A survey on Content based Image Retrieval using Vector Quantization

Prerana Karnik
P. G. Student
KKWIEER Nashik

Nitin Shahane
Associate Professor
KKWIEER Nashik

ABSTRACT
Image retrieval and image compression are active field of research. But rare advances have been made to consider these both problems simultaneously. In this paper we present a survey on content based image retrieval based on vector quantization. Vector Quantization (VQ) is a technique usually used for data compression. But VQ is used as a feature descriptor for retrieving images. By harnessing the characteristics of VQ the retrieval technique can capture the spatial relationship among pixels. We discuss here several ways in which VQ has been used for content based image retrieval in the past.

General Terms
Image indexing, Image retrieval, Image compression.

Keywords
Content Based Image Retrieval (CBIR), Vector Quantization (VQ), Linde-Buzo-Gray (LBG) Algorithm, Image compression, Image indexing, Codebook (CB) generation.

1. INTRODUCTION
Image retrieval is an active area of research. One of the main reason is advancement in technology of image capturing devices, thus empowering the storage of digital images, thus creating the need of image retrieval based on content of image. Earlier the retrieval systems used a traditional approach such as image numbering and text description. In such systems text based query for image retrieval was fired. Such systems have a several limitations. Such as a retrieval result may not contain the desired images, or result is altogether not valid in required context. [4]

So to overcome the drawbacks of such systems image retrieval based on the contents is actively persuade. CBIR is a technique where image features such as color, texture, edge density are used for image indexing and retrieval. For example in a color based CBIR system, a histogram of pixels values is used as a parameter for matching and retrieval. But these systems still have many limitations such as it cannot capture spatial relationship among pixels. But till date such systems are still famous. Based on this discussion it is clear that a color based CBIR systems is required, which can also capture spatial relationship among pixels.[5]

Image compression is also an important field for research. Now a days transmission of image over a network is done on large scale, thus image compression techniques are actively studied. But very few advances have been done to consider the above fields simultaneously. In this paper we present a survey on CBIR using Vector Quantization (VQ). Where Vector Quantization is a technique used for image compression [1]. VQ is used as a feature descriptor for CBIR. By harnessing the characteristics of VQ, the above technique can capture the color feature as well as spatial relationship among pixels.[5][9][13]

The paper organization is as follows. Section 2 and 3 describe technique of vector quantization and codebook generation algorithm. Section 4 describes how VQ is used for image retrieval. Following section 5 concludes the paper.

2. VECTOR QUANTIZATION
Vector Quantization is a simple approach, usually but not exclusively used, for data compression. It has been used for image compression for many years. VQ includes dividing an image into several equal sized vectors (or blocks), and each vector is mapped with a codeword from codebook to encode an image. A vector Quantizer Q of dimensions K and N can be defined as a mapping of K-dimensional Euclidean space R^K into a finite subset of Y of R^K i.e.

$$Q : R^K \rightarrow Y$$

Where $Y = (x_i ; i = 1, 2, \ldots, N)$ and x_i is the ith vector in Y

Y is set called the VQ Codebook or VQ table. This is a set of reproduction vectors of N, meaning it has N distinct elements each a vector in R^K. At the encoder each vector x is approximately matched with a codeword in a codebook, and index of the codeword is transmitted. Euclidean distance or a Manhattan distance can be used for computing a best match codeword. Thus an encoded image containing the indexes is produced. At decoder side the indexes are mapped back to codebook, and original image is generated. The transmitter and receiver should have same codebook. [1-3]

To summarize there are 3 steps involved in Vector Quantization

I. Generating Codebook:
   For the codebook generation, size of the vector must be decided. And divide the image into vectors. After this apply codebook generating algorithm.
II. Encoding Image:
   Encoding means find the best match codeword for each vector and replace the codeword with the index of the matched codeword.
III. Decoding Image :
   Decoding means using the same codebook, for each index, replace it with the respective codeword. And assemble all codewords.

The figure below shows the working of VQ process [13]

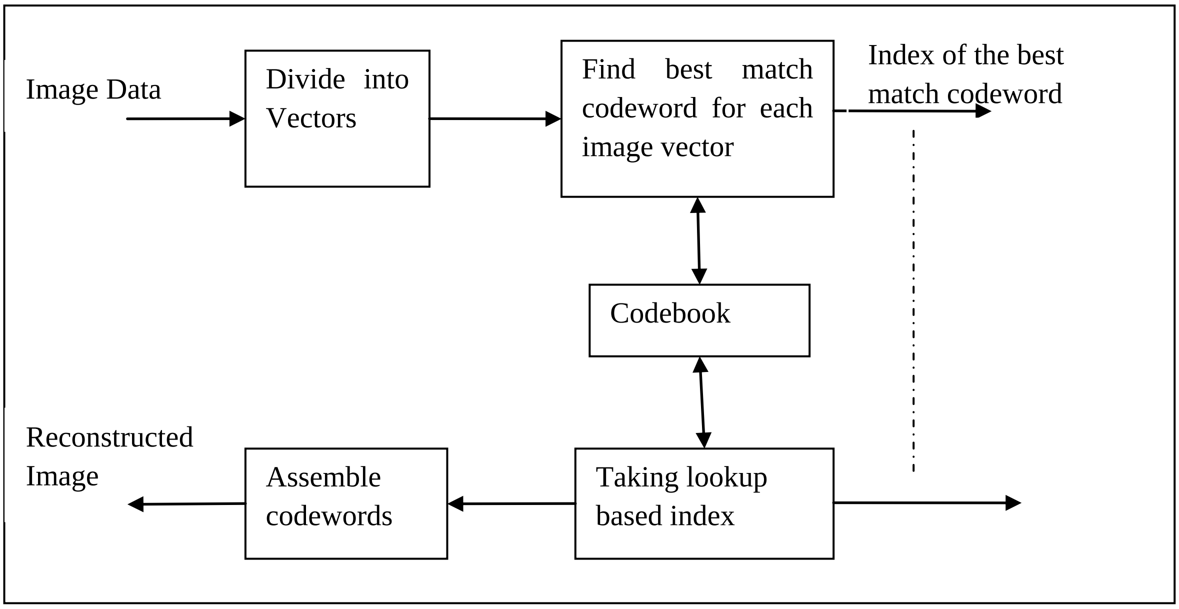

Fig 1: VQ encoding and decoding process

3. CODEBOOK GENERATION ALGORITHMS

The most important part in a process of VQ is the codebook generation. The codebook should be such that it should produce minimum distortion between the original and the reconstructed image. Also the process of codebook generation is time consuming which is another important issue. There are many algorithms suggested over a period of time on VQ, but most commonly used is Linde-Buzo-Gray (LBG) algorithm. For CBIR using vector quantization mostly LBG is used for codebook generation.

LBG was given by . Linde, A. Buzo, and R. M. Gray in 1980 which is also called Generalized Lloyd Algorithm. Before initiating the algorithm image must be divided into several vectors of size (*NXN*)

This algorithm includes the following steps:

1. Randomly generate the initial codebook.
2. Set i=0
3. Perform following for each training vector
   a. Compute Euclidean distance between each training vector and codeword.
      $$d(X,C)=\sqrt{\sum_{i=0}^{k}(x_i - c_i)^2}$$
   b. Search the nearest codeword.
4. Partition the codebook into N cells
5. Compute the centroid of each cell

LBG thus is an easy algorithm for codebook generation.[2-3]

4. VQ FOR IMAGE RETRIEVAL

After an image undergoes process of Vector Quantization a codebook is generated which is representative of the image. As the image is divided into several blocks, an interesting property of this generated codebook is that it captures the spatial correlation among pixels. Which is the key factor used for image indexing and retrieval. Image retrieval works on following condition that, same images will yield same codebook. Thus comparing the codebooks will channel the retrieval. This method is very similar to color feature technique, but has a major advantage. As here a spatial relation among pixels are considered giving a meaningful search. In the following section we have described how this method is used by various authors for image retrieval.

4.1 Method I

Storage and retrieval of compressed images was first introduced by Idris and Panchnathan in 1995. Here they have used VQ for compressing the image database. Initially the image is converted to a gray-scale and is divided into vectors of size (*mXm*), then VQ is applied to an image and a codebook is generated. This generated codebook is used as a training set of codewords for rest of images in the database. Thus all images undergo VQ and one single codebook is evolved over the time. This codebook is a universal codebook. LBG is used as a codebook generating algorithm. Then the images are encoded with this codebook. Hence a set of indexed images is produced. Thus a database of encoded images is produced. In other words image database is compressed using VQ over a single universal codebook. Hence creating VQ-compressed database.

For retrieval query image is also converted to gray-scale and undergoes vector quantization using the same universal codebook. And the image is encoded. Similarity of two images is then found out by measuring the distance between them using following formula

$$S = \sum_{i=0}^{k}(e_i \oplus e_Q)$$

Where S is the similarity distance and e_i indicate encoded image from database and e_Q encoded query image. $\oplus$ This symbol is used to indicate XOR. Thus x-oring the indices of query and images from database the similarities are found. Hence a sorted list of distance between all the images in database is populated, and top *N* images with minimum distance are selected for retrieval. Then these top *N* images are decoded and reconstructed. Thus retrieval is performed. This work was further developed for storage and retrieval of videos.[5-6]

4.2 Method II

Although method given by Idris and Panchnathan yields better result than simple color histogram technique, it still has major drawback. The quality of decoded image was seriously compromised. Also the results were poor for large databases. Thus content based retrieval of compressed images was given by G. Schaefer in 2002. Here author pursues same approach but in different way. VQ based image compression is done for retrieval.

The codeword size is fixed here i.e 4X4 pixels. And so the total vector length is 48 for a color image. No need to convert an image into grayscale. The major difference in this and previous technique is that a codebook per image is generated in contrast to a universal codebook. A codebook for separate image ensures the quality of reconstructed image. There is no need of distributing the universal codebook now. The key feature for this technique is that similar images will have similar codebooks. Thus comparison of codebooks will yield the retrieval result. Hence in this method the author proposed to generate codebook against each image and image compression takes place using respective codebook. Thus a compressed image database in produced and another database containing the codebooks of all images is created. Here also LBG algorithm is used for codebook generation.

For retrieval the query image undergoes the vector process and a codebook is generated C_Q. This codebook is matched with the codebooks in the database say C_D by Modified Huasdroff distance HD defined as

$$HD_{mod}=\max(hd_{mod}(C_D,C_Q),hd_{mod}(C_Q,C_D))$$

With

$$hd_{mod}(C_D,C_Q)= \frac{1}{N}\sum_{i=1}^{N}\min_{j}||C_D(i)-C_Q(j)||$$

Where ||.|| denote some underlying norms. Rather than taking the maximum of minima the author has taken the average of minimum. After calculating the distance with all images, these are ranked and listed in ascending order beginning with minimum distance. And top *N* images are retrieved from compressed image database and reconstructed.[7-8]

4.3 Method III

Approach similar to G. Schaefer is given by A. Daptardar and J. Storer in 2005. Although the above given approach is suggested for compressed database, this approach is suggested for uncompressed database. The working principle of this method is encoding an image with a codebook of a similar image will yield a better representation. Hence here firstly the database is clustered and the each cluster undergoes VQ separately. Before applying VQ all images are converted to gray-scale format, so as to reduce size of codebook.

When one image undergoes VQ, then rest of images of same cluster also undergo VQ using same codebook. Thus this codebook is evolved over the time. This codebook is representative of the cluster, hence *M* codebooks for *M* clusters will be generated. Algorithm used for codebook generation is LBG. There is no restriction on size of codebook.

For a query image, this image undergoes VQ. Then is encoded firstly using the generated codebook. After it is encoded using the codebooks representing clusters. Then the distance between these encoded images is found. To determine the similarity between them a simple retrieval distance which is mean square error is used. This is explained as follows let *Q* be query image and I be image from database. Now *Q* is encoded by *I*'s codebook. Let C_Q and C_I be codebooks of query image and image from database, then the Encoding the distortion distance (EDD) is given by

$$d_{EDD}(Q,I) = \max\{d_{ED}(Q,C_I),d_{ED}(I,C_A)\},$$

Where

$$d_{ED}(X,C)= \frac{1}{Nk}\sum_{i=1}^{N}\min_{j}\|x_i-y_j\|^2$$

is average per component squared Euclidean distance between source vector *X* and code vector *C*. Hence the where the distance is minimum query image is searched in that cluster.

This work was extended further to reduce complexity of image retrieval by same authors. In this works authors include a extended feature vector for retrieval performance. Extended feature vector means additional positional information is calculated and this is added to the feature vector. This positional information is calculated by taking the mean XY coordinates of all the blocks under a particular codeword entry in codebook. These XY co-ordinates are added to the image indicating the position and color of that block. Thus the retrieval complexity is reduced.[9-11]

4.4 Method IV

Another approach for image retrieval by query content is suggested by S. Tang and G. Lu in 2007. According to the authors techniques using color histogram and color correlogram does not consider spatial relation between pixels. So to overcome their drawbacks they have used VQ as a feature descriptor. Unlike previously discussed methods here images are not compressed but the results available on image compression are used. The key factor for working is that two similar images will have similar codebooks, and their encoded images will also look alike. Thus histogram obtained by indexed images is similar for similar images.

The method is described as follows. Size of a codeword is fixed (*mXm*). An image is divided into blocks of pixels as per the decided size. This image than undergoes VQ and a codebook is generated. Now for each image in database VQ is applied, but same codebook is updated. Thus by applying VQ to the database one universal codebook is generated. This codebook is used for retrieval. After the universal codebook is produced each image is encoded using this codebook. As discussed previously encoding means replacing each block in an image with a corresponding index for a codeword in a codebook entry. Thus a database of indexed images is produced. Now for each indexed image a histogram based on the occurrence of each index is calculated. This is used for similarity matching. Algorithm used is LBG, but with slight modification. As a color image is used all the 3 plains R, G, B should be included in Euclidean distance measurement as shown

$$d(X,Y)=\sum_{i=0}^{H-1}\sqrt{\left(X_i^R-Y_i^R\right)^2+\left(X_i^G-Y_i^G\right)^2+\left(X_i^B-Y_i^B\right)^2}$$

Where X and Y are two vectors and H is *mXm* codeword block size.

Query image is encoded with the universal codebook. And an indexed query image is generated, and histogram for this image is calculated. Now similarity measure is computed by calculating the distance between histogram of query image and images in databases. Let a histogram of any indexed image from the database be $H(v_1, v_2, v_3, \ldots\ldots\ldots\ v_i, \ldots\ldots\ldots\ v_n)$, where v_i is the number of times codevector i is used and n is total number of codevectors in a codebook. Let a histogram of query image be $H(q_1, q_2, q_3, \ldots\ldots\ldots\ q_i, \ldots\ldots\ldots\ q_n)$. Then the distance between query image and image in database is calculated as follows

$$D(Q,V)=\sum_{i=1}^{n}\left|H(q_i)-H(v_i)\right|$$

Based on the results of similarity measures, top *N* images with minimum distance are retrieved. These images are retrieved directly from image database. Hence no need to preserve the indexed image database, once the histograms are calculated. According to the authors this method suggested above overcomes the drawbacks of color histogram and color correlogram as well as drawbacks of previously discussed techniques. The major advantage of this method is there is no need of decompression. So errors introduced by decoding are absent. As matching takes place over the indexes there is no need to generate an accurate codebook. Hence the size of codebook can be reduced.[12-13]

4.5 Method V

All the above discussed methods use vector quantization in various ways for effective retrieval. But no such method was proposed which focus on codebook improvement. Such method is proposed by B. Janet, A. Reddy and S. Domnic in 2010. They proposed an incremental codebook generation method for image retrieval. Also they suggest a index cube model for images indexing and retrieval.

In this method special attention is given to the database images. Firstly images are converted to gray-scale and VQ is applied to all images. Each image thus will produce a codebook respectively.LBG algorithm is used for codebook generation. These codebooks are called local codebooks. Now applying LBG algorithm again on all the local codebooks one locally global codebook is generated. This codebook represents all the images in database. Next step is to find the encoding distortion distance d_{Dij} between codebooks of all images. And it is stored as a feature vector. Using locally global codebook encode all images and calculate the histogram of all. Now calculate the histogram intersection d_{Hij} distance between all the images. This is also saved as a feature vector. Similarity measure between two images i and j is given by

$$S= d_{Dij} + d_{Hij}$$

Retrieval is done on the bases of this distance. Major advantage of this method is that any new image can be added to the database at any time. And retrieval complexity is also reduced as less time is required for searching and retrieving images.[14-15]

5. CONCLUSION

This paper presents a survey on content based image retrieval using vector quantization. Although Vector Quantization is a image compression technique, we discussed a various way in which VQ is used for image retrieval. We have shown that retrieval efficiency depends on the accuracy of codebook. Thus codebook generation is area of concern. Further study includes incremental codebook generation process. Also performance of this technique on different databases.

Future research will be to carry out image retrieval on compressed image database.

6. REFERENCES

[1] R. M. Gray "Vector quantization",IEEE Acoustics, speech and Signal Processing Magazine, 1984.

[2] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design,",IEEE Transactions on Communications, pp. 702-710,January 1980..

[3] Tzu-Chuen Lu, "A survey of VQ Codebook Generation",Journal of Information Hiding and Multimedia Signal Processing, July 2010

[4] Ritendra Datta, Dhiraj Joshi, Jia Li, And James Z. Wang, "Image Retrieval: Ideas, Inuences, and Trends of the New Age,",ACM Computing Surveys, 2008.

[5] Idris and S. Panchanathan, "Storage and Retrieval of Compressed Images",IEEE Transactions on Consumer Electronics, 43:5,pp. 937-941,1995.

[6] Idris and S. Panchanathan, "Image and Video Indexing Using Vector Quantization",Machine Vision and Applications, 1997,10:2, pp. 43-50.

[7] Schaefer, G, "Compressed Domain Image Retrieval by Comparing Vector Quantization Codebooks",In: Proc. of the SPIE Visual Communications and Image Processing, 2002, Vol. 4671, pp. 959-966.

[8] Gerald Schaefer, "Content-based Retrieval of Compressed Images",J. Pokorny, V. Snasel, K. Richta (Eds.): Date 2010

[9] Daptardar Ajay H., Storer James A, " Content Based Image Retrieval via Vector Quantization" ,LNCS 3804, pp 502-509 ISVC 2005.

[10] Ajay H. Dapatardar, James A. Storer, "Reduced Complexity Content Based Image Retrieval using Vector Quantization",IEEE transaction 2006

[11] Ajay H. Daptardarand James A. Storer, "VQ based Image Retrieval using Color and Position Features", IEEE transaction Dec 2008.

[12] Shyl Wei Teng, Guojun Lu, "A Novel Image Retrieval Technique based on Vector Quantization", Feb 2004.

[13] Shyl Wei Teng, Guojun Lu , "Image Indexing and Retrieval based on Vector Quantization",ScienceDidect Jan 2007.

[14] B. Janet, R. V. Readdy, S.Domnic "Incremental Codebook Generation for Vector Quantization in Large Scale Content Based Image Retrieval",Communications in Computer and Information Science BAIP, CCIS 70, pp. 978-1-4244-5967-4/10- 2010 IEEE.

[15] B. Janet, A.V. Reddy, and S. Domnic, "Vector Quantization based Index Cube Model for Image Retrieval",Proc. PSIVT 2010.