http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (22): 6763-6769, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.6763.6769

An Adaptive Spectral Clustering Algorithm for Image Clustering and Segmentation

Gu Ruijun, Chen Shenglei and Wang Jiacai
School of Information Science, Nanjing Audit University,Nanjing, 211815, China

Abstract: Based on clustering consistency, the proposed method first emphasizes the flexibility of the local scale, which means each sample has a corresponding scale parameter. Further more, it overcomes the limitations of traditional methods in all samples with the same global scale parameter. Secondly, it stresses the convenience of parameter selection. It can determine the value of a local scale for one sample by computing the sum of weighted distances of N neighbors. Therefore, it can determine the scale parameter automatically. This study illustrates the proposed algorithm not only has inhibition for certain outliers but is able to cluster the data sets with different scales. Finally, experiments on both, artificial data and UCI data sets show that the proposed method is effective. Some experiments were also performed in image clustering and image segmentation to demonstrate its excellent features in application.

Key words: Spectral clustering, spectral graph theory, neighbour adaptive scale, image segmentation

INTRODUCTION

Clustering is one popular data analysis method and has been widely used in pattern recognition, data mining and image processing. DBSCAN (Ester *et al*., 1996), based on variable density of dataset, can deal with datasets with arbitrary shape but an inappropriate choice of parameters may yield poor result. And how to define the density in high dimensional space is also the weakness of DBSCAN. Based on spectral partition theory, spectral clustering algorithms have seen an explosive development over the past years and been successfully used in image segmentation (Shi and Malik, 2000) and image clustering. However, spectral clustering also has some weaknesses (Lihi and Pietro, 1995; Von Luxburg *et al*., 2005) for example, it is sensitive to the datasets which include clusters with distinctly different densities and the parameters of algorithm must be selected cautiously. These limit the application fields of spectral clustering algorithm.

This study proposes an improved spectral clustering algorithm who fully considers the local structure of dataset like DBSCAN. Neighbour adaptive scale simplifies the selection of parameters and makes the improved algorithm insensitive to both density and outliers. To extend its application domain, a robust image clustering using image distance was proposed. Experiments on images returned by search engine show that the proposed method is effective.

The rest of the study is organized as follows. In Section 2, Spectral clustering algorithm is summarized briefly. In Section 3, the proposed adaptive spectral clustering algorithm is described. In Section 4, experiments are presented and the results are discussed. Finally, a conclusion is provided in Section 5.

OVERVIEW OF SPECTRAL CLUSTERING

Spectral clustering can be interpreted in multi-view, such as graph cut theory (Shi and Malik, 2000; Chung, 1997) random walks point of view(Meila and Shi, 2001) and perturbation theory (Von Luxburg, 2007). No matter which theory is based on, spectral clustering in the end is the problem of finding eigenvectors of Laplacian matrix and then clustering eigenvectors into clusters. Normalized Cut (NCut) based-on spectral clustering algorithm (Ng *et al*., 2002) is a representative one (Verma and Meila, 2003). Given a set of n points $X = \{x_1, \ldots, x_n\}$ in $\Re^l$, cluster them into c clusters as follows:

- Computer affinity matrix $A \in \Re^{n \times n}$, in which $A_{ii} = 0$ and

$$A_{ij} = \exp(-\|x_i - x_j\|^2 / \sigma^2)\ (i \neq j) \tag{1}$$

- Construct Laplacian matrix $L = D^{-1/2}AD^{-1/2}$, in which D is diagonal matrix defined as:

$$D_{ii} = \sum_{j=1}^{n} A_{ij}$$

- Find $f_1, \ldots, f_c$, the c largest eigenvectors of matrix L and form the matrix $F = [f1, \ldots, fc] \in \mathbb{R}^{n \times c}$ (normalization when required).
- Normalize the rows of F to be unit length, i.e.:

$$\overline{F}_{ij} = F_{ij} / (\sum_{j=1}^{c} F_{ij}^2)^{1/2}$$

- Treating each row of $\overline{F}$ as a point in $\mathbb{R}^k$, cluster into clusters using k-means algorithm or any other sensible clustering algorithm.
- Assign the original point x_i to cluster j if and only if row i of $\overline{F}$ was assigned to cluster j.

Corresponding Author: Gu Ruijun, School of Information Science, Nanjing Audit University, 211815, Nanjing, China

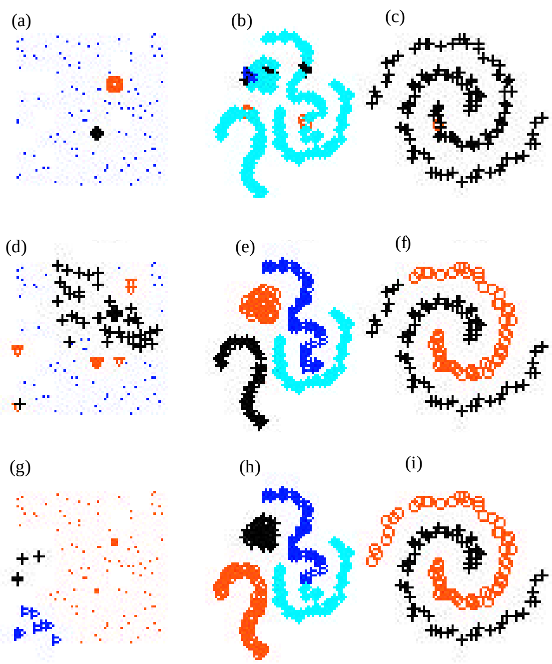

Fig. 1(a-i): Influence of σ to spectral clustering algorithm. (a)-(c), $\sigma = 0.005$. (d)-(f), $\sigma = 0.03$. (g)-(i), $\sigma = 0.08$

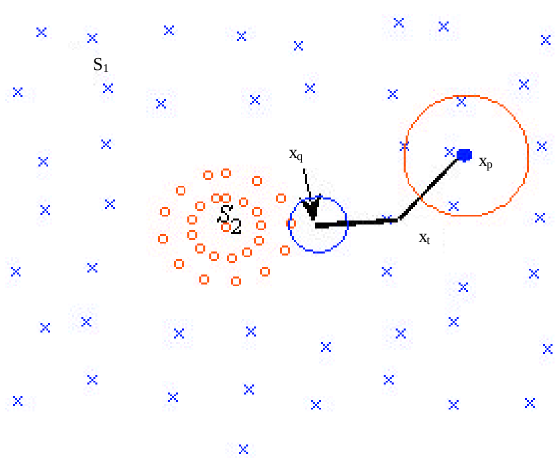

Fig. 2: Principle of neighbour adaptive scale

From above, we can see that scale parameter σ and clusters number c influence the final clustering result. In (Ng *et al.*, 2002) the parameter σ is chosen based on that value of σ that gives the least distorted clusters by repeated experiments. It is obviously that the value of σ depends on specific problem and detecting the range of σ is difficult. Moreover, repeated experiments will increase the compute complexity.

It is can be seen from Fig. 1 that only dataset 1 obtains right result when σ = 0.005 and 0.03 is only suitable for dataset 2 and 0.08 for dataset3. Through a lot of repeated experiments, the optimal intervals of σ for dataset 1~3 are 0.002~0.009, 0.02~0.06, 0.04~0.12 respectively. In conclusion, standard spectral clustering is sensitive to σ and specially not robust to dataset with variable densities. Besides, spectral clustering can find clusters in non-spherical-shape distribution with appropriate σ, however, k-means etc can not do.

Adaptive spectral clustering algorithm based on neighbour adaptive scale

From above analysis we know that maybe a varying σ is more appropriate for spectral clustering. Inspired by DBSCAN, for each point x_i an adaptive scale σ_i is set, which considers the local distribution of neighbours of x_i. Let:

$$\sigma_i = \frac{1}{k}\sum_{m=1}^{k}\|x_i - x_m\| \tag{2}$$

where, σ_i is the average distance between x_i and its k nearest neighbours, called the Neighborhood Adaptive Scale (ASC) of point x_i. However, self-tuning spectral clustering in (Lihi and Pietro, 1995) only considers the kth neighbour (usually k = 7) and tends to be affected by outliers. Similar to (Lihi and Pietro, 1995), affinity between point x_i and x_j is defined as:

$$\bar{A}_{ij} = \exp(-\|x_i - x_j\|^2 / \sigma_i\sigma_j) \tag{3}$$

because scale σ_i varies with neighbour distribution and is adaptive to local structure. This trait enlarges the affinity between two points in the same cluster and reduces that in different clusters as Fig. 2 shows.

In Fig. 2, there are two clusters S_1 and S_2 with different densities. Point x_t, x_p are located in sparse cluster S_1 and x_q in dense cluster S_2.

For convenience, assume that $\|x_t\text{-}x_p\| = \|x_t\text{-}x_q\|$. If constant scale σ is used, according Eq. 1, we can obtain:

$$A_{tp} = \exp(-\|x_t - x_p\|^2/\sigma^2) = \exp(-\|x_t - x_q\|^2/\sigma^2) = A_{tq}$$

which shows that if arbitrary pair points have the same distances they have the same affinities. In fact, x_t and x_p are located in one cluster whereas x_t and x_q in different clusters, so if $A_{tp}>A_{tq}$, spectral clustering will benefit from this. If neighbour adaptive scale is adopted, then $\sigma_p>\sigma_q$ and:

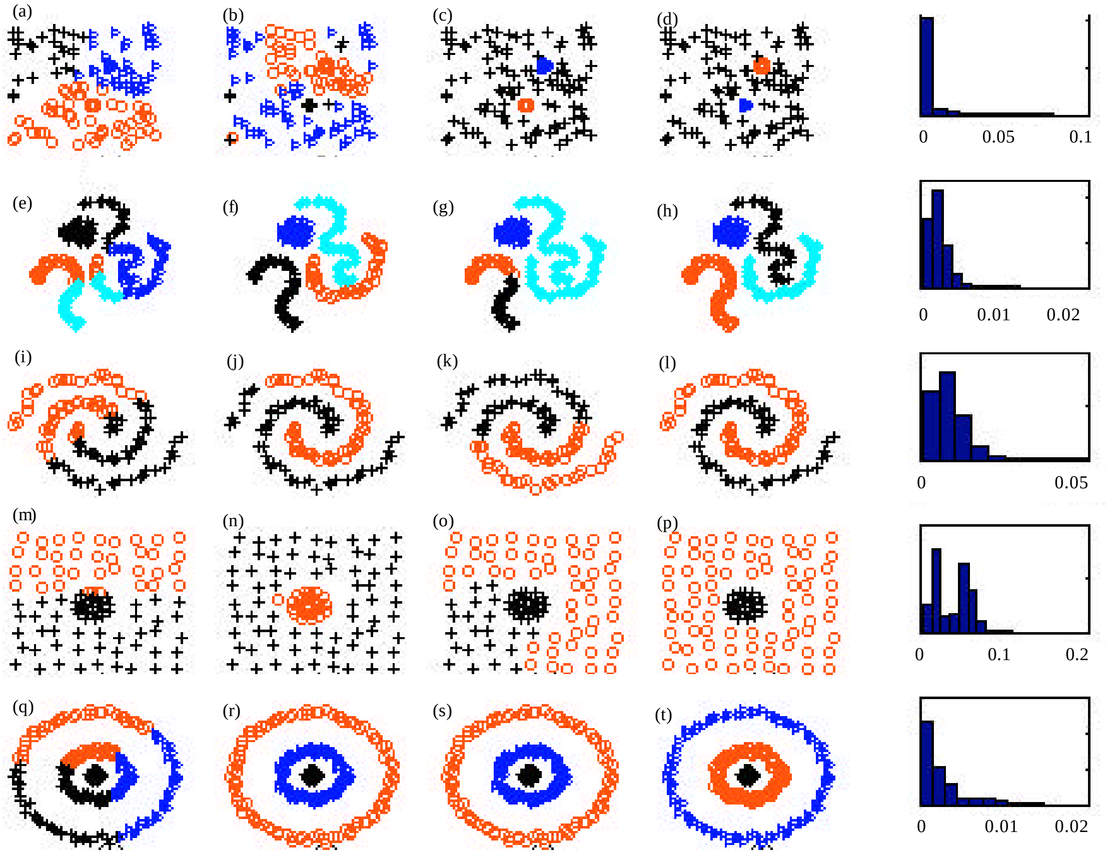

Fig. 3(a-): Diffrenct clusters is distinguished by diffrenct colors. Column 1-4 were obtained by k-means clustering, standard spectral clustering, self-tuning spectral clustering and ASC respectively. The right most column is the statistical histogram of neighbour adaptive scale σ_i

$$\begin{aligned}\bar{A}_{tp} &= \exp(-\|x_t - x_p\|^2/\sigma_t\sigma_p) = \exp(-\|x_t - x_q\|^2/\sigma_t\sigma_p) \\ &> \exp(-\|x_t - x_q\|^2/\sigma_t\sigma_q) = \bar{A}_{tq}\end{aligned}$$

It is obvious that neighbour adaptive scale makes points closer in one cluster and farther in different clusters. Replace the first step of standard spectral clustering with above algorithm, the improved spectral algorithm is named Adaptive Spectral Clustering (ASC). ASC can distinguish clusters with different densities. ASC utilizes average distance to tune affinity by means of the idea of DBSCAN. ASC is resistant to outliers, because σ_i is determined by the average distance, which is more reliable than the nearest neighbour distance or kth neighbour (Lihi and Pietro, 1995) distance. ASC does not have a bias towards a particular cluster shape or size compared with k-means.

Experimental results and analysis: To validate the efficiency of the ASC, we conducted extensive experiments using two publicly available datasets and then compared ASC with k-means, standard spectral clustering and self-tuning spectral clustering.

Experiment with the "Challenging Problems": The data of this experiment mainly come from (Lihi and Pietro, 1995; Ng *et al.*, 2002; Feil and Abonyi, 2007) and some outliers were added to increase the difficulty. Comparative experiments were carried out using k-means clustering, standard spectral clustering, self-tuning spectral clustering and ASC. The parameters of these algorithms were set as follows:

- Standard spectral clustering: $\sigma = 0.03$;
- Self-tuning spectral clustering: $k = 7$; (Lihi and Pietro, 1995)
- ASC: $k = 3$

The clustering result shown in Fig. 3 Column 1 indicates that k-means is unfit for these special clusters because k-means suffers from non-spherical-shape

Table 1: Clustering comparison on UCI database

Dataset	Datasets: Sample size	Categories	Attributes	k-means	Standard spectral clustering: $\sigma = 0.1d_r$	$\sigma = 0.2d_r$	Self-tuning spectral clustering	ASC: k = 3	k = 5
Glass	214	7.0000	9.0000	0.6808	0.6282	0.7061	0.6927	0.6973	0.7084
WDBC	569	2.0000	30.0000	0.7504	0.7554	0.7612	0.7707	0.7973	0.7965
Image segment	210	7.0000	19.0000	0.8448	0.8346	0.8120	0.8374	0.8505	0.8525
Iris	150	3.0000	4.0000	0.8797	0.8797	0.8737	0.8823	0.8859	0.8859
Ionosphere	351	2.0000	34.0000	0.5191	0.5191	0.5191	0.5004	0.5938	0.5938
Average	0.7350	0.7234	0.7344	0.6139	0.7650	0.7674			

distribution. Standard spectral clustering got right partition on two datasets((f), (r)) but failed on other three datasets, which demonstrate that the global scale σ is problem-specific and one single value can only be applied to specific dataset. Similar to standard spectral clustering, self-tuning spectral clustering succeeded on two datasets ((c), (s)) and failed on others. However, ASC obtained accurate partition on all datasets as column 5 shows. Compare Fig. 3 (n) and Fig. 3 (p) and you will find that only one outlier was partitioned by mistake, which indicate that ASC is robust to outliers just because it fully considers the local distribution.

Experiment on the UCI dataset: Next, five real datasets selected from UCI (Blake *et al*., 1998) database were used to evaluate ASC. To evaluate clustering performance quantitatively, we adopt popular Rand index as evaluation, which is defined as follow (Rand, 1971).

The Rand index has a value between 0 and 1, with 0 indicating that the two data clusters do not agree on any pair of points and 1 indicating that the data clusters are exactly the same.

The value of σ is typically set to 10 to 20 percent of the total range of the feature distance function d_r (d_r = max (D)- min (D)) .Without loss of generality, the parameters of these algorithms were set as follows:

- Standard spectral clustering: $\sigma = 0.1d_r$, $\sigma = 0.2d_r$
- Self-tuning spectral clustering: k = 7
- ASC: k = 3, k = 5

Table 1 shows the attributes of UCI database and the experimental results of four clustering algorithms. As can be seen, ASC almost achieved all optimal RIs, moreover, the values are very near when k = 3 or k = 5. We can conclude that ASC is the most stable algorithm of them and insensitive to k. Although the RIs of standard spectral clustering varied widely, at most one RI is larger than that of k-means clustering. Self-tuning spectral clustering is also not stable, because it just considers the seventh neighbour and is sensible to outliers.

Experiment on image clustering: The goal of image clustering is to find a mapping of the archive images into clusters such that the set of clusters provide nearly the same information about the entire image collection (Goldberger *et al*., 2006; Kim *et al*., 2012). To improve the accuracy of image clustering, we introduce a novel image distance to measure the similarity between the images. Then neighbour spectral clustering is performed on the image similarity matrix.

Unlike the traditional Euclidean distance, IMD (Wang *et al*., 2005) takes into account the spatial relationships of pixels. Therefore, it is robust to small perturbation of images. Let vetor x, y be two m by n images, $x = (x_1, x_2, \ldots, x_{mn})$, $y = (y_1, y_2, \ldots, y_{mn})$, the Euclidean distance $d^2_E(x, y)$ is given by:

$$d^2_E(x,y) = \sum_{i=1}^{mn} (x_i - y_i)^2 \tag{4}$$

Suppose $e_1, e_2, \ldots, e_{mn}$ is a base of mn-dimensional image space, the metric coefficients g_{ij} is defined as:

$$g_{ij} = \langle e_i, e_j \rangle = \sqrt{\langle e_i, e_i \rangle}\sqrt{\langle e_j, e_j \rangle} \cdot \cos\theta_{ij} \tag{5}$$

where, <,> is the scalar product and θ_{ij} is the angle between e_i and e_j. Then, the IMD of two images x, y is written by:

$$d^2_{IMD}(x,y) = \sum_{i,j=1}^{mn} g_{ij}(x_i - y_i)(x_j - y_i) = (x-y)^T G(x-y$$

where, $G = (g_{ij})_{mn\times mn}$ is a symmetric positive definite matrix. g_{ij} can be represented as a Gaussian function:

$$g_{ij} = \frac{1}{2\pi\sigma^2}\sum_{i,j=1}^{mn} \exp\{-|P_i - P_j|^2/2\sigma^2\}) \tag{6}$$

where, $|P_i\text{-}P_j|$ is the pixel distance between P_i and P_j. Let σ = 1, then IMD between image x and image y is defined as:

$$d^2_{IMD}(x,y) = \frac{1}{2\pi}\sum_{i,j=1}^{mn} \exp\{-|P_i - P_j|^2/2\}(x_i - y_i)(x_j - y_i)$$

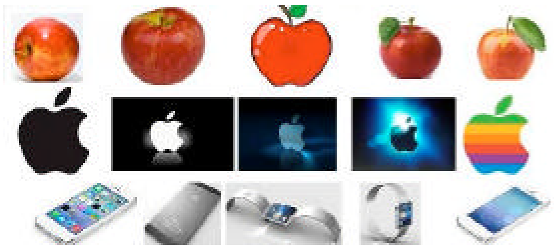

Fig. 4: Image clustering result for keyword "apple" by IMD-ASC

Fig. 5: Image clustering result for keyword "Affrica" by IMD-ASC

Fig. 6: Image clustering result for keyword "Nanjing" by IMD-ASC

To reduce computation complexity, G is decomposed by:

$$G = G?\, G^T = (G?^{1/2} G^T)(G?^{1/2} G) = G^{1/2} G^{1/2} \quad (7)$$

where, Λ is a diagonal matrix whose elements are eigenvalues of G and Γ is an orthogonal matrix whose column vectors are eigenvectors of G. Let $u = G^{1/2}x$, $v = G^{1/2}y$, so:

$$d^2_{IMD}(x, y) = (x - y)^T G^{1/2} G^{1/2} (x - y) = (u - v)^T (u - v)$$

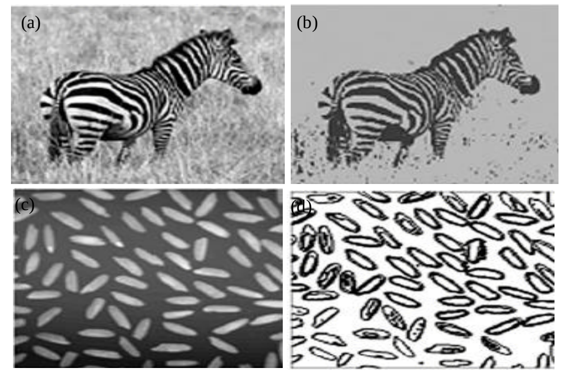

Fig. 7: Gray image segmentaion result by ASC (k = 5)

which has the form of Euclidean distance. Because G is only associated with the size of an image, so can be computed before transform is performed.

Utilizing the property of IMD and spectral clustering, we propose a robust image clustering method named IMD-ASC , by replacing Euclidean distance with IMD. IMD- ASC can be descried as follows (Pughineanu and Balan, 2011):

Step 1: For N images (m by n), compute $G^{1/2}$ according to Eq. 7 and 8

Step 2: For image x_i, x_j (I, j = 1, 2, …, N), let $u_i = G^{1/2}x_i$, $u_j = G^{1/2}x_j$ then:

$$d^2_{IMD}(x_i, x_j) = (x_i - x_j)^T G(x_i - x_j) = (u_i - u_j)^T (u_i - u_j)$$

Step 3: Replace the Euclidean distance in ASC by IMD to perform spectral clustering

Utilizing Google search engine, images were collected by inputting keywords like apple, Africa, Nanjing separately. Then the former 500 images were selected as the input of IMD-ASC. Clustering results were partially figured out in Fig. 4-6.

In Fig. 4, images were clustered into three main clusters identified by apples for eating, Apple logo and Apple products. In Fig. 5, elephants, landscape and eagles are the typical Africa elements. Ancient buildings, Nanjing maps and Mausoleum are representatives of Nanjing city in Fig. 6.

Experiment on image segmentation: Image segmentation (Batra *et al.*,2010; Chai *et al.*,2011) is the process of dividing an image into multiple parts, which is a difficult problem in computer vision. Image segmentation(Jadin and Taib, 2012) is typically used to identify objects or other relevant information in digital images. ASC can also be used on image segmentation. Figure 7 shows the result

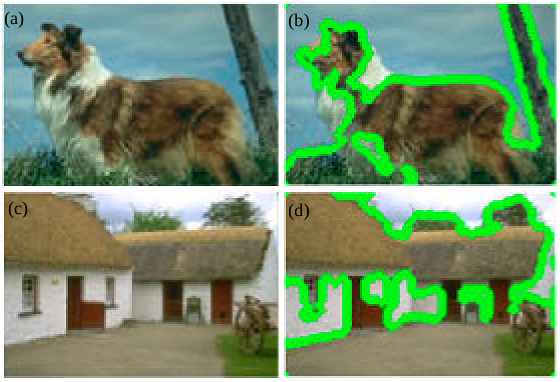

Fig. 8: Color Image segmentaion result by ASC (k = 5)

using ASC (k = 5) on two typical gray images. Notice that for both images, our method can clearly segmented the object from the background. Figure 8 shows the result using ASC (k = 5) on two typical color images.

CONCLUSION

Unlike standard spectral clustering algorithm that has a global unique scale, ASC algorithm makes each point with a different neighbour adaptive scale computed according to the mean distance to its k nearest neighbors. Neighbour adaptive scale simplifies the selection of parameters and makes the improved algorithm insensitive to both density and outliers. Experimental results show that, compared with k-means and standard spectral clustering, our algorithm can achieve better clustering effect on artificial datasets and UCI public databases. However, ASC is sensible to noises and the number of clusters is difficult to select. Besides, computational complexity is increased. For NAS, some quantitative comparison has been performed. But for image clustering and segmentation, only qualitative results were concluded. It is expected that using evaluation to improve the performance further.

ACKNOWLEDGMENT

This research was funded by a grant (No.71271117/G0112) from the National Natural Science Foundation of China and the Priority Academic Program Development of Jiangsu Higher Education Institutions(Auditing Science and Technology).

REFERENCES

Batra, D., A. Kowdle, D. Parikh, J. Luo and T. Chen, 2010. iCoseg: Interactive co-segmentation with intelligent scribble guidance. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June13-18, 2010, San Francisco, CA., pp: 3169-3176.

Blake, C., E. Keogh and C. Merz, 1998. UCI Repository of Machine Learning Database. University of California, Irvine, CA.

Chai, Y., V. Lempitsky and A. Zisserman, 2011. Bicos: A bi-level co-segmentation method for image classification. Proceedings of the IEEE International Conference on Computer Vision, November 6-13, 2012, Barcelona, pp: 2579-2586.

Chung, F.R.K., 1997. Spectral graph theory. Vol. 92, AMS Bookstore, USA., ISBN: 9780821889367, Pages: 207.

Ester, M., H.P. Kriegel, J. Sander and X. Xu, 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, August 2-4, 1996, Portland, pp: 226-231.

Feil, B. and J. Abonyi, 2007. Geodesic Distance Based Fuzzy Clustering. In: Soft Computing in Industrial Applications: Recent and Emerging Methods and Techniques, Saad, A., K. Dahal, M. Sarfraz and R. Roy (Eds.). Vol. 39, Springer, Berlin, Heidelberg, ISBN: 978-3-540-70704-2, pp: 50-59.

Goldberger, J., S. Gordon and H. Greenspan, 2006. Unsupervised image-set clustering using an information theoretic framework. IEEE Trans. Image Process., 15: 449-458.

Jadin, M.S. and S. Taib, 2012. Infrared image enhancement and segmentation for extracting the thermal anomalies in electrical equipment. Electr. Electr. Eng., 120: 107-112.

Kim, E., H. Li and X. Huang, 2012. A hierarchical image clustering cosegmentation framework. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI., pp: 686-693.

Lihi, Z.M. and P. Pietro, 1995. Self-Tuning Spectral Clustering. In: Advances in Neural Information Processing Systems 7, Tesauro, G., D.S. Touretzky and T. Leen (Eds.). MIT Press, Cambridge, MA., pp: 1601.

Meila, M. and J. Shi, 2001. A random walks view of spectral segmentation. Proceedings of the 8th International Conference on Artificial Intelligence and Statistics, (AIS'01), Key West, FL., USA.

Ng, A., M. Jordan and Y. Weiss, 2002. On Spectral Clustering: Analysis and an Algorithm. In: Advances in Neural Information Processing Systems 14, Dietterich, T.G., S. Becker and Z. Ghahramani (Eds.). MIT Press, Cambridge, MA.

Pughineanu, C. and I. Balan, 2011. Parallel algorithm evaluation in the image and clustering processing. Electr. Electr. Eng., 110: 89-92.

Rand, W., 1971. Objective criteria for the evaluation of clustering methods. J. Am. Stat. Assoc., 66: 846-850.

Shi, J. and J. Malik, 2000. Normalized cuts and image segmentation. IEEE. Trans. Pattern Anal. Mach. Intell., 22: 888-905.

Verma, D. and M. Meila, 2003. A comparison of spectral clustering algorithms. University of Washington, Tech Report No. UW-CSE-03-05-01.

Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput., 17: 395-416.

Von Luxburg, U., O. Bousquet and M. Belkin, 2005. Limits of Spectral Clustering. In: Advances in Neural Information Processing Systems 17, Saul, L.K., Y. Weiss and L. Bottou (Eds.). MIT Press, Cambridge, MA., pp: 857-864.

Wang, L., Y. Zhang and J. Feng, 2005. On the euclidean distance of images. IEEE Trans. Pattern Anal. Mach. Intell., 27: 1334-1339.