EM-IRC, which are represented by their mean vectors and covariance matrices, KL divergence can measure the similarities among them. On the other hand, the performance of spectral clustering is also highly dependent on the structure of the similarity matrix. If the similarity matrix is closer to block-structured, the performance will be better [18]. To achieve this goal, we strengthen the block structure of the similarity matrix using a revised Floyd's algorithm developed from the Zadeh's operations [19].

In this paper, a spectral clustering method based on GMMs for image segmentation is proposed. First, we model the joint distribution of color and position features of the image with a mixture of Gaussians, and use the EM-IRC algorithm to estimate the number of Gaussian components and the model parameters. Second, we measure the similarities among the Gaussian components using KL divergence, and use a revised Floyd's algorithm to improve the structure of the similarity matrix. Finally, the Gaussian components are merged based on this improved similarity matrix using spectral clustering, and the corresponding image segments are regrouped to obtain the final segmentation result.

The proposed method has the following attractive characteristics. First, we propose a two-step algorithm, producing Gaussian components by the EM-IRC algorithm in the first step, which are then merged by spectral clustering in the second step. The EM-IRC algorithm not only produces much fewer Gaussian components compared to the traditional EM-BIC method, but also is robust to noise and alleviates the problem that the EM algorithm is sensitive to initialization. Second, the improved similarity matrix generated by the Floyd's algorithm is notably block-structured, which can ensure the effective extraction of spectral features. Finally, since spectural clustering is performed on the Gaussian components (i.e., the image segments) but not the image pixels, the similarity matrix is much smaller and thus can be easily decomposed. All these characteristics make our method more robust and suitable for image segmentation.

The rest of this paper is structured as follows. A brief introduction of GMM-EM clustering and spectral clustering is given in Section 2. The new Image Reconstruction Criterion for finding an optimal number of Gaussian components and the novel image segmentation method are presented in Section 3. Section 4 describes our experimental results, and Section 5 summarizes and concludes this paper.

2. Background

2.1. Gaussian mixture models and EM algorithm

The basic idea of Gaussian Mixture Models (GMMs) is that the characteristics of the samples are described by M states, and each state will be approximated by a Gaussian distribution. If the characteristics of a sample are described as the random variable x, then the kth state ($0 < k \le M$) distribution of the sample can be expressed as follows:

$$p_k(x_t|k,\phi_k) = \frac{1}{(2\pi)^{n/2}|\sum_k|^{1/2}} e^{-\frac{1}{2}(x_t-\mu_k)^T\sum^{-1}(x_t-\mu_k)} \tag{1}$$

where $\phi_k = \{\mu_k, \sum_k\}$ denotes the expectation and covariance matrices of the kth state. For M states, the probability density function of the Gaussian mixture is

$$p(x_t|\Phi) = \sum_{k=1}^{M} \pi_k p_k(x_t|k,\phi_k) \tag{2}$$

where π_k is the weight of the kth state in the Gaussian mixture model. It represents the prior probability of the kth Gaussian distribution, and $\sum_{k=1}^{M}\pi_k = 1$. Φ is the set of all parameters, defined as $\Phi = \{\pi_1, \pi_2 \ldots \pi_M;\quad \phi_1, \phi_2 \ldots \phi_M\}$, and all parameters need to be estimated from the observed values of x.

GMM parameters are usually estimated using the EM algorithm. The EM algorithm first calculates an optimal lower bound according to current value of model parameters (E-step), then maximizes this bound to obtain an improved parameter estimation (M-step). The process is repeated until the difference of the log likelihood between two consecutive iterations is below a threshold. The EM algorithm for GMMs consists of the following two steps:

E-step:

$$p(i|x_l,\Phi^h) = \frac{\pi_i p(i|x_l,\Phi^h)}{\sum_{j=1}^{M}\pi_j p_j(x_l|\Phi^h)}, \quad i = 1,\ldots,M \tag{3}$$

M-step:

$$\pi_i = \frac{1}{N}\sum_{l=1}^{N} p(i|x_l,\Phi^h) \tag{4}$$

$$\mu_i = \frac{\sum_{l=1}^{N} x_l p(i|x_l,\Phi^h)}{\sum_{l=1}^{N} p(i|x_l,\Phi^h)} \tag{5}$$

$$\sum_i = \frac{\sum_{l=1}^{N} p(i|x_l,\Phi^h)(x_l-\mu_i)(x_l-\mu_i)^T}{\sum_{l=1}^{N} p(i|x_l,\Phi^h)}, \quad i = 1,\ldots,M \tag{6}$$

where $p(i|x_l,\Phi^h)$ indicates the membership probability of x_l in the ith distribution, Φ^h represents the old parameter values, π_i, μ_i and $\sum_i$ are the new parameter values.

2.2. Spectral clustering

Spectral clustering transforms the original dataset into a new one in a lower-dimensional eigenspace by utilizing eigenvalues and eigenvectors of a similarity matrix derived from the dataset. Then the traditional clustering algorithms can be performed on this new dataset to obtain the final clustering result. In spectral clustering, a neighborhood graph on the data points is first constructed based on some criteria, such as the fully connected graph or the K-nearest neighbor graph. A weighted similarity matrix $S \in R^{N\times N}$ is then defined, whose (i,j) element s_{ij} reflects the similarity between data point x_i and x_j. The measuring of the similarity between samples is still an open problem in spectral clustering, and Gaussian kernel function is adopted in most references to measure the similarity of samples. Specifically,

$$s_{ij} = \begin{cases} exp(-\|x_i - x_j\|^2/2\sigma^2) & i \ne j \\ 0 & i = j \end{cases} \tag{7}$$

where $\|\cdot\|$ denotes the Euclidean norm, and σ is the Gaussian kernel parameter. Many studies show that this function is sensitive to σ, and [20] provided a solution with certain limitations.

Given a neighborhood graph with similarity matrix S, we compute the first k generalized eigenvectors $z_1, \ldots, z_k$, corresponding to the k smallest eigenvalues, of the generalized eigenvalue

problem

$$(D-S)z=\lambda Dz \tag{8}$$

where D is an $N\times N$ diagonal matrix with $D_{ii}=\sum_j s_{ij}$, λ is the eigenvalue and z is the corresponding eigenvector. Then a traditional clustering algorithm (such as K-means, FCM) is performed on the row vectors of $Z=[z_1,\ldots,z_k]\in\Re^{N\times k}$ to obtain the final clustering result.

3. The new image segmentation method

3.1. GMM-based image representation and the EM-IRC algorithm

In our framework we chose the color and position of the pixels as features (other features such as texture can be easily adopted too). Like the earlier works ([22,23]) we extract the color features in the RGB color space, because distances in this space are meaningful [24]. So the images were represented with a set of data points in a 5-dimensional (3 for color and 2 for position) feature space. These data points are modeled by a Gaussian mixture model. The model parameters are estimated using EM, as described in Section 2.1. We propose an Image Reconstruction Criterion for finding the optimal number of components of the GMM. In the proposed EM-IRC algorithm, the way to estimate the number of Gaussian components t is to start with 1 component and increase the number of components. At each t, the reconstructed image obtained from the GMM is compared to the original image to compute the image reconstruction error. The optimal value of t is obtained when the reconstruction error reaches the threshold.

Let $H(:,:,k),\quad (k=1,2,3)$ be the RGB matrix of the original image, whose size is $M\times N\times 3$, and $G_t(:,:,k),\quad (k=1,2,3)$ the reconstructed image, where t is the number of Gaussian components. Note that the reconstructed image is obtained by replacing the image pixel color value with the mean value of the Gaussian component to which the pixel belongs to. We define the image reconstruction error as the color difference between the reconstructed image and the original image

$$Delta(t)=\sum_{k=1}^{3}\sum_{i=1}^{M}\sum_{j=1}^{N}|G_t(i,j,k)-H(i,j,k)| \tag{9}$$

Since the EM algorithm is sensitive to the initialization, for each t, the EM algorithm is performed 20 times, which gives us 20 different GMMs independently, and $Delta(t)$ is the average difference. The Image Reconstruction Criterion (IRC) is normalized as follows:

$$IRC(t)=\frac{|Delta(t)-Delta(t+1)|}{Delta(t+1)} \tag{10}$$

when $IRC(t)$ reaches a certain threshold, and t is the optimal number value of Gaussian components.

Fig. 1(f)–(i) shows the reconstructed images by GMMs with 4, 5, 6, 12 components, correspondingly. From these figures, we can see that the reconstructed image is closer to the original one when the number of Gaussian components is larger. Table 1 gives the values of $IRC(t)$ and the minimum, maximum, mean and variance of $Delta(t)$ when t is 3 to 12. The threshold is empirically chosen as 0.005, and the optimal number of Gaussian components is therefore 8 in this case.

We also visualize the GMM model in Fig. 1. The components are represented by a uni-colored region (blob) as shown in Fig. 1(b)–(e). We projected the 5-dimensional components into a 2-dimensional space for visualization and assigned each pixel of the original image to the most probable two-dimensional Gaussian component (i.e., the component in which it has the largest likelihood). The color of the region is computed from the mean of the corresponding data points.

Because the EM-IRC algorithm is based on the reconstruction of the image, the number of Gaussian components chosen by the EM-IRC algorithm is generally greater than the number of image segments perceived by humans. As showed in Fig. 2, the optimal number of Gaussian components automatically chosen by the EM-IRC algorithm is 10, but the desired numbers of image segments is 3 (here we do not try to separate the two elephants). Fig. 2(b) shows

Table 1
The $IRC(t)$ and the minimum, maximum, mean and variance of $Delta(t)$.

$Delta(t)$	3	4	5	6	7	8	9	10	11	12	
Min	9.1727	6.1866	5.8544	6.1742	5.5503	5.6019	5.0743	5.1857	5.3228	5.0241	$\times e^5$
Max	9.1727	8.1122	9.1950	6.4140	6.5902	6.5235	7.2216	6.4601	6.7075	5.9825	$\times e^5$
Mean	9.1727	7.1649	6.9605	6.3376	6.2998	5.9578	6.1384	5.9216	5.9323	5.6363	$\times e^5$
Var	0.0000	5.4371	9.5741	0.0428	1.6087	1.2736	5.3697	1.1643	1.4302	1.0653	$\times e^9$
$IRC(t)$	0.2802	0.0294	0.0983	0.0060	0.5740	0.0022	0.0040	0.0018	0.0525	–	

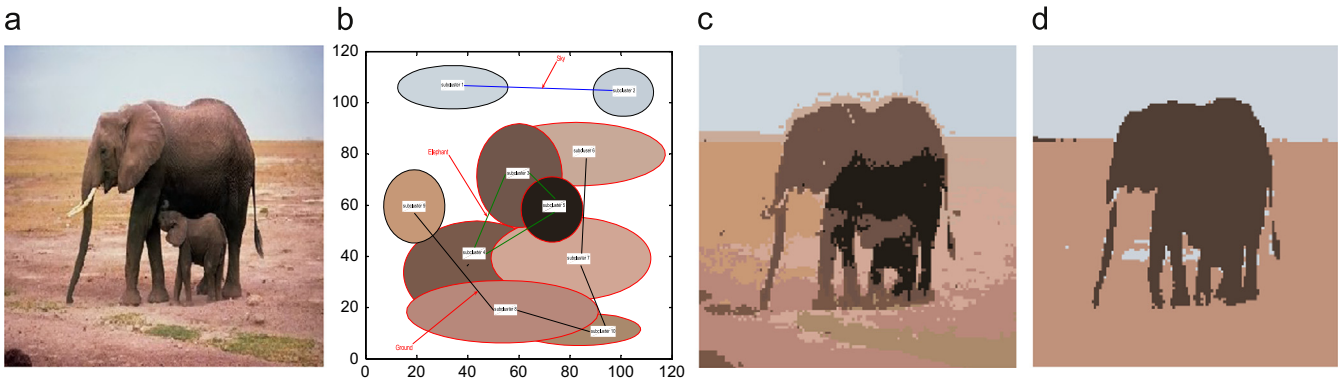


Fig. 2. The visualization of the GMM representation and image segmentation. (a) Original image; (b) the GMM representation with 10 components; (c) the corresponding reconstructed image of 10 components; and (d) the image segmentation result when the desired numbers of image segments is set to 3.

the GMM representation when the component number is 10, and the corresponding reconstructed image is shown in Fig. 2(c). Fig. 2(d) showed the image segmentation result when the target numbers of image segmentation is manually chosen as 3. Therefore, the next issue is how to automatically merge the Gaussian components for image segmentation at object level.

3.2. Spectral clustering for merging Gaussian components

We use spectral clustering to merge the Gaussian components. As shown in Fig. 2(b) and (d), the subclusters 1 and 2 are merged as the sky, the subclusters 3–5 are merged as the elephants, and the subclusters 6–10 are merged as the ground. To achieve this, there are two crucial issues to be considered: the measurement of the similarity between the Gaussian components and the construction of the similarity matrix.

3.2.1. Kullback–Leibler divergence

The KL divergence is a well-known measure of the difference between two distributions [21]. If F and G are continuous random variables, and $f(x)$ and $g(x)$ are their density functions, the KL divergence is defined to be

$$D(F\|G)=\int f(x)log\frac{f(x)}{g(x)}dx \tag{11}$$

Although the KL divergence is not a metric since it is not symmetric and it does not satisfy the triangle inequality, it still has some good properties for being used as a dissimilarity measure [21].

In particular, the KL divergence between two Gaussians has an explicit form based on their means and covariance matrices. Let F and G represent two n-dimensional Gaussian distribution $N(\mu,\Gamma)$ and $N(\nu,\sum)$, respectively, the KL-divergence of them is

$$D(F\|G)=\frac{1}{2}\left\{log\frac{|\sum|}{|\Gamma|}+Tr(\Gamma(\sum^{-1}-\Gamma^{-1}))+(\mu-\nu)^T\sum^{-1}(\mu-\nu)\right\} \tag{12}$$

where $Tr(A)$ denotes the trace of matrix A.

Since most clustering algorithms use symmetric distance, we define the similarity measure in our framework as the resistor average of the KL divergence. Other methods (e.g., arithmetical average, geometrical average) can be used too. The resulting measurement is

$$R(F,G)=\frac{1}{1/D(F\|G)+1/D(G\|F)} \tag{13}$$

where $D(F||G)$is the KL-divergence as showed in Eq. (10), given the GMM parameters estimated by the EM-IRC algorithm.

3.2.2. The improved similarity matrix

Although spectral clustering algorithms are simple and efficient, their performance is highly dependent on the construction of the similarity matrix. According to [18], we can obtain better clustering performance if the similarity matrix is block-structured. Therefore, the generation of a similarity matrix that is closer to the block structure is critical to the success of spectral clustering. To achieve this goal, we try to amplify the block structure of the similarity matrix.

Given the similarities of all the pairs of Gaussian components calculated by the similarity measure, we build the relationship graph of the similarity as, for example, shown in Fig. 2(b). In this figure, there are ten components (subclusters). We defined r_{ij} as the similarity between subcluster i and j. In the similarity relationship graph of subclusters, the set of all subclusters in their feature space can be represented as the weighted undirected graph $G=\{V, E, S\}$. V is the set of vertices and E is the set of edges. The weight r_{ij} on each edge is a function of the similarity between vertices v_i and v_j. When constructing similarity graph the goal is to model the local neighborhood relationships between subclusters.

From Fig. 2(b), we can know that there are more than one path from subcluster i to j. In essence, the block structure of similarity graph is the graph conductivity. Graph conductivity is defined following conductivity for electrical networks, i.e., the conductivity of two points depends on all paths between them. For example, the improved similarity between subcluster 1 and 2 in Fig. 2(b) can be formulated as

$$r_{12}^{*}=max\{r_{12},min(r_{13},r_{32}),\ldots,min(r_{13},r_{34},r_{42}),\ldots, min(r_{13},r_{34},r_{45},r_{56,}r_{62}),\ldots\} \tag{14}$$

We can see that the essence of improving the similarity matrix is to find the maximum among all the possible paths as the similarity between the subclusters and choose the minimum in a path as the weight of itself. Obviously, to find the improved similarity of two subclusters is an optimal path problem in the similarity graph.

In this paper, the Floyd's algorithm is adopted to establish the improved similarity among subclusters. That is, applying max and min operators of the Zadeh's operations between two subclusters to get the optimum similarity between them [19]. The Floyd's algorithm is a classical method to find the shortest path in a graph. More precisely, insert vertex v_k as an intermediate (all vertices in the path except for the starting vertex and the terminating vertex) between any two vertices and compare the known shortest distance from v_i to v_j with the possible distance from v_i to v_j which comes out while the existence of the inserted vertex v_k. Displace the original value with the smaller one to get the new weighted adjacency matrix. Repeat the procedure when all the vertices have been in turn inserted between and as the intermediate, the new weighted adjacency matrix can tell every shortest distance from one vertex to another. It is the distance matrix of graph G.

To subclusters i and j, V is the set of all the subclusters, R_1 is the original similarity between subcluster i and V, R_2 is the original similarity between subcluster j and V. If an optional element of V is set as the intermediate vertex, the improved similarity of subclusters i and j is $R_1\circ R_2$

$$(R_1\circ R_2)(subclusteri,subclusterj)\triangleq\underset{v\in V}{\vee}(R_1(subclusteri,v)\wedge R_2(v,subclusterj)) \tag{15}$$

If the number of samples is n, the optimal similarity matrix R_{opt} of inserting one intermediate vertex can be derived from original similarity matrix R. Let $R=(r_{ij})_{n\times n}$,$R_{opt}=(b_{ij})_{n\times n}$, then R_{opt} is

$$R_{opt}=R\circ R=(b_{ij})_{n\times n} \tag{16}$$

where $b_{ij}=\vee_{k=1}^{n}(r_{ik}\wedge r_{kj})$.

In this paper, the improved similarity matrix by the Floyd's algorithm is regarded as the similarity matrix to generate the Laplacian matrix used in spectral decomposition, which has a clear block structure.

3.2.3. The component merging method

The improved similarity matrix can reflect the similarity among subclusters more objectively than the matrix simply derived from the Gaussian kernel function. Given the improved similarity matrix, we need to merge those subclusters using spectral clustering. In the graph Laplacian theory, the eigenvectors corresponding to the k smallest eigenvalues contain the main classification information. In this paper, we select the eigenvectors corresponding to the second and third minimal eigenvalues as the spectral features based on the method in [3], and FCM is used to cluster the

spectral features in the end. The entire proposed method is outlined as follows:

Algorithm 1. The proposed method

Input: The input image and the desired number of image segments, the distance function2 ln n, the termination criterion ζ.

Output: The segmentation of the image.

1. Extract color and position features from the input image.
2. Perform the EM-IRC algorithm to get the GMM parameters using Eqs. (3)–(6). The number of components of GMM is generated by thresholding Eq. (10).
3. Consider the GMM parameters as the features of spectral clustering. Get the similarity matrix of the components R via the KL divergence by using Eq. (13).
4. Adopt the Floyd's algorithm in Eq. (16) to modify the original similarity matrix R into the improved similarity matrix R_{opt}.
5. Build the Laplacian matrix according to Eq. (8). Decompose the Laplacian matrix to get the spectral features.
6. Generally, the eigenvectors corresponding to the second minimal eigenvalues and the third minimal eigenvalues are chosen to be the spectral features. Cluster spectral features by using the FCM algorithm to get the merging results.
7. Reclassify the image pixels using the merged components for final image segmentation.

3.3. *Analysis on computational complexity*

In this section, we study the computational complexity of our method. In our method, image is first divided into small regions modeled by a Gaussian Mixture Model, and the problem of partition pixels is converted to a problem of merging components, the computational load is low.

In the EM part, the algorithm runs iteratively to refine the subclusters. The time complexity of matrix inversion, matrix determinant, and matrix decomposition is $O(q^3)$, where q is the feature dimension; thus, the time complexity of the matrix operations for h subclusters is $O(hq^3)$. Besides, the time complexity of computing the membership probabilities of each data object for each subclusters is $O(q^2)$. For all data objects and all subclusters, the total time complexity of EM clustering is $O(ihnq^2+ihq^3)$, where n is the number of samples, and i is the number of iteration. In the merge part, the time complexity computing the initial difference matrix is $O(h^2q^3)$, where $O(q^3)$ is time spent on computing the distance between each pair of subclusters; the time complexity of transitive closure is $O([\log_2 h]+1)$, and the time complexity of the spectral clustering algorithm based on subclusters is $O(h^3)$. Thus, the total runtime of the second phase is $O(h^3)+O(h^2q^3)+O([\log_2 h]+1)$. So our method's time complexity is $O(h^3)+O(h^2q^3)$.

4. Experiments and analysis

In this section, the proposed method was tested on the IRIS dataset and some real-world images. We first apply it to the IRIS dataset classification problems. In the second experiment, we compare the effect of using the original and improved similarity matrix in spectral clustering for image segmentation. In order to visualize the performance, the relationship graphs of the original similarity matrix and the improved similarity matrix of components are shown. Lastly, the real image segmentation and robustness to noise of our method is compared with the conventional EM algorithm, FCM and FGFCM [25], so that the superiority of robustness and efficiency of our method can be clearly seen.

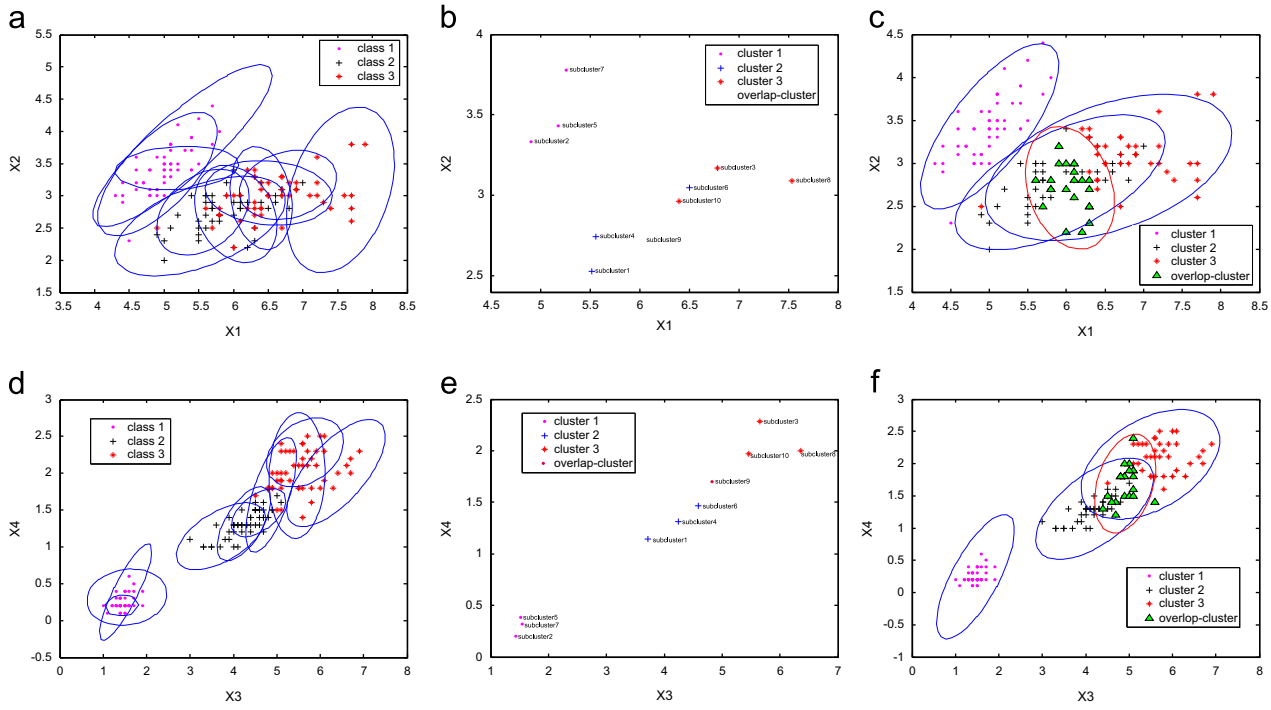


Fig. 3. Classification of the IRIS dataset. (a) The subclusters constructed by the EM-IRC algorithm in X1, X2 attributes; (b) the centers of all subclusters in X1, X2 attributes; (c) the final result by our method in X1, X2 attributes; (d) the subclusters constructed by the EM-IRC algorithm in X3, X4 attributes; (e) the centers of all subclusters in X3, X4 attributes; and (f) the final result by our method in X3, X4 attributes. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

4.1. Classification of the IRIS dataset

In this section, the experiment used the Iris dataset that is available from the UCI machine learning repository. It consists of information on 150 Iris flowers, 50 each from one of three iris species: Setosa (class 1), Versicolour (class 2), and Virginica (class 3). Each flower is characterized by four numeric attributes, denoted as X1–X4. By projecting the data onto two dimensional planes (X1 and X2) and (X3 and X4), respectively, as shown in Fig. 3(a) and (d) (the red, pink, and black dots represent three different classes), we can see that there are two large clusters: one consisting of samples of class 1 and the other consisting of samples from class 2 and 3. The second cluster can further be divided into two subclusters, one composed of samples from class 2 and the other from class 3.

Similar to [17], we set $h=10$ in this experiments. The subclusters constructed by the EM-IRC algorithm are shown in Fig. 3(a) and (d) as blue ellipses. We then plot the centers of all subclusters in Fig. 3(b) and (e) (Class 1 is partitioned into subclusters 2, 5, 7, Class 2 is partitioned into subclusters 1, 4, 6, and Class 3 is partitioned into subclusters 3, 8, 10, and subcluster 9 is an overlapping cluster).

We show the spectral features of all subclusters in Table 2. From Table 2, we can see that the three merged classes (subclusters 2, 5, 7 are merged into Class 1, subclusters 1, 4, 6 are merged into Class 2, subclusters 3, 8, 10 are merged into Class 3, and subcluster 9 is left as a single class). After merging those 10 subclusters using FCM, we get the final clustering result in Table 3. From Table 3, we can see the resulting clusters fit the true class label very well, although there is some overlap between two clusters. The first specie can be correctly clustered without any wrongly assigned data points. For the second and third species, we can see that subcluster 9, which contains 8 samples of Class 2 and 12 samples of Class 3, is the overlap-cluster. The result shows that

Table 2
The spectral features by the improved similarity matrix.

	Subcluste 1	Subcluste 2	Subcluste 3	Subcluste 4	Subcluste 5	Subcluste 6	Subcluste 7	Subcluste 8	Subcluste 9	Subcluste 10
Feature 1	0.0472	0.5728	0.0472	0.0472	0.5728	0.0472	0.5728	0.0472	0.0472	0.0472
Feature 2	−0.3559	0.0000	0.2453	−0.4064	0.0000	−0.1390	0.0000	0.7792	−0.1450	0.0219

Feature 1: the eigenvectors corresponding to the second minimal eigenvalues; feature 2: The eigenvectors corresponding to the third minimal eigenvalues.

Table 3
The final classification result by our method.

	Subcluste 1	Subcluste 2	Subcluste 3	Subcluste 4	Subcluste 5	Subcluste 6	Subcluste 7	Subcluste 8	Subcluste 9	Subcluste 10
Class 1	0	37	0	0	6	0	7	0	0	0
Class 2	14	0	0	12	0	16	0	0	8	0
Class 3	0	0	13	0	0	0	0	10	12	14

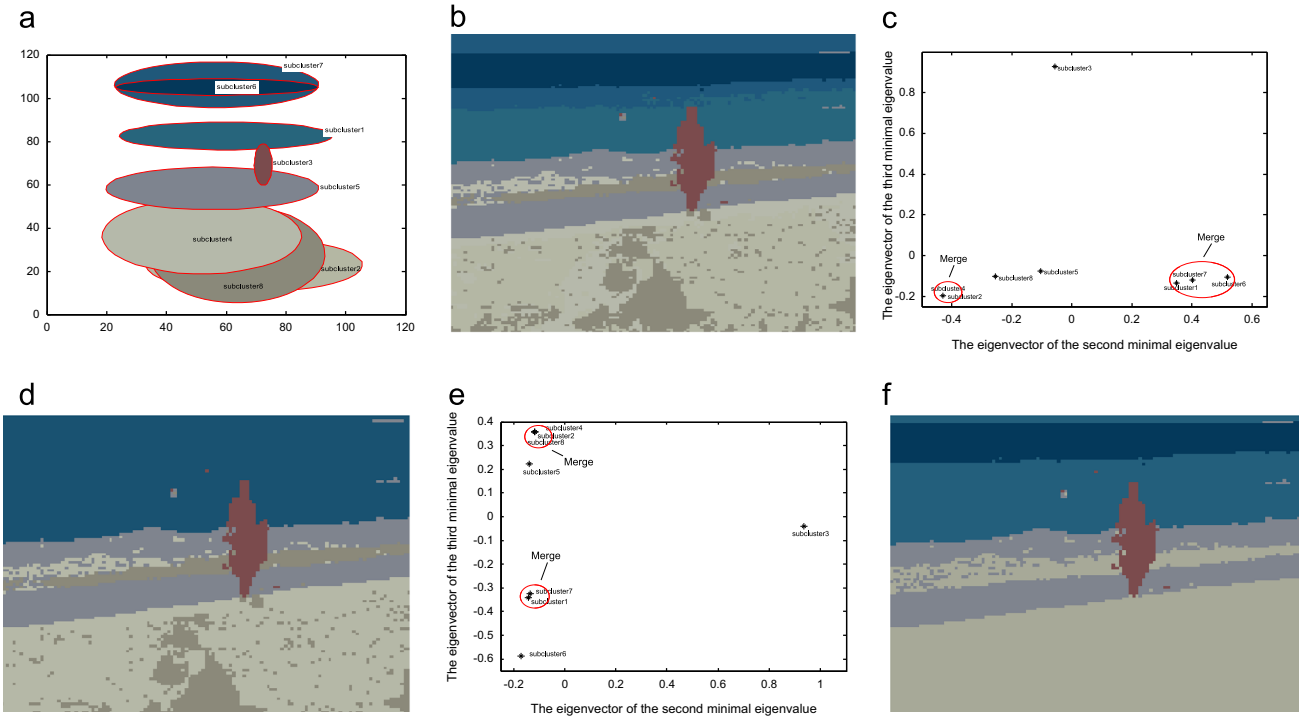


Fig. 4. The real-world image segmentation. (a) The GMM representation when the number of Gaussian components is 8; (b) the corresponding image reconstructed from 8 components; (c) the spectral features generated from the original similarity matrix; (d) the final segmentation result based on the original similarity matrix; (e) the spectral features generated from the improved similarity matrix; and (f) the final segmentation result based on the improved similarity matrix.

our method might have difficulty in clustering highly overlapped clusters.

4.2. Compare the original and improved similarity matrix in image segmentation

In this experiment, we compare the effect of using the original similarity matrix and the improved similarity matrix in spectral clustering for image segmentation.

From Table 1, we can see that the reconstructed image is close enough to the original one when the number of components of GMM is 8. Fig. 4(a) showed the GMM representation when the components numbers of GMM is 8, and the corresponding image reconstruction of 8 components is showed in Fig. 4(b).

When the KL divergence is used as the similarity measure among components, the original and the improved similarity matrix among components are showed in Fig. 5. Fig. 5(a) showed the values of the original similarity matrix; Fig. 5(b) showed the values of the improved similarity matrix; Fig. 5(c) showed the graph of the original similarity matrix; and Fig. 5(d) showed the graph of the improved similarity matrix. From Fig. 5, we can see that the improved similarity matrix is more notably block-structured than the original similarity matrix is. According to the theory of spectral clustering, the improved similarity matrix, which will be used to generate the Laplacian matrix in spectral decomposition, can better ensure the effective extraction of spectral features.

Fig. 4(c) showed the spectral features which are generated from the original similarity matrix, i.e., the eigenvectors corresponding to the second and the third minimal eigenvalues. From Fig. 4(c), subclusters 2, 4 are merged into the same class and subclusters 1, 6, 7 are also merged into another class. Fig. 4(d) showed the final segmentation result based on the original similarity matrix.

Fig. 4(e) showed the spectral features which are generated from the improved similarity matrix. From Fig. 4(e), subclusters 2, 4, 8 are merged into the same class and subclusters 1, 7 are also merged into another class. Fig. 4(f) showed the final segmentation result based on the improved similarity matrix. It can be seen that Fig. 4(f) is clearly better than Fig. 4(d) in terms of image segmentation perceived by humans.

If adopting the conventional EM algorithm with the number of Gaussian components being 5, the result is shown in Fig. 1(e). Obviously, this segmentation result is not satisfactory compared to ours (Fig. 4(f)). Even Fig. 4(d) is arguably better than Fig. 1(e).

4.3. Experiments on noisy images

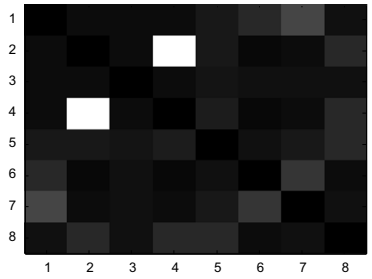
In the following experiments, to investigate the robustness of our algorithms, we firstly test our algorithm, the standard GMM-EM algorithm, FCM, and FGFCM with 4×4 window on a real-world image (Fig. 6(a)), then we test these algorithms on the same image corrupted by Gaussian (0, 0.005) (Fig. 7(a)) and salt–pepper noise (0.025) (Fig. 8(a)), respectively. Figs. 6(b), 7(b), and 8(b) show the GMM representation when the number of components of GMMs is set to 7 for the standard EM algorithm, and the corresponding image segmentation results are shown in Figs. 6(c), 7(c), 8(c). Figs. 6(d), 7(d), and 8(d) show the GMM representation estimated by the EM-IRC algorithm when the number of components of initial GMMs is 12, and the final image segmentation results where 7 clusters are left after merging are shown in Figs. 6(f), 7(f), 8(f). Figs. 6(e), 7(e), and 8(e) show the spectral features of the 12 suclusters. Figs. 6(g), 7(g), and 8(g) show the segmentation results by FCM algorithm and Figs. 6(h), 7(h), and 8(h) show the segmentation results by FGFCM with 4×4 window.

From Fig. 6(c) and (g), one can see that the image segmentation results by standard EM and FCM are not satisfactory. Fig. 6(c)

a

$R=[r_{ij}]_{c\times c}$	1th	2th	3th	4th	5th	6th	7th	8th
1th	0	0.0245	0.0239	0.0251	0.0456	0.0712	0.1159	0.0331
2th	0.0245	0	0.0233	0.4339	0.0464	0.0174	0.0223	0.0732
3th	0.0239	0.0233	0	0.0237	0.0360	0.0279	0.0303	0.0331
4th	0.0251	0.4339	0.0237	0	0.0499	0.0177	0.0231	0.0729
5th	0.0456	0.0464	0.0360	0.0499	0	0.0327	0.0460	0.0733
6th	0.0712	0.0174	0.0279	0.0177	0.0327	0	0.0909	0.0250
7th	0.1159	0.0223	0.0303	0.0231	0.0460	0.0909	0	0.0296
8th	0.0331	0.0732	0.0331	0.0729	0.0733	0.0250	0.0296	0

c



b

$R=[r_{ij}]_{c\times c}$	1th	2th	3th	4th	5th	6th	7th	8th
1th	0	0.0456	0.0360	0.0456	0.0460	0.0909	0.0712	0.0456
2th	0.0456	0	0.0360	0.0729	0.0732	0.0327	0.0460	0.0729
3th	0.0360	0.0360	0	0.0360	0.0331	0.0327	0.0360	0.0360
4th	0.0456	0.0729	0.0360	0	0.0729	0.0327	0.0460	0.0732
5th	0.0460	0.0732	0.0331	0.0729	0	0.0460	0.0456	0.0499
6th	0.0909	0.0327	0.0327	0.0327	0.0460	0	0.0712	0.0331
7th	0.0712	0.0460	0.0360	0.0460	0.0456	0.0712	0	0.0460
8th	0.0456	0.0729	0.0360	0.0732	0.0499	0.0331	0.0460	0

d

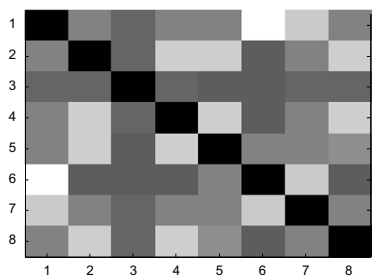


Fig. 5. The original and improved similarity matrix among subclusters in Fig.4(a). (a) The values of the original similarity matrix among subclusters; (b) the values of the improved similarity matrix among subclusters; (c) the gray graph of the original similarity matrix; and (d)the gray graph of the improved similarity matrix.

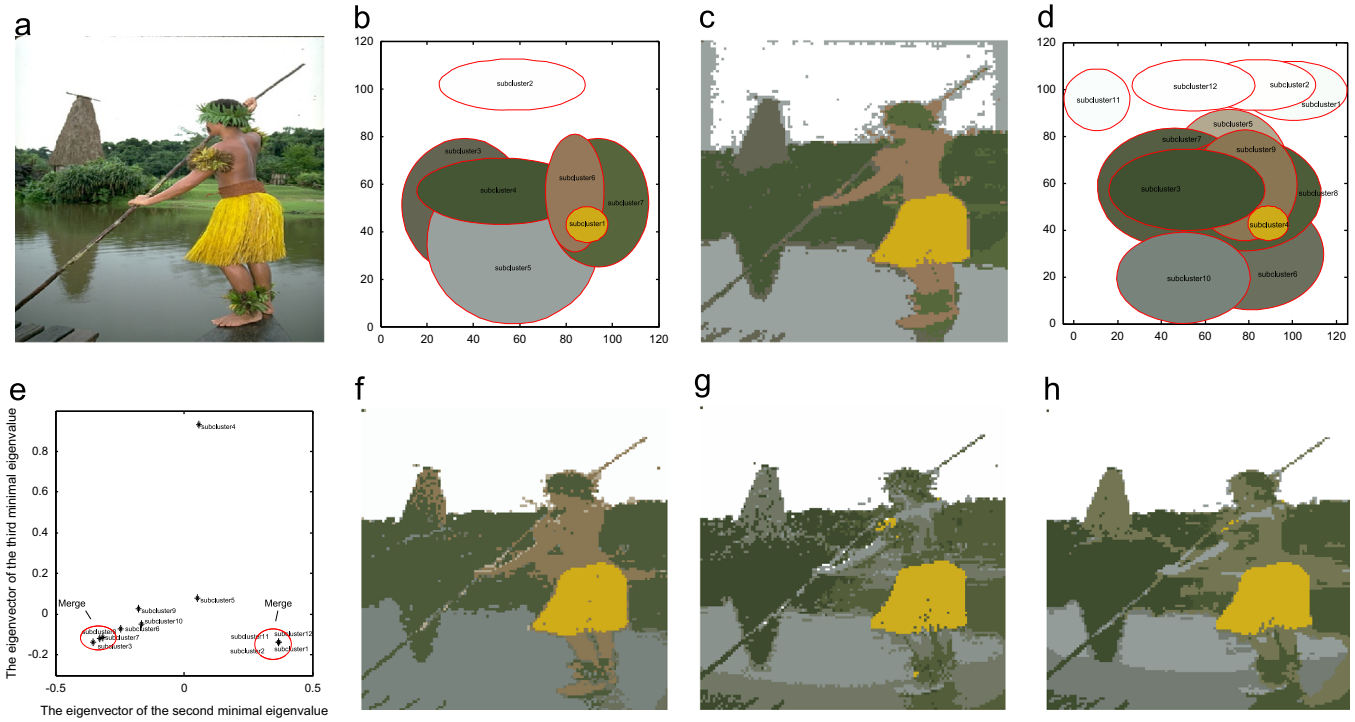


Fig. 6. The segmentation results of the real image. (a) Original image; (b) the GMM representation when the components numbers of GMM is 7; (c) the final segmentation result by EM algorithm with the number of components of GMM is 7; (d) the GMM representation by the EM-IRC algorithm with the number of components of GMM is 12; (e) the spectral features of 12 subclusters; (f) the final segmentation results that images were divided into seven categories by merging subclusters; (g) the segmentation result by FCM algorithm; and (h) the segmentation result by FGFCM with 4×4 window. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

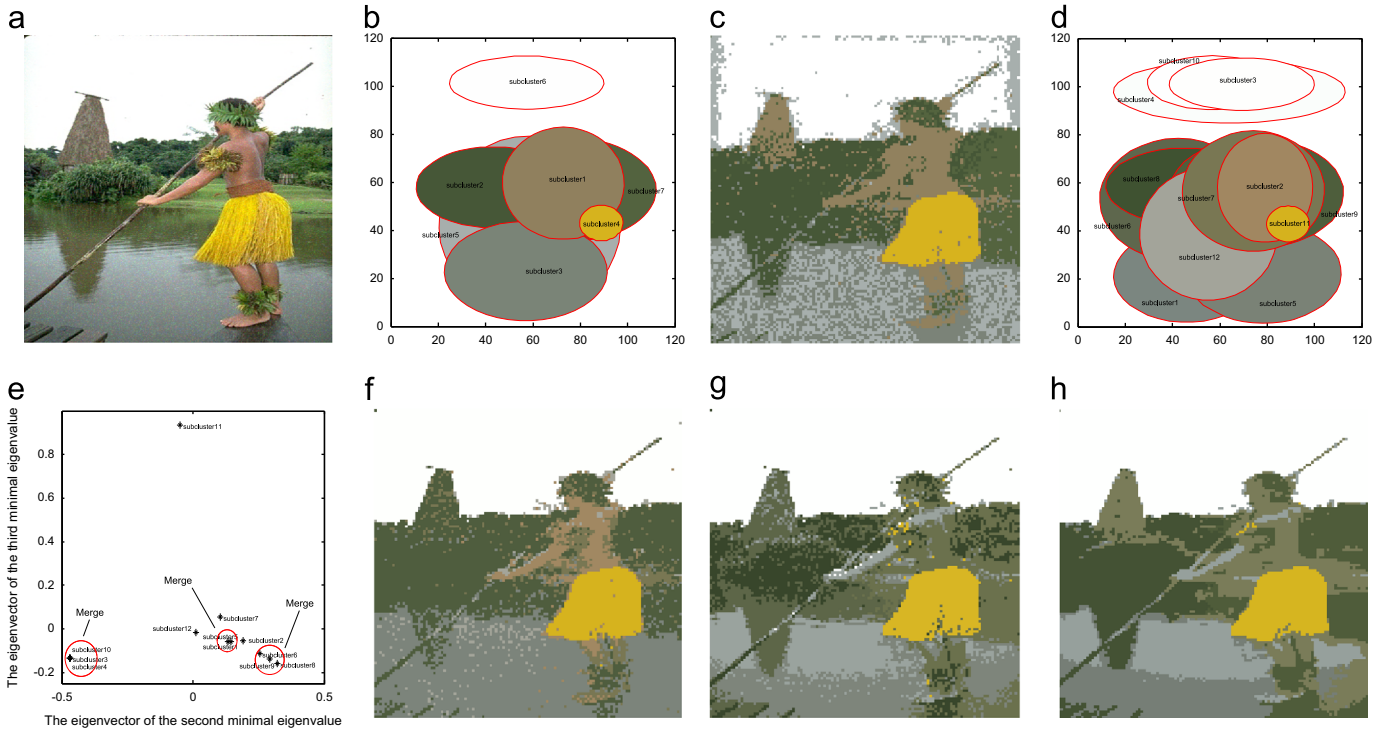


Fig. 7. The segmentation results of the real image corrupted by Gaussian noise (0, 0.005). (a) The noisy image; (b) the GMM representation when the components numbers of GMM is 7; (c) the final segmentation result by EM algorithm with the number of components of GMM is 7; (d) the GMM representation by the EM-IRC algorithm with the number of components of GMM is 12; (e) the spectral features of 12 subclusters; (f) the final segmentation results that images were divided into seven categories by merging subclusters; (g) the segmentation result by FCM algorithm; and (h) the segmentation result by FGFCM with 4×4 window.

shows that some of the sky was divided into the lake, while Fig. 6(g) shows that part of the human being was divided into the lake. In Fig. 6(h), FGFCM obtains better results than FCM algorithm, but worse than our method. From Fig. 6(e), the subclusters 3, 7, 8 are merged into same class (green background) and the subclusters 1, 2, 11, 12 are also merged into another class (sky).

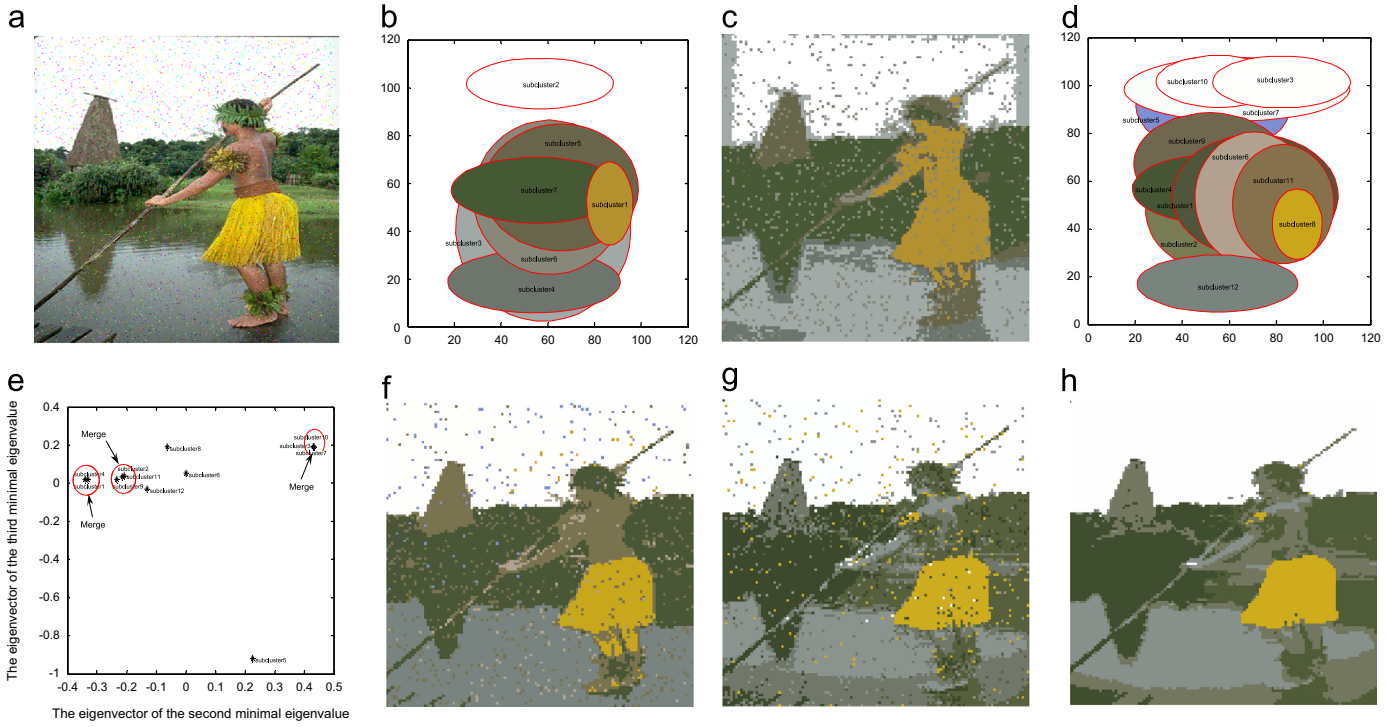


Fig. 8. The segmentation results of the real image corrupted by salt–pepper noise (0.025). (a) The noisy image; (b) the GMM representation when the components numbers of GMM is 7; (c) the final segmentation results by EM algorithm with the number of components of GMM is 7; (d) the GMM representation when the EM-IRC algorithm with the number of components of GMM is 12; (e) the spectral features of 12 subclusters; (f) the final segmentation results that images were divided into seven categories by merging subclusters; (g) the segmentation result by FCM algorithm; and (h) the segmentation result by FGFCM with 4×4 window.

Fig. 6(f) shows that our method obtains the best segmentation results and is more effective than the other methods.

For the images corrupted by Gaussian and salt–pepper noise (Figs. 7(a) and 8(a)), similar conclusions can be drawn. The original EM algorithm is the most sensitive to the noise. As Figs. 7(c) and 8(c) show, there are many snowflakes in the segmentation results. FCM (Figs. 7(g) and 8(g)) is slightly less affected by the noise but the results are still very noisy. From Figs.7(h) and 8(h), one can see that FGFCM deals with the noise much better, but the results are not satisfactory because there are a lot of wrongly assigned regions on the human and the lake. Figs.7(f) and 8(f) clearly show that the results of our algorithm is overall the best, although the remaining noise level is slightly higher than that of FGFCM. Our method might fail when the noise reaches a certain high level.

5. Conclusions

In this paper, a novel image segmentation method using spectral clustering and GMMs is presented. We first use a newly proposed EM-IRC algorithm to estimate the GMM representing the image, which can better preserve image details while reducing the number of Gaussian components to a controllable level. In the second phase, we use spectral clustering on the Gaussian components to obtain the final segmentation, which is highly efficient and effective. The proposed method achieves satisfying segmentation results on the complex real-world images. In the future, we plan to further improve the robustness of our method to noise.

Acknowledgments

This work is supported by National Natural Science Foundation of China under Grants 61303116, 61105014 and 61179032.

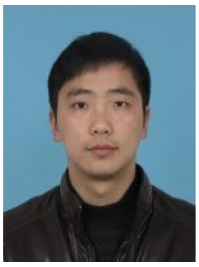
References

[1] Y.H. Wang, Tutorial: Image Segmentation. Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan.
[2] S.C. Chen, D.Q. Zhang, Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 34 (4) (2004) 1907–1916.
[3] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern. Anal. Mach. Intell. 22 (8) (2000) 888–905.
[4] H.Q. Liu, F. Zhao, L.C. Jiao, Fuzzy spectral clustering with robust spatial information for image Segmentation, Appl. Soft Comput. 12 (2012) 3636–3647.
[5] N. Ueda, R. Nakano, Z. Ghahramani, G.E. Hinton, SMEM algorithm for mixture models, Neural Comput. 12 (9) (2000) 2109–2128.
[6] P. Bradley, C. Reina, U. Fayyad, Clustering very large databases using EM mixture models, in: Proceedings of the 15th International Conference on Pattern Recognition (ICPR), 2000, pp. 76–80.
[7] F. Zhao, Fuzzy clustering algorithms with self-tuning non-local spatial information for image segmentation, Neurocomputing 106 (2013) 115–125.
[8] M.A. Balafar, A.R. Ramli, S. Mashohor, Medical brain magnetic resonance image segmentation using novel improvement for expectation maximizing, Neurosciences 16 (2011) 242–247.
[9] K.S. Chuang, H.L. Tzeng, S. Chen, J. Wu, T.J. Chen, Fuzzy c-means clustering with spatial information for image segmentation, Comput. Med. Imaging Gr. 30 (2006) 9–15.
[10] H. Tanga, J. Dillensegerb, X.D. Baoa, L.M. Luoa, A vectorial image soft segmentation method based on neighborhood weighted gaussian mixture model, Comput. Med. Imaging Gr. 33 (2009) 644–650.
[11] F. Rahman, Z. Behnam, Image segmentation using gaussian mixture model, Int. J. Eng. Sci. (IUST) 19 (1) (2008) 29–32.
[12] C. Fraley, A.E. Raftery, Model-based clustering, discriminant analysis, and density estimation, J. Am. Stat. Assoc. 97 (458) (2002) 611–631.
[13] C. Biernacki, G. Celeux, G. Govaert, Assessing a mixture model for clustering with the integrated completed likelihood, IEEE Trans. Pattern Anal. Mach. Intell 22 (7) (2000) 719–725.
[14] J.P. Baudry, A.E. Raftery, G. Celeux, K. Lo, R. Gottardo, Combining mixture components for clustering, J. Comput. Gr. Stat. 19 (2) (2010) 332–353.
[15] C. Hennig, Methods for merging gaussian mixture components, Adv. Data Anal. Classif. 4 (2010) 3–34.
[16] A. Pastore, S.F. Tonellato, A merging algorithm for Gaussian mixture components, University Ca'Foscari of Venice, Department of Economics Research Paper series No. 04/WP/2013, 2013.
[17] A. Tung, X. Xu, B. Ooi, CURLER: Finding and visualizing nonlinear correlation clusters, in: Proceedings of ACM SIGMOD International Conference on Management of Data, 2005, pp. 467–478.

[18] I. Fischer, J. Poland, Amplifying the Block Matrix Structure for Spectral Clustering, Technical Report No. IDSIA-03-05, Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA), Switzerland, 2005.
[19] K.T. Shukla, Fuzzy floyd's algorithm to find shortest route between nodes under uncertain environment, Int. J Math. Comput. Appl. Res. (IJMCAR) 3 (5) (2013) 43–54.
[20] Y.N. Andrew, I.J. Michael, W. Yair, On Spectral Clustering: Analysis and algorithm, Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA (2002) 849–856.
[21] J. Goldberger, S. Gordon, H. Greenspan, An efficient image similarity measure based on approximations of KL-divergence between two Gaussian mix, Proc. Int. Conf. Comput. Vis. (ICCV) (2003) 487–493.
[22] C. Carson, S. Belongie, H. Greenspan, J. Malik, Blobworld: image segmentation using expectation -maximization and its application to image querying, IEEE Trans. Pattern Anal. Mach. Intell. 24 (8) (2002) 1026–1038.
[23] M. Hanmandlu, O.P. Verma, S. Susan, V.K. Madasu, Color segmentation by fuzzy co-clustering of chrominance color features, Neurocomputing 120 (2013) 235–249.
[24] G. Wyszecki, W. Stiles, Color science, concepts and methods, quantitative data and formulae, Wiley, 1982.
[25] W. Cai, S. Chen, D. Zhang, Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation, Pattern Recognit. 3 (2007) 825–838.

Shan Zeng received the B.E. degree in mechanical engineering in 2003 and the M.E. degree in mechatronics engineering in 2009 from Wuhan Polytechnic University. He joined Huazhong University of Science and Technology, where he received the Ph.D. degree in pattern recognition and intelligent systems in 2012. He is currently a faculty member in Wuhan Polytechnic University, Wuhan, China. His research interests include computationally modeling of biological vision perception and applications in computer vision, image and video based object segmentation and recognition, medical image processing and analysis, statistical learning models and fuzzy clustering models.

Rui Huang received the B.S. degree in computer science from Peking University in 1999. After earning the M.E. degree in intelligent systems from Chinese Academy of Sciences in 2002, he joined Rutgers University, where he received the Ph.D. degree in computer science in 2008 and worked as a postdoctoral research associate. He is currently a faculty member in Huazhong University of Science and Technology. His research interests are deformable models, graphical models, image and video based object segmentation and recognition, shape analysis, etc.



Kang Zhen received the B.Sc. degree in Information and Computing Science from Wuhan Polytechnic University in 2010, and the M.Sc. in Mechanical and Electronic Engineering from Wuhan Polytechnic University in 2013. Currently, he is an Experimentalist at Wuhan Polytechnic University. His current research interests fall primarily into the areas of Data Mining, Machine Learning.

Nong Sang graduated from Huazhong University of Science and Technology and received his B.E. degree in computer science and engineering in 1990, M.S. degree in pattern recognition and intelligent control in 1993, and Ph.D. degree in pattern recognition and intelligent systems in 2000. He is currently a professor at the Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China. His research interests include computationally modeling of biological vision perception and applications in computer vision, image analysis and object recognition based on statistical learning, medical image processing and analysis, interpretation of remote sensing images, and intelligent video surveillance.