ijecs open access www.ijecs.in

International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 12 December 2014, Page No. 9647-9651

An Efficient Content Based Image Retrieval System for Sketches by Using PAM Algorithm

A. Ravi Kumar, A. Yuva Krishna
M.Tech Student
, Assistant Professor
CSE Department, PVP Siddhartha Institute Of Technology, Kanuru, Vijayawada

Abstract

In this paper we propose a new and effective image retrieval scheme using color, texture and shape features based on K-medoids. Image is predefined by using fast color quantization algorithm with cluster merging. A small number of dominant colors and their percentage can be obtained. The spatial features are extracted using clustering methods. This offers an efficient and flexible approximation of early processing in the human visual system (HVS). It provides better feature representation and more robust to noise then other representations. Finally, the combination of the color, shape and texture feature provides a robust feature set for image retrieval. Experimental result shows that the proposed method provides a better color image retrieval. It provides accurate and effective in retrieving the user-interested images.

Introduction

A cluster could be a assortment knowledge of knowledge of information points that square measure the same as each {other} among identical cluster and dissimilar to data points in other clusters [1]. bunch could be a methodology of unattended classification, wherever information points square measure sorted into clusters supported their similarity. The goal of a bunch algorithmic program is to maximize the intra-cluster similarity and minimize the inter-cluster similarity.

Clustering algorithms may be broadly speaking classified into 5 types:

1. Partitioned bunch, 2. graded bunch, 3. Density-based bunch, 4. Grid-based bunch and five. Model-based bunch [1]. Partitioned and graded bunch square measure the foremost wide used types of bunch. In partition bunch, the set of n information points square measure divided into k non-empty clusters, wherever $k \leq n$. within the case of graded bunch, the info points square measure organized into a hierarchical data structure, leading to a binary tree or dendogram [2].

In this paper, we have a tendency to propose a brand new bunch algorithmic program, which might come back below the class of partitioned bunch algorithms. 2 unremarkably used ways for partitioning information points embody the k-means methodology [3, 4] and therefore the k-medoids methodology [5]. With in the k-means methodology, every cluster is described by its center of mass or the mean of all information points within the cluster. With in the case of the k-medoids methodology, every cluster is described by an information purpose near the center of mass of the cluster. Except for these ways, there has been voluminous work on fuzzy partitioning ways

[6] and partition ways for large-scale datasets [7].We use the notion of 'contribution of an information point' for partitioned bunch. The resultant algorithmic program needs solely 3 passes and that we show that the time quality of every pass is same as that of one iteration of the k-means bunch algorithmic program. Whereas the k-means algorithmic program optimizes solely on the intra-cluster similarity, our algorithmic program additionally optimizes on the inter-cluster similarity.

Content-Based Image Retrieval

Content-based image retrieval could be a technique that uses visual content to look pictures from large-scale databases per users' interest [8]. a standard technique of querying a content-based image retrieval system is to produce AN example image. The system then retrieves all pictures within the information that ar similar in content with the question image.

In this paper, we have a tendency to target the appliance of cluster to content-based image retrieval. an outsized assortment of pictures is partitioned off into variety of image clusters. Given a question image, the system retrieves all pictures from the cluster that's nearest in content to the question image. the general system is shown in Fig. 1. we have a tendency to apply the planned contribution-based cluster formula to image retrieval and compare its performance therewith of the k-means formula.

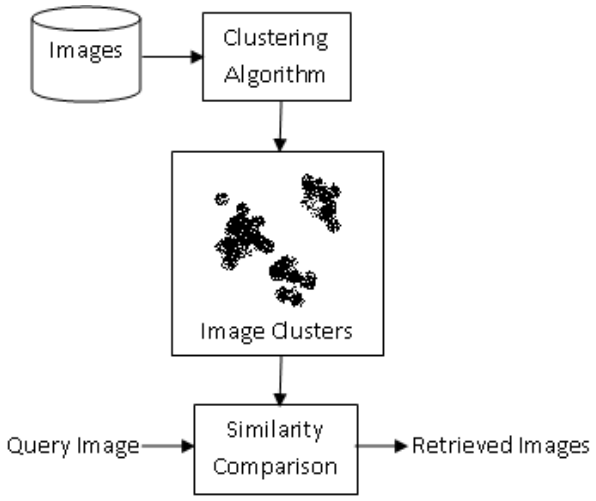

Fig. 1. Content-based Image Retrieval (CBIR) System.

Each image within the information is portrayed by a visible content descriptor consisting of a collection of visual options [8]. A similarity/dissimilarity live is then wont to retrieve pictures whose options ar nearest to it of the question image. a standard distance/dissimilarity metric is that the euclidian distance, that is employed in our work. To represent the visual content of a picture, we have a tendency to use a RGB color bar chart. the colour coordinates of the RGB color area ar uniformly quantal into variety of bins. In our work, we have a tendency to use eight bins every for the Red, Green and Blue coordinates, resulting in 512 bins/features.

K-Medoids

The k-medoids algorithmic program could be a cluster algorithmic program associated with the k-means algorithmic program and therefore the medoidshift algorithmic program. each the k-means and k-medoids algorithms area unit partitional (breaking the dataset up into groups) and each decide to minimize the space between points labeled to be during a cluster and some extent selected because the center of that cluster. In distinction to the k-means algorithmic program, k-medoids chooses datapoints as centers (medoids or exemplars) associate degreed works with an discretionary matrix of distances between datapoints rather than l_2. This technique was planned in 1987[1] for the work with l_1 norm and alternative distances.k-medoid could be a classical partitioning technique of cluster that clusters the information set of n objects into k clusters famed a priori. A useful gizmo for deciding k is that the silhouette.It is a lot of sturdy

to noise and outliers as compared to k-means as a result of it minimizes a add of pair wise dissimilarities rather than a add of square geometrician distances. A k-medoid is outlined because the object of a cluster, whose average un similarity to any or all the objects within the cluster is token i.e. it's a most centrally situated purpose within the cluster.

The most common realisation of k-medoid cluster is that the Partitioning Around Medoids (PAM) algorithmic program and is as follows

1.Initialize: at random choose (without replacement) k of the n information points because the medoids

2.Associate every information to the nighest medoid. ("Closest" here is outlined mistreatment any valid distance metric, most ordinarily geometrician distance, Manhattan distance or Hermann Murkowski distance)

3. For each k-medoids m

A. For each non-medoids information o

B. Swap m and o and reckon the whole value of the configuration

4.Select the configuration with very cheap value.

5.Repeat steps a pair of to four till there's no amendment within the k-medoid.

Proposed method

Since image retrieval is in step with the similarities between the question image and pictures in image info, ignoring the similarities between pictures in image info. The paper applies the bunch formula to additional explore the similarities between pictures in image info for reducing the image retrieval house.

In CBIR, Image feature vectors is pictured by a true matrix G, every row gi of that represents the feature vector of a picture in info, and every column represents one reasonably feature price. part G(i,k) represents the feature price of the I th image underneath the kth feature. The relationship among images can be pictured by affinity A ¼ (aij), wherever aij represents the similarity between the ith image and also the jth image. The similarity may be measured by Euclidean distance or different metrics. The image retrieval system hair care bunch formula is shown as Fig.2. The system may be any real price stellate image retrieval system. First, extract the image options of every image in image info and apply the bunch formula to analysis the similarities of pictures within the info for constructing the photographs bunch info, then, input the question image, extracting its options and examination the similarities between options of it and people of pictures in image bunch info, and output the simplest matching results.

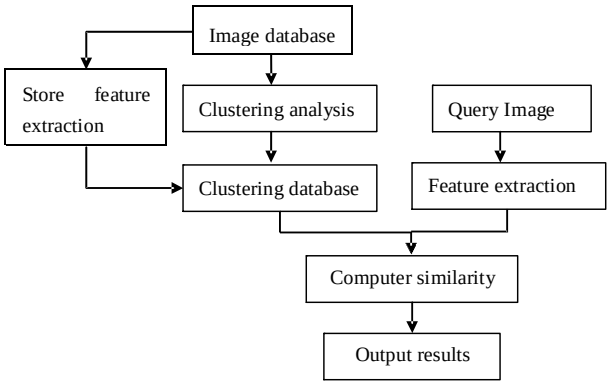

Fig2. Framework of image retrieval system combined with clustering algorithm.

SIMULATION RESULTS

We performed experiments in the main on pictures testing set, like flowers, flags and winter, exploitation color feature solely and color and form feature for retrieval, combining the k-medidos cluster rule with pictures retrieval.

In order to point out k-medidos cluster rule performance in image retrieval system, we tend to style a series of take a look at on the cluster performance. Fig. 3 and Fig.4 show the retrieval

results with and while not k-medidos cluster rule. The higher left image is that the question image. the proper half is that the retrieved pictures. Fig.3 a pair of displays the retrieval results while not k-medidos cluster rule, while Fig.4 displays the results with k-medidos cluster rule that pictures almost like one another and to question . In Fig. 3, the question image could be a denser and have vector composed of color and texture. we tend to observe that application k-medidos cluster rule within the pictures retrieval will throw away some pictures that ar visually unsuitable to the question image for reducing pictures retrieval house. A dinosaur image is displayed within the firstly retrieval leads to Fig. 3.thanks to low-level color options. Whereas through doing k-medidos cluster analysis for image retrieval, such dinosaur image isn't displayed within the leads to Fig.4.

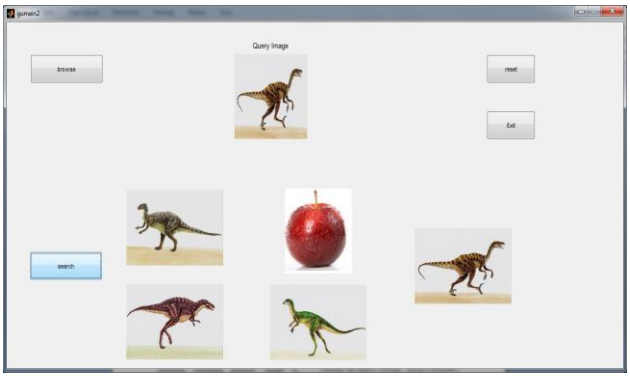

Fig.3

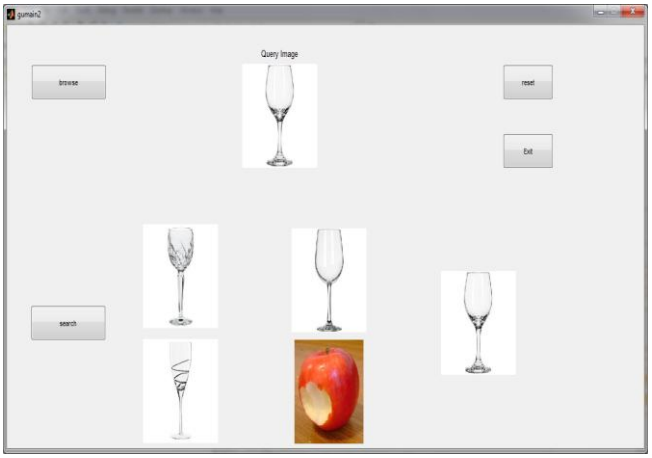

Fig.4

CONCLUSION

In conventional approaches we used to follow the k means clustering approach for image retrieval process. But due to inappropriate retrieval process we opt for advanced CBIR approach for data retrieval approach. In this paper we propose a new and effective image retrieval scheme using k medoids clustering method. Image is predefined by using fast color quantization algorithm with cluster merging. A small number of dominant colors and their percentage can be obtained. This offers an efficient and flexible approximation of early processing in the human visual system (HVS). It provides better feature representation and more robust to noise then other representations.

REFERENCES

[1] D. Comaniciu, and P. Meer, "Robust analysis of feature spaces: color image segmentation," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 750–755, June 1997.

[2] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Conference on Computer Vision and Pat1tern Recognition*, pp. 886–893, July 2005.

[3] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: an experimental comparison," *Information Retrieval*, vol. 11, pp. 77–107, December 2007.

[4] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa, "An evaluation of descriptors for large-scale image retrieval from sketched feature lines," *Computers and Graphics*, vol. 34, pp. 482–498, October 2010.

[5] R. Fabbri, L.D.F. Costa, J.C. Torelli, and O.M. Bruno, "2D Euclidean distance transform algorithms: a comparative survey," *ACM Computing Surveys*, vol. 44, pp. 1–44, February 2008.

[6] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Hiang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: the QBIC

system," *IEEE Computer*, vol. 28, pp. 23–32, 2002.
[7] Gy. Gy¨or¨ok, "Embedded hybrid controller with programmable analog
circuit," *IEEE 14th International Conference on Intelligent Systems* pp.
59.1–59.4, May 2010.
[8] R. Hu, M. Barnard, and J. Collomosse, "Gradient _eld descriptor for
sketch based image retrieval and localization," *International Conference
on Image Processing*, pp. 1–4, 2010.
[9] A.K. Jain, J.E. Lee, and R. Jin, "Sketch to photo matching: a feature-based
approach," *Proc. SPIE, Biometric Technology for Human Identi_cation VII*,
vol. 7667, pp. 766702–766702, 2010.
[10] A.K. Jain, J.E. Lee, R. Jin, and N. Gregg, "Graf_ti-ID: matching retrieval
of graf_ti images," *ACM MM, MiFor'09*, pp. 1–6, 2009.
[11] A.K. Jain, J.E. Lee, R. Jin, and N. Gregg, "Content based image retrieval:
an application to tattoo images," *IEEE International Conference on Image
Processing*, pp. 2745–2748, November 2009
[12] T. Hashimoto, A. R¨ovid, G. Ohashi, Y. Ogura, H. Nakahara, and A.R. V´arkonyi-K´oczy, "Edge detection based image retrieval method by sketches," *Proc. of the Internatonal Symposium on Flexible Automation*, pp. 1–4, 2006.