Semantic Memory Learning in Image Retrieval Using k Means Approach

Pushpa B. Patil[1(✉)] and M. B. Kokare[2]

[1] BLDEA's V. P. Dr. P.G. Halakatti College of Engineering and Technology, Vijayapur 586103, Karnataka, India
cs.pushpa@bldeacet.ac.in

[2] SGGS IOT, Nanded, Maharashtra, India

Abstract. To reduce the conceptual gap in content-based image retrieval (CBIR) and small training problem in relevance feedback (RF), this paper attempts to focus on the semantic memory learning in image retrieval using proposed 2-means clustering. In this system, initial retrieval results of CBIR are obtained, and then user's opinion is given to the system as relevant/irrelevant to the user. With this user feedback, we can easily make the relevant image cluster and the irrelevant image cluster directly instead of random selection. Hence with initial known clusters and number of clusters, computational time is highly reduced for finding cluster center. We have also reduced the burden of clustering by considering only relevant cluster repeatedly for each feedback iteration. We experimented on two different data sets using proposed system. Results are found better compared to the earlier approaches.

Keywords: Image retrieval · k-means · Relevance feedback · Complex wavelets

1 Introduction

In this internet era, there is rapid enhancement and changes in digital technology. Hence there is a big collection of different variety of digitized images with respect to different applications like medical, entertainment, and biometric etc. So it is like galaxy with millions or billions of stars. With such huge information, user needs to search, browse, and retrieve relevant information. For fulfillment of the user, there is a need for efficient and effective retrieval systems. So in past, researchers introduced two kinds of retrieval systems namely based text and content of image. Initially, there was only image retrieval based text. The drawback of this system is manual labeling of huge image collection. It leads increase in labor cost and difficult to maintain user perception. To address these problems, researcher introduced Content Based Image Retrieval (CBIR) in the year 1990. It retrieves the images based on low level features like texture, color and shape etc. Therefore it is called as content based image retrieval. Earlier, a few marketable products and experimental models were developed, such as

K. C. Santosh and R. S. Hegadi (Eds.): RTIP2R 2018, CCIS 1035, pp. 520–533, 2019.
https://doi.org/10.1007/978-981-13-9181-1_46

Virage, QBIC, SIMPLIcity, VisualSEEK, Netra and Photobook. Detailed surveys on CBIR presented in [2,4,17,18] In addition to these approaches, recently ontology based annotation tool used for image retrieval [13].

1.1 Related Work

Relevance feedback is a semantic classification approach. Here, user feeds both relevant and/or irrelevant data, and then it learns from that input to divide all data into appropriate and in appropriate groups with respect to query image. Hence many supervised machine learning algorithms are useful to design RF, namely Bayesian learning [19], decision tree learning [9], support vector machines and Gaussian mixture models [10], boosting, Re-weighting and, Query refinement [1] so on. The learning procedure is very hard job in RF, because of three causes, firstly training data set size is small, secondly imbalance in training data set images, finally RF takes more real time since both testing and training process has to be performed online for every feedback iteration.

Liu and Yu [8] used k-means to cluster images in the image database then similarity is applied to the clustered database instead of feature database in order to reduce retrieval time. With known number of image categories it works better. Murthy et al. [11] used hierarchical and the k-means clustering algorithms to group the images into clusters based on the color content. Initially images in the database are filtered using Hierarchical clustering and then applied the clustered images to k-means for better retrieval performance. Mishra et al. [12] used k-means to classify the coherent pixels and incoherent pixels for color images. With these observation of the above literature review in context to k-means clustering in CBIR, we found that k-means is applied for low level content in earlier approaches rather than semantic learning. Santosh et al. [21,22] proposed dynamic time warping for matching radon features. In 2018, Engin, and Cavusoglu proposed rotation invariant features using curvelet transform for retrieval of images [23].

The major goal of this paper is briefed out here, we have presented new semantic learning in image retrieval using 2-means clustering algorithm. With this we try to solve the small training data problem (number of training data is less than feature vector dimension) and real time problem (since RF is an online process). Relevance feedback works on two known image group namely relevant image group and irrelevant image group, this motivates us to use k-means algorithm to find the clusters centers directly instead of random computation. Hence we achieved better retrieval performance with less computational time. The practical results of anticipated method perform superior, compared to earlier method.

The remaining part of the paper is structured as follows. We discuss the image descriptors in brief in Sect. 2. The general k-means approach and proposed semantic memory learning in image retrieval using modified k-means algorithm called 2-means clustering are discussed in Sect. 3. Results are discussed in Sect. 4 and last section concludes.

2 Image Descriptors

We extracted the image features using combined "rotated complex wavelet filters (DT-RCWF)" and "complex wavelets (DT-CWT)". As a result, we will get information in twelve different directions. However results of depend on visual features [14] and similarity metrics [5]. Both the wavelets are discussed in the following section.

2.1 Complex Wavelet Transforms

DWT has drawbacks, namely it gives only four directional information with lack of shift invariance. To address these problems of DWT, We used dual tree complex wavelet transform (DT-CWT) [7]. It provides six directional information namely $(-15°, -45°, -75°, +15°, +45°\ and\ +75°)$ shown in Fig. 1.

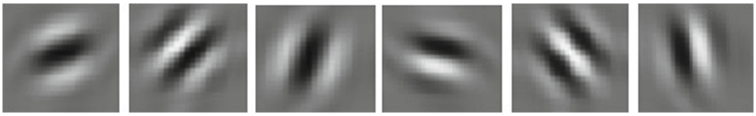

Fig. 1. Six orientations of the Wavelet Filters $(15°, 45°, 75°, -15°, -45°\ and\ -75°)$ of Complex Wavelet.

2.2 Rotated Complex Wavelet Transforms

In 2005, authors [6] created 2D rotated complex wavelet filters(RCWF) which gives six different direction information's, which is 45° away from each other from decomposition of DT CWT. Hence we have another six different directional information oriented at $(30°, 0°, +30°, +60°, 90°\ and\ 120°)$. The six orientations associated with this are illustrated in Fig. 2. For similarity measure we used the Canberra distance measure. With these image features we developed the new relevance feedback framework using k means clustering in the following section.

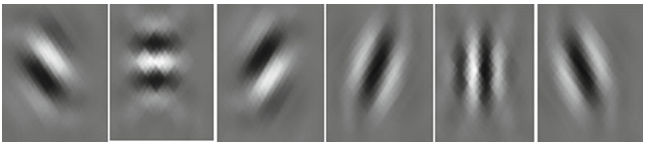

Fig. 2. Six orientations of Rotated Complex Wavelets $(-30°, 0°, +30°, +60°, 90°\ and\ 120°)$

3 Proposed Semantic Memory Learning Framework

The fundamental pace of k means is used to decide number of cluster k and then assume the centroid or center of these clusters. Since k is unknown for classification of objects. Therefore it considers any arbitrary objects as the initial centroids. An algorithm 1 will depicts the sequence of steps until convergence in general k-means clustering

Algorithm 1. *General k -means algorithm*

```
Inputs:o={o1,o2,...ok}   {objects to be clustered}
       k:number of clusters
Outputs: C={c1,c2,...ck} (cluster centroids)
         m:O->{1...k}(cluster membership)
    Begin
        Set  C to initial value (e.g. random selection of O )
        For each oi belongs to O do
             m(oi)=     argmin         dist(oi,cj)
                   j belongs to{1...k}
         End
        While m  has changed do
              For each  i belongs to {1...k}  do
                   Recompute the ci   as the centroid of {o|m(o)=i}
               End
             For each oi belongs O do
                m(oi)=     argmin         dist(oi,cj)
                   j belongs to{1...k}
              End
         End
    End
```

Here we presented novel RF approach using k means. The k-means plays an important role in relevance feedback in CBIR. However cost of computation of the initial cluster center is more using random objects as we have seen in the Algorithm 1. In this section we introduced the k-means algorithm for image retrieval, which takes least amount of computational time for generation of the cluster centers.

Figure 3 depicts the basic modules of the proposed semantic memory learning framework for interactive image retrieval. In order to reduce the conceptual gap between low level content and high level perception, the traditional CBIR system is enhanced by introducing relevance feedback (RF) loop in it. The relevance feedback is an online process, it takes the feedback (relevant/irrelevant) from the user and refines results using supervised/unsupervised [9,10,19,20] or query point movement or re-weighting approach [1]. It is continued till the user fulfillment or the output does not improve further.

In the proposed system, we used k-means clustering to retrieve the user perception information. However the k-means is unsupervised learning and cost of computation of the cluster center is more. If number of cluster centers k and dimension of vector d is constant, then the cluster center can be computed in

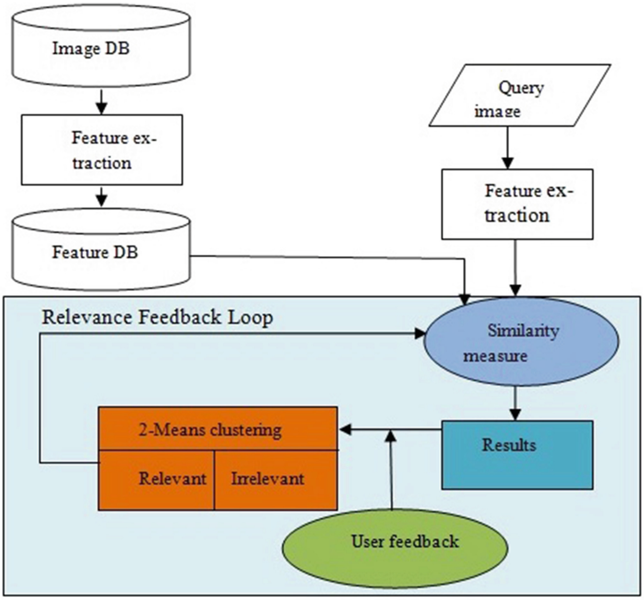

Fig. 3. Proposed system architecture

time $O(n^{dlogn})$, where n is the number of images to be clustered. Hence in this system k and d values are fixed and small values. That is $k = 2$ and $d = length$ of the feature vector. It motivates us to use k-means clustering for semantic learning with known number of cluster and hence it is easy to compute the cluster center in constant time(best case efficiency). In Sect. 3.1, the proposed semantic memory learning in image retrieval system is discussed in detail.

3.1 Proposed 2 Means Approach

The performance of the k means limits due to random selection of initial centroid by the user. It motivates to propose clustering algorithm which computes centroids appropriately with known relevant and irrelevant groups; As a results this, we will get the real and proper creation of the clusters. The proposed system uses the results obtained from CBIR as the initial training set. Then training set is annotated by user either relevant or irrelevant. Thus the training data set consists of N input vectors $(X_1, X_2, X_3, \ldots, X_N)$ with corresponding labels $(t_1, t_2, t_3, \ldots, t_N)$ and new data's are classified using k-means clustering. From user feedback, we have relevant data set $fr = \{fr_1, fr_2, \ldots, fr_p\} \subset X$ and irrelevant $fn = \{fn_1, fn_2, \ldots, fn_q\} \subset X$such that $f_n \cap f_r = \phi$ (i.e null set). Where p and q are the number of related and unrelated images respectively. Hence the number of clusters $k = 2$. We determine the two cluster center for relevant image group $\overline{fr}$ and irrelevant image group $\overline{fn}$ using Eqs. (1) and (2) respectively.

$$\overline{fr} = \frac{1}{p}\sum_{i=1}^{p} fr_i \tag{1}$$

$$\overline{fn} = \frac{1}{q}\sum_{i=1}^{q} fn_i \tag{2}$$

We determine the similarity distance with the database images and relevant image group centriod $\overline{fr}$ and irrelevant image group centriod $\overline{fn}$ using Eqs. (3) and (4) respectively.

$$dr = \sum_{i=1}^{d} \frac{|x_i - \overline{fr}|}{|x_i| - |\overline{fr}|} \tag{3}$$

$$dn = \sum_{i=1}^{d} \frac{|x_i - \overline{fn}|}{|x_i| - |\overline{fn}|} \tag{4}$$

Hence it makes us to categories the relevant image group and irrelevant image group based on minimum distance. To speed up the testing time, here we concentrated on relevant image group and neglected the irrelevant group in every iteration of the feedback. We used the memory learning concepts to produce the results in feedback iteration.

4 Experimental Results

We conducted experiments with known number of category in the database and number of images in each category, we have designed RF framework to obtain the user feedback automatically. In this design, images belongs to the category of the query image are considered as relevant. In RF, we can carry out the number of rounds repetitively till there is no improvement in results/user satisfaction. Since, the numbers of rounds are directly proportional to the retrieval performance. A system tested for evaluation of retrieval performance by taking into account of top 20 images for each iteration. For performance evaluation the approach, we employed both Brodatz texture [6] and Corel color image dataset [3].

4.1 Image Data Set

We have used two standard image databases namely Brodatz texture data set and Corel Image Data Set. Brodatz texture data set comprises 116 variety textures. Size of image is 128 × 128. Database includes 1856 such images. Figure 4 shows the sample image of the each category from Brodatz texture database. Corel image set comprises 1000 color images of size 384 × 256 pixels, includes a various natural to artificial scenes [3]. The data set is divided into ten categories, each with 100 images. Ten categories are namely Dinosaurs, African people, Flowers, Beach, Building, Buses, Elephants, Horses, and Food. Figure 5 shows the example image of the each category from Corel natural color image database.

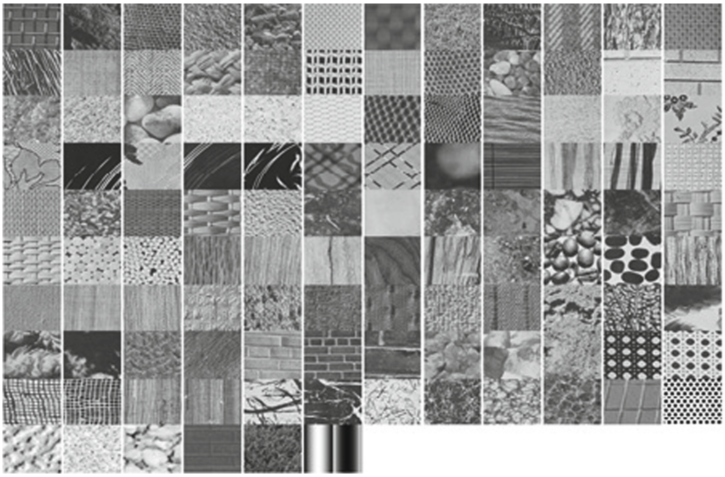

Fig. 4. Example Image of each category: Brodatz texture dataset-116 categories

Fig. 5. Example Image of each category: COREL image dataset-10 categories

4.2 Performance Parameters

For retrieval performance analysis, it is important to define a appropriate metric for performance evaluation. Therefore following performance measures are used.

$$recall = \frac{No.\quad of \quad relevant \quad images \quad retrieved}{Total \quad No. \quad of \quad relevant \quad images \quad in \quad the \quad Database} \tag{5}$$

$$precision = \frac{No. \quad of \quad relevant \quad images \quad retrieved}{Total \quad No. \quad of \quad relevant \quad images \quad retrieved} \tag{6}$$

A system performance is tested for 116 images from the texture database. The average accuracy is computed for all tested images. In each experiment, from each category randomly one image was selected as a query image. Thus, we have retrieved images. Then, the users has to identify images which are relevant from the retrieved images. This user selection image set is fed to the RF system for next round. Finally, average accuracy of all the categories in the database is computed. The number of iterations were performed up to 8 times for texture

Table 1. Average accuracy on each feedback iterations for texture data set

Approach	CBIR	1st iter	2nd iter	3rd iter	4th iter	5th iter
k-means (proposed)	78.50	89.83	91.41	92.27	92.60	92.82
SVMRF	78.50	89.27	91.75	92.18	92.29	92.29
ADABoostRF	78.50	88.52	91.32	91.70	91.70	91.70

database and 13 times for Corel database. Since the feedback process is repeated until result doesn't improve further.

Figure 6 depicts comparative retrieval results obtained using, AdaBoostRF, SVMRF and proposed k means RF on every feedback. The proposed k-means RF is compared with SVMRF [15] and AdaBoostRF [16]. From Fig. 6 we can observe that, the proposed k-means RF framework is better over AdaBoostRF and SVMRF However there is a quick boost in retrieval performance with each feedback of RF using both methods. Finally, the accuracy of k-mreansRF is also higher than that achieved by the SVMRF and AdaBoostRF, starting from the first iteration. Results are depicted in Table 1. Note that retrieval results of AdaBoostRF and SVMRF remain same after 3th and 4th iteration respectively, however results of the kmeansRF increases the retrieval performance from the previous iteration to the next iteration. It can be also observed that the performances achieved by Bordatz dataset are usually higher than those of the Corel data set. The reason of this performance is related to the different semantic of the images contained in the two datasets, and to their subdivision into categories. Similarly we computed results for Corel image data set, which consists of 10 categories of images and in each category 100 natural colour images. For testing we have selected randomly 5 images from each category as query images

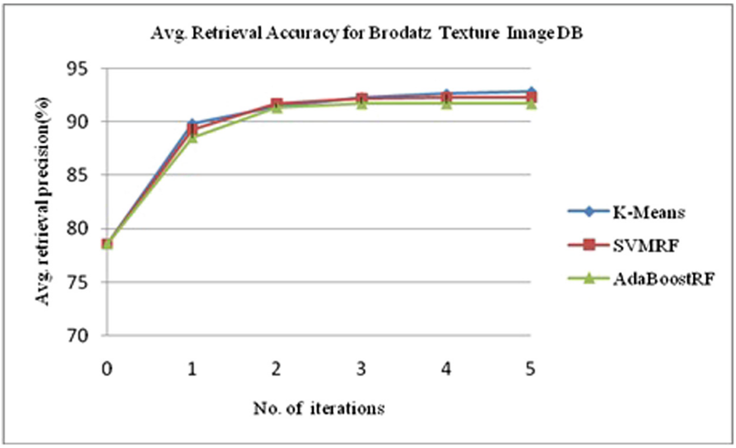

Fig. 6. Number of iteration versus average accuracy for texture images

(altogether 50 images). The reported results of average precision are obtained by taking an average over the 50 queries. Figure 7 depicts the complete assessment of average retrieval precision got from using SVMRF [15], AdaBoostRF [16] and proposed k-means RF on every feedback iteration for Corel Images. We can observe from the Fig. 7, the proposed approach produced superior retrieval performance than the Single_RBF and RBFGaussFunction anticipated by Ding et al. [3]. Results are listed in Table 2.

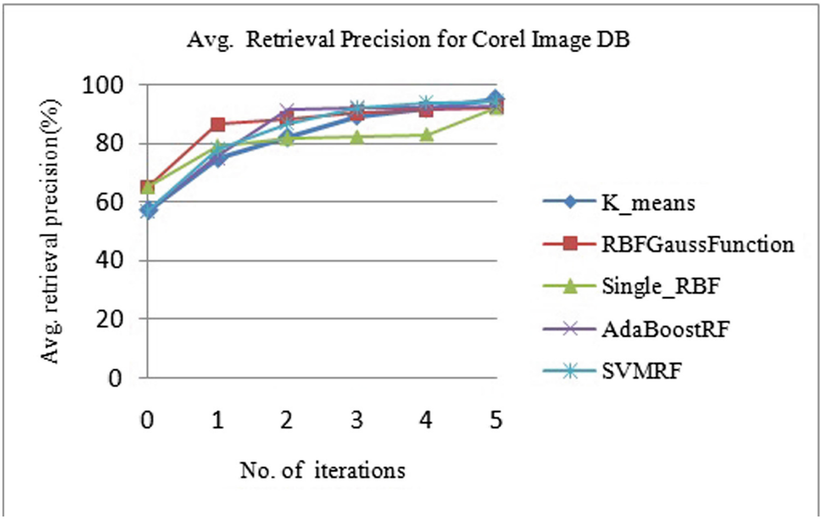

Fig. 7. Average precision versus iteration curves for Corel images

As stated before we have conducted experiments on five set of images, in each set we selected randomly an image from each category. In total there are ten categories in the database. Thus we tested 10 images from each image set. Hence, total number of testing images is fifty (5×10). For more clarity, Figs. 8(a)–(d) and 9 shows the precision versus the iterations curves of five testing image sets separately. For more clarity observe catg_2 (buildings image) graph in Fig. 8(b), where CBIR retrieval precision of catg_2 image is 20% and then it increases to 70% in the first iteration, 85% in the second iteration, 90% in the third iteration, 100% in the fourth iteration, and finally from the fifth iteration on words result remains same. For catg_5 (i.e Dinosaur image) in all five image sets (see Figs. 8(a)–(d) and 9) retrieval precision is 100% without relevance feedback. Furthermore, from all image sets in the Figs. 8(a)–(d) and 9, more than five categories images reached 100% precision at the fifth iteration of system.

Table 2. Average precision of each feedback iteration for Corel image database

Approach	CBIR	1st iter	2nd iter	3rd iter	4th iter	5th iter
RBFGaussFunction	65.2	86.5	88.4	90.4	91.5	92.3
SingleRBF	65.2	79.2	81.9	82.3	83.1	84.6
AdaBoostRF	57.2	75.4	91.32	91.70	91.70	92.5
SVMRF	57.2	78	86.9	92.2	94.0	94.6
k means (proposed)	57.2	74.9	81.96	89.35	91.89	95.07

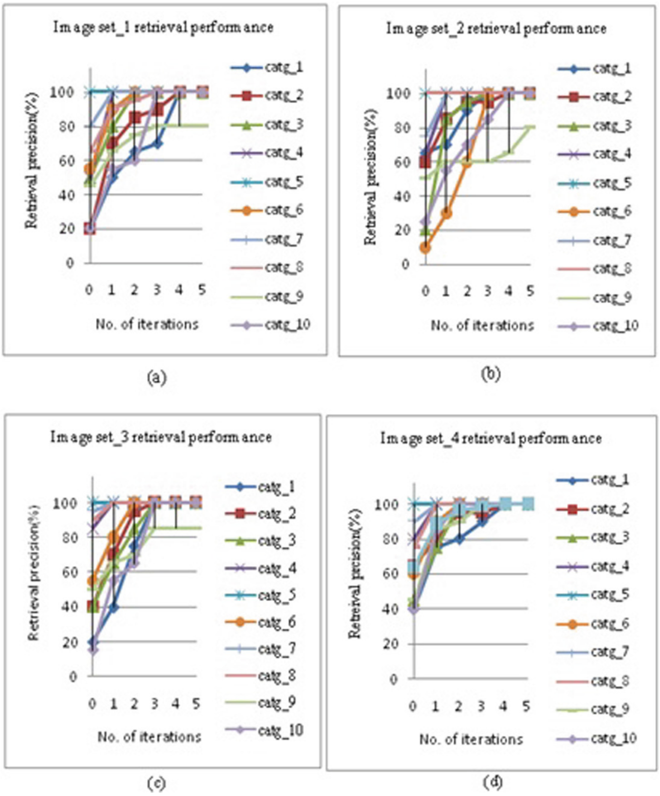

Fig. 8. Precision of all 10 categories images of Corel DB versus iteration curves five image sets: (a) Image set 1, (b) Image set 2, (c) Image set 3, (d) Image set 4

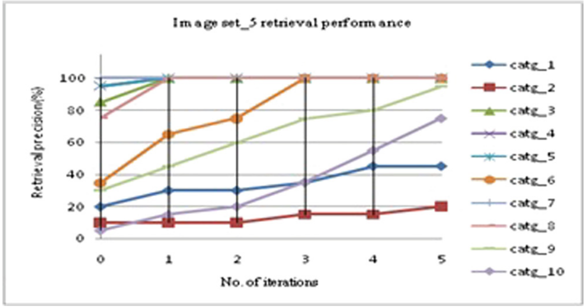

Fig. 9. Precision of all 10 categories images of Corel DB versus iteration curves five image sets: Image set 5

4.3 Image Retrieval Examples

Figures 10, 11 and 12 shows improvement of retrieval performance from the initial CBIR results to second feedback iteration. We can observe from Fig. 10, retrieval precision of initial CBIR is 55% and then retrieval precision is increased from 55% to 85% in the first round of the relevance feedback (see Fig. 11). Finally, we can observe from Fig. 12, improvement of retrieval precision reaches to the 100% in second iteration.

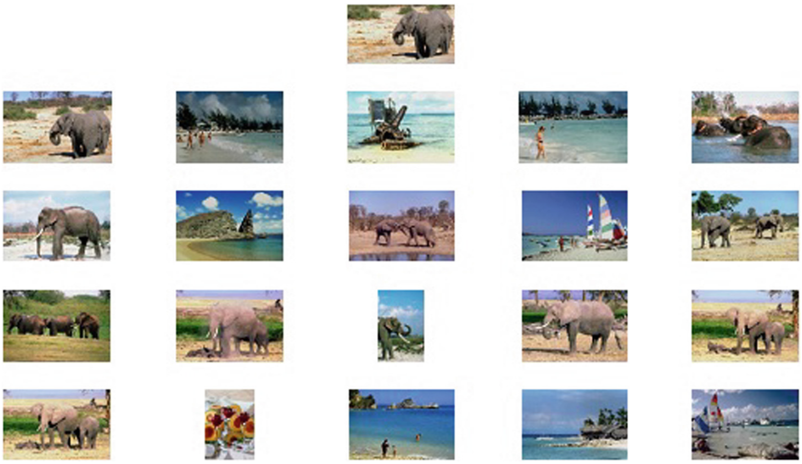

Fig. 10. Initial retrieval results of CBIR (11/20)

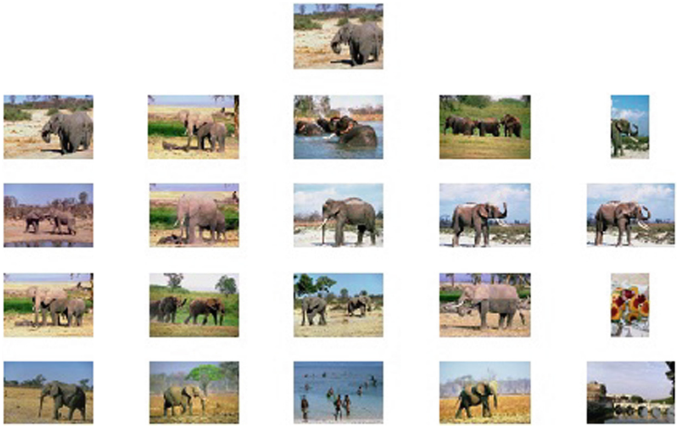

Fig. 11. Results after first feedback iteration (17/20)

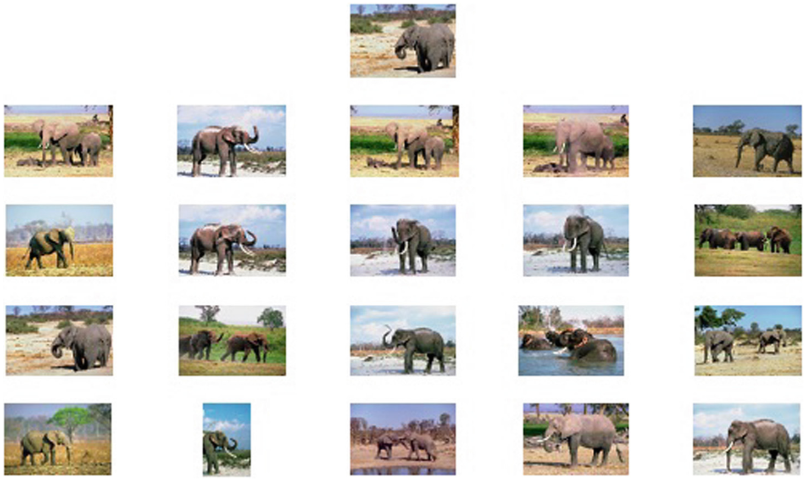

Fig. 12. Results after second feedback iteration (20/20)

5 Conclusion

We have developed new semantic memory learning in image retrieval using k-means clustering. It is an interactive online RF process, hence the utilization of k-means become effective and efficient with known number clusters fed from the user in every round of feedback. In order to boost the retrieval time we considered only positive image group. It uses Canberra distance to classify the relevant and irrelevant image group. It is experimented on both e texture image database and natural image database. Proposed system gives very promising retrieval accuracy and precision. Proposed system can be extended to develop

RF for different unsupervised clustering algorithm Fuzzy C means, k-memoids, and to support the linear composition of the clustering as future work.

References

1. Chiang, C.C., Wu, J.Y., Yang, M.T.: Independent query refinement and feature re-weighting using positive and negative examples for content-based image re-trieval. Multimedia Tools Appl. **41**, 27–43 (2009)
2. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: ideas, in-fluences, and trends of the new age. ACM Comput. Surv. **40**(2), 5:1–5:60 (2008). Article no. 5
3. Ding, R., Ji, X., Zhu, L.: Research on the relevance feedback based image retrieval in digital library. PWASET **25**, 48–53 (2007). ISSN 1307–6884
4. Kokare, M.B., Chatterji, B.N., Biswas, P.K.: A survey on current content-based image retrieval methods. IETE J. Res. **48**(3 and 4), 261–271 (2002)
5. Kokare, M.B., Chatterji, B.N., Biswas, P.K.: Comparison of similarity metrics for texture image retrieval. In: IEEE TENCON, pp. 571–575 (2003)
6. Kokare, M.B., Chatterji, B.N., Biswas, P.K.: Texture image retrieval using new rotated complex wavelet filters. IEEE Trans. Syst. Man Cybern. Part B: Cybern. **35**(6), 1168–1178 (2005)
7. Kingsbury, N.G.: Image processing with complex wavelet. Phil. Trans. Roy. Soc. London A **357**(1999), 2543–2560 (1999)
8. Liu, H., Yu, X.: Application research of k-means clustering algorithm in image retrieval system. In: Proceedings of the Second Symposium International Computer Science and Computational Technology (ISCSCT 2009), China, 26–28, pp. 274–277 (2009)
9. MacArthur, S.D., Brodley C.E., Shyu, C.R.: Relevance feedback decision trees in content-based image retrieval. In: Proceedings of the IEEE Work-Shop Content-Based Access of Image and Video Libraries, pp. 68–72 (2000)
10. Marakakis, A., Galatsanos, N., Likas, A., Stafylopatis, A.: Combining Gaussian mixture models and support vector machines for relevance feedback in content based image retrieval. In: Iliadis, Maglogiann, Tsoumakasis, Bramer, (eds.) IFIP International Federation for Information Processing, vol. 296, pp. 249–258. Springer, Boston (2009). https://doi.org/10.1007/978-1-4419-0221-4_30
11. Murthy, V.S., Vamsidhar, E., Swarup Kumar, J.N.V.R., Rao, P.S.: Content based image retrieval using hierarchical and k-means clustering techniques. Int. J. Eng. Sci. Technol. **2**(3), 209–212 (2010)
12. Mishra, J., Sharma, A., Chaturvedi, K.: An unsupervised cluster-based image retrieval algorithm using relevance feedback. Int. J. Manag. Inf. Technol. (IJMIT) **3**(2), 10–16 (2011)
13. Horvat, M., Grbin, A., Gledec, G.: Labeling and retrieval of emotionally annotated images using WordNet. Int. J. Knowl. Based Intell. Eng. **17**(2), 157–166 (2013)
14. Patil, P.B., Kokare, M.B.: Interactive content -based texture image retrieval. In: IEEE ICCCT2011, MNNIT, Allahabad, 15–17 September, pp. 71–76 (2011)
15. Patil, P.B., Kokare, M.B.: Relevance feedback in content-based image retrieval. In: Proceeding of 4th International Conference Computer Application in Electrical Engineering-Recent Advances, pp. 364–367. IIT Roorke (2010)
16. Patil, P.B., Kokare, M.: Semantic learning in interactive image retrieval. In: Nagamalai, D., Renault, E., Dhanuskodi, M. (eds.) DPPR 2011. CCIS, vol. 205, pp. 118–127. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-24055-3_12

17. Rui, Y., Huang, T.S., Chang, S.F.: Image retrieval: current techniques, promising di-rections and open issues. J. Vis. Commun. Image Represent. **10**, 39–62 (2011)
18. Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content - based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1349–1380 (2000)
19. Su, Z., Zhang, H., Li, S., Ma, S.: Relevance feedback in content-based image retrieval: Bayesian framework, feature subspaces, and progressive learning. IEEE Trans. Image Process. **12**(8), 924–936 (2003)
20. Tieu, K., Viola, P.: Boosting image retrieval. In: Proceedings of the IEEE Conference Computer Vision Pattern Recognition, pp. 228–235 (2003)
21. Zhou, X.S., Huang, T.S.: Relevance feedback in image retrieval: a comprehensive review. Multimedia Syst. **8**(6), 536–544 (2003)
22. Santosh, K.C., Lamiroy, B., Wendling, L.: DTW-radon-based shape descriptor for pattern recognition. Int. J. Pattern Recogn. Artif. Intell. (IJPRAI) **27**(3), 30 (2013)
23. Santosh, K.C., Lamiroy, B., Wendling, L.: DTW for matching radon features: a pattern recognition and retrieval method. In: Blanc-Talon, J., Kleihorst, R., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2011. LNCS, vol. 6915, pp. 249–260. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23687-7_23