*Journal of Information Hiding and Multimedia Signal Processing* ©2018 ISSN 2073-4212
**Ubiquitous International** Volume **9**, Number **4**, July **2018**

# Image Retrieval Using VQ Index Histograms in Classified DCT Domain

Ya-Pei Feng[1] and Zhe-Ming Lu[2*]

School of Aeronautics and Astronautics
Zhejiang University
Hangzhou, 310027, P. R. China
[*] Corresponding author: zheminglu@zju.edu.cn

Received October, 2017; revised February, 2018

Abstract. *An image retrieval scheme, called Classified DCT based Vector Quantisation Index Histogram (CDCT-VQIH) is proposed to extract features from the DCT frequency domain. Our algorithm takes advantage of both the energy-compaction property of DCT and the high compression ratio of classified vector quantization (CVQ). The input colour image is decomposed into Y, Cb, and Cr components, and each component image is divided into non-overlapping* $8 \times 8$ *blocks. Then we transform the image blocks from the spatial domain to the DCT domain. As a result, we can get a transformed vector consisting of DC and AC coefficients. After that, we perform CVQ on the AC coefficients in each block to classify all the* $8 \times 8$ *blocks into four categories. Our CDCT-VQIH based feature combine the class index, the DC index and the VQ index together, which represent the edge classification information, energy information and texture information of an image respectively. The retrieval simulation results show that, compared with the traditional spatial-domain colour-histogram (SCH) based feature and the existing DCT-VQIH based feature, our feature performs much better in terms of recall and precision.*
**Keywords:** Classified DCT based vector quantization, Vector quantization, Image indexing, Edge classification information, Energy information, Texture information

1. **Introduction.** With the rapid increase of demand on data transfer and data storage, more and more visual information is compressed and stored in databases. As for image compression, the main purpose is to reduce the redundant information of the images. Generally speaking, there are three types of redundancies, spatial redundancy, spectral redundancy and temporal redundancy [1]. The spatial redundancy is due to correlation among neighbouring pixels which means neighbouring pixels are not statistically independent. The spectral redundancy is due to different colour planes and spectrum bands. The temporal redundancy is due to correlation of different frames of an image. Image compression aims to reduce the number of bits required to represent an image by removing the spatial and spectral redundancies of an image.

There are mainly two kinds of compression techniques: frequency domains based and the spatial domain based techniques. The compression techniques based on frequency domains transform digital data into coefficients, and the values of most of the coefficients are very small, requiring only a few binary bits to represent them. The coefficients that have high values are quantified to decrease their values, thereby accomplishing the compression of the data. Typical kinds of compression methods include the discrete cosine-transform (DCT) [2] based, discrete wavelet-transform (DWT) [3] based, and block truncation coding (BTC) [4] based schemes. The second kind of compression technique is based on the spatial domain, such as Huffman coding [5] and vector quantization (VQ) [6]. In 1952, Huffman proposed a variable length code to solve the problem [5]. In this method, the symbols appearing at lower frequencies are encoded as longer variable length codes. In contrast, the symbols appearing at higher frequencies are encoded as shorter variable length codes. The Huffman encoding method can efficiently compress symbols and successfully decode compression codes. In 1984, Gray proposed a VQ-based encoding method to compress images [6]. In this method, the image is divided into several blocks, which are further mapped to the most similar codewords in the selected VQ codebook and the index of the codeword is used as a compression code. In other words, all the pixels in the image block are replaced by a VQ index. Therefore, the VQ compression method has excellent ability to compress images. After each image block is compressed, we can get the VQ index table. In general, the difference between the image block and its adjacent blocks is small.

As one of popular image compression techniques, VQ has been used in several image retrieval schemes because of its natural classification and clustering characteristics. For example, classified VQ (CVQ) [7], DCT domain VQ (DCTVQ) [8], block truncation coding based VQ (BTCVQ) [9] and edge orientation pattern based VQ (EOPVQ) [10], etc. In the CVQ based scheme [11], a classifier is used to determine the class for each block, and the block is then coded with a VQ codebook designed specifically for that class. In the DCTVQ based scheme [8], researchers extract 12 DCT-domain vector quantisation index histograms based on the YCbCr colour space. Reference [9] proposed an effective feature for colour image retrieval based on two pattern co-occurrence matrices generated from the BTC compressed Y image and VQ compressed Cb and Cr images, respectively. Reference [10] presented a new edge orientation (EO) detector compatible with VQ-based feature extraction. VQ provides a number of attractive features for image coding with high compression ratios. However, the initial study using VQ's image coding revealed several difficulties, most notably edge degradation and high computational complexity. Overall consideration, DCT has the excellent energy compact property and the superiority of CVQ is high compression ratio and low complexity. To get better image retrieval efficiency, we choose Classified DCT based Vector Quantisation Index Histogram (CDCT-VQIH) to overcome the difficulties.

2. **Related works.** As we all know, the analog signal is a continuous value, and the computer can only handle digital signals. When converting an analog signal to a digital signal, we can use a value in the interval instead of an interval, for example, all the values on [0, 1] become 0, and all the values on [1, 2] become 1, and so on. This is the process of a VQ. VQ is a classical quantization technique that allows the modelling of probability density functions by the distribution of prototype vectors. It works by dividing a large set of vectors into groups having approximately the same number of points closest to them. Each group is represented by its centroid point, as in k-means and some other clustering algorithms. VQ also provides an image coding method with a high compression ratio and a simple table lookup decoding structure which has been used in many research areas, such as image compression, image retrieval, image recognition and image restoration.

As for VQ based compression, there are three main steps in the VQ process: (i) the codebook design, (ii) the encoding process and (iii) the decoding process. Initially, the image is divided into non-overlapping blocks of sub-images of the same size. During codebook design process, codebooks should be generated based on the large training vector set using the well-known Linde-Buzo-Gray (LBG)[12] algorithm and the most similar codeword is chosen for each training vector. The codebook is a lookup table that contains the codeword and its index, that is, the location in the lookup table in a prescribed order. During the encoding process, each training vector is replaced by the index of the most appropriate representative codeword. In the decoding process, decoder has the same codebook as the encoder. For each index, the decoder merely performs a simple table look-up operation to reconstruct the input training vector. To sum up, in VQ, each image block, rather than a single pixel, can be quantized and coded by an index. Thus, VQ can provide an effective means for image indexing and image retrieval.

However, the study of image coding using VQ alone reveals several difficulties, most notably edge degradation and high computational complexity. So we need to explore new ways to make up for these shortcomings of VQ.

2.1. **DCT-VQ Compression Algorithms.** DCT has the excellent energy compact property, thus we can throw the high-frequency information and only perform VQ on the low-frequency coefficients. DCT transforms usually produce signal energy distribution in a small set of transform coefficients. Assume an image is divided into non-overlapping blocks of size $N \times N$, and f(x,y) represents the pixel values of the original block, then the two-dimensional DCT can be defined as follows[12]:

$$F(i,j) = \frac{2}{n} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} c(i)c(j)f(x,y) \cos\frac{(2x+1)i\pi}{2N} \cos\frac{(2y+1)j\pi}{2N} \tag{1}$$

Where $c(i) = c(j) = \begin{cases} 1/\sqrt{2} & \text{i=j=0} \\ 1 & \text{others} \end{cases}$, and $F(i,j)$ represent the coefficients in the block after applying DCT, and $0 \le i, j \le N-1$. Supposing an image is divided into blocks of size $8 \times 8$, the two-dimensional DCT is performed on each block. As a result, we can obtain a DCT matrix for each block according to the zigzag scanning order [18],in which there are 64 coefficients, and the coefficient $F(0,0)$ is defined as the DC coefficient and the remaining 63 coefficients are called AC coefficients. The image is reconstructed from the transformed image $F(i,j)$ by applying the inverse 2D-DCT according to the following equation.

$$f(x,y) = \frac{2}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} c(i)c(j)F(i,j) \cos\frac{(2x+1)i\pi}{2N} \cos\frac{(2y+1)j\pi}{2N} \tag{2}$$

The existing DCT-VQ based image compression takes advantage of the DCT and VQ. For example, reference [13] studied the use of energy histograms of the low frequency DCT coefficients as a feature to retrieve DCT compressed images. Reference [8] proposed a feature for colour image retrieval based on DCT-domain vector quantisation (VQ) index histograms. In this algorithm, authors divided the DCT blocks into four parts and encoded them with different codewords and extract 12 histograms from the 12 DCT-VQ index sequences with high compression ratio and high image retrieval efficiency. It is proved that DCT-VQ can obtain better performance than the spatial-domain VQ.

2.2. **Classified VQ algorithms.** In CVQ, the image is divided into blocks, and the blocks are classified into various classes. A classifier determines the class of each block, and the block is then coded with a vector quantiser designed specifically for that class. The size of codebook for each class can vary, which makes up the total codebook. Reference [14] proposed a method of natural image classification by an effective color quadtree segmentation together with a more effective codebook with the colour local thresholding classifier for content-based image retrieval, where quadtree segmentation and local thresholding classifier are applied to classify all the blocks. Blocks with distinct perceptual features, such as edges, are generated from different sub-sources, i.e., belonging to

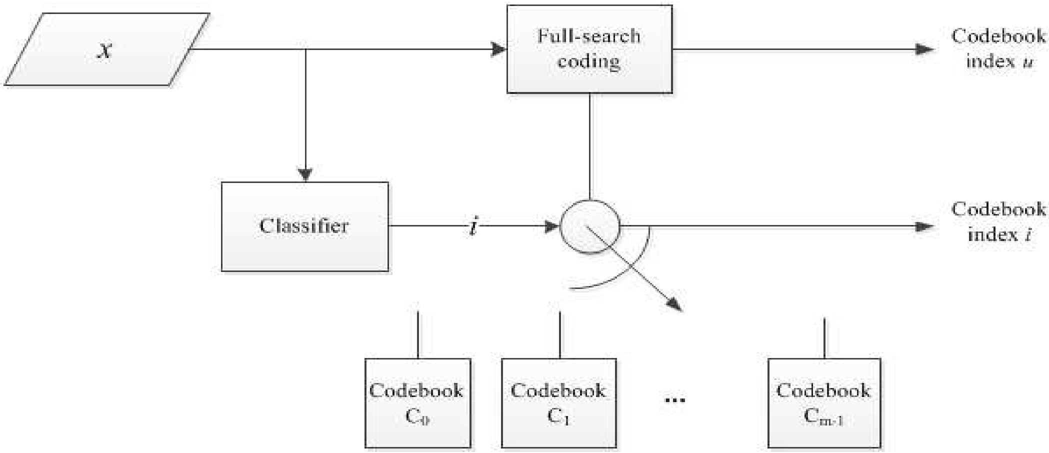


FIGURE 1. The block diagram of classified VQ

different classes. Time complexity is lower since the size of codebook for each class is smaller. Supposing there are $M$ classes, and if the input $X$ belongs to Class $i$, the i-th subcodebook $C_i$ of size $N_i$ is employed to encode $X$. The classifier generates a codebook index $i$, $i$ varies from 0 to $m-1$. The input vector $x$ search for the nearest code in the codebook $C_i$, and generate the index $u$. The index of the nearest codeword is transmitted to the decoder. The decoder simply retrieves the corresponding codeword from its codebooks, generating the output $X$. Reference[19] proposed the classified side-match vector quantization to achieve low-bit-rate image encoding. Blocks are classified into two level categories according to the variances of the upper and left codewords of the input vector. Besides,two different master codebooks are used for generating the state codebook according to the variance of the input vector. Experimental results prove the effectiveness of the proposed algorithm.

In CVQ, the image is divided into blocks, and the blocks are classified into various classes. The classifier determines the class for each block, and the block is then coded with a vector quantizer designed specifically for that class. The size of codebook for each class can vary, which make up to the total codebook. Time complexity is lower since the size of codebook for each class is smaller. The CVQ coder is depicted in Fig. 1.

Chen et al. [7] performed the quadtree segmentation and introduced the CVQ to improve the capability of describing the variety of texture blocks that is constrained by the VQ. They proposed a color image retrieval scheme that introduces hue and gray-level information to the quadtree segmentation and extract the low-detail and high-detail regions. Then they design edge binary templates to classify high-detail regions and use the color local thresholding classifier which uses the color edge detection technique, the distortion measure parameters so that it can provide better natural-scene image classification performance and more effective descriptions to the image. In a word, CVQ is an efficient and low time-complexity image processing tool.

## 3. Proposed Scheme.

### 3.1. CDCT-VQ based compression.

In this section, a new image retrieval scheme combing the $DCT-VQ$ and the CVQ is proposed. Image compression schemes apply the CVQ coding in the DCT domain have made great progress. Kim et al. [15] proposed an image coding scheme employing the CVQ in the DCT domain to realize image compression. Authors proposed transformed classified vector quantization which is a hybrid image compression method that takes advantage of the energy-compaction property of

DCT and the high compression ratio of CVQ. In order to achieve high compression ratio, the transformed classified VQ uses two stages of compression. In the first stage, image is transformed from the spatial domain to the frequency domain and then quantized. In the second stage, the quantized nonzero AC coefficients are compressed again using the CVQ by classifying the image blocks into different categories and coding them with different codebooks specifically.

In addition, Tseng and Chang [11] proposed a hybrid compression algorithm based on CVQ in the DCT domain, which divides DCT blocks into shade blocks and edge blocks according to the coefficients of DCT blocks. In their algorithm, according to the 63 AC coefficients, they classified the image blocks into four classes. As the DC coefficients of the adjacent blocks are close to each other, the DC coefficients are encoded by DPCM, while the AC coefficients are encoded with CVQ. According to the distribution characteristics of AC coefficients and the block classification algorithm, each block is classified into one of the following four classes: shade block, horizontal block, vertical edge block and diagonal edge block. The classification steps during the codebook training process can be briefly illustrated as follows [11]:

Step 1.The coefficients of a DCT matrix are named as $C_0$, $C_1$, $\ldots C_{63}$, compute $V = max(|C_1|, |C_5|, |C_6|, |C_7|, )$ and $H = max(|C_2|, |C_3|, |C_8|, |C_9|, )$.

Step 2. Set the threshold as 45, and all the overlapping $8 \times 8$ image blocks are classified into four categories, and the classifier is designed according to the following equation:

$$Theinputblock \begin{cases} shadeblock & V < \Gamma, H < \Gamma \\ diagonalblock & V \geq \Gamma, H \geq \Gamma, \frac{MAX(V,H)}{min(V,H)} < 2 \\ horizontalblock & H \geq V \\ verticalblock & otherwise \end{cases}$$

Step 3. Different edge blocks are trained to generate their corresponding codebooks. In this procedure, we apply the Linde-Buzo-Gray(LBG) algorithm which approximates the optimal regenerative codebook by training the vector set and a certain iterative algorithm. After obtaining codebooks, the coding process is mainly divided into three steps, i.e., classification of each input block in the DCT domain, encoding each DCT block using its corresponding DCT domain codebook, and inverse transformation of each DCT block into its spatial block.

The block diagram of CDCT-VQ based compression is illustrated in Fig. 2. The above method takes advantage of both DCT and CVQ and shows good compression performance at a low bit rate. In image retrieval, it is vital to encode the training images and the database images to gain more similar information between them. So an efficient image compression algorithm plays a vital role. Inspired by the CDCT-VQ based compression algorithm, we apply some of its idea in our image retrieval scheme.

3.2. **CDCT-VQ based feature extraction.** In this section, the proposed algorithm for image retrieval is described in detail. In our algorithm, simple features combining the CDCT based features with the traditional VQ based features are proposed. In this paper, YCbCr images are considered. For each colour component (such as Y Cb Cr) of an input colour image, we divide it into non-overlapping blocks of size $8 \times 8$, and then perform CDCT classification on each block. Thus, we can classify all blocks into four classes, and we can obtain a CDCT index histogram (CDCTIH) based on the number of blocks belonging to each class. The CDCTIH consists of the histogram of classification indices and the histogram of the DC coefficients which are cascaded. In addition, we encode all blocks with a predesigned VQ codebook consisting of $N$ codewords to obtain an index sequence, and then calculate a VQ index histogram (VQIH) from this sequence.

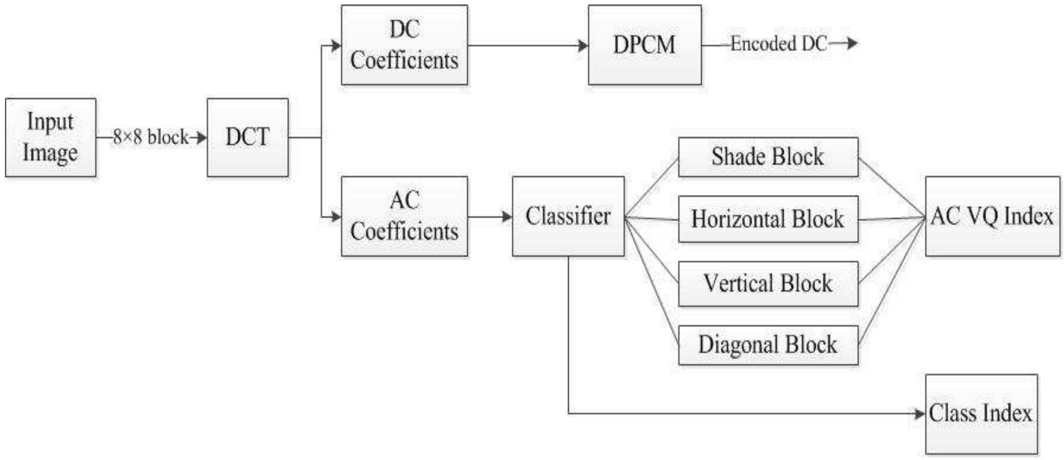


FIGURE 2. Block diagram of CDCT-VQ based compression

Finally, we can obtain a feature by combining CDCTIH with VQIH for each colour image with three colour components. The detailed steps of the proposed retrieval process can be outlined as follows:

Step 1. We divide every input image into overlapping blocks of size $8 \times 8$, and perform the two-dimensional DCT on each $8 \times 8$ block and thus get a DCT matrix consisting of 64 coefficients: one DC coefficient and 63 AC coefficients.

Step 2. Encode the DC coefficients of all blocks with scalar quantisation. Just as the classification method used in Tseng and Changs algorithm, we classify the blocks into four categories and encode them with their corresponding codebooks.

Step 3. Count the number of blocks belonging to each category for each image, obtaining its class index histogram. Similarly, we can also obtain DC index histogram. In addition, we encode all blocks with a predesigned corresponding codebook with $N$ codewords to obtain an index sequence, and then calculate a VQIH from this sequence. Combined with the VQ index histogram, our CDCT-VQ based feature consists of three parts: the class index histogram, DC index histogram and VQ index histogram. It is worth mentioning that our class index and DC index are cascaded together to express the classification feature and the energy information of an image. Overall, our CDCT-VQ based feature is a kind of multi-feature image retrieval method or a fusion of edge, energy and texture features.

After we have extracted the image needed features, the next step is to match the training images with the images from our selected database. As our features are presented in the form of index histograms, we choose the precision and recall as our evaluation criterion. Here, the precision and recall are defined as follows:

$$precision = \frac{NO.relevantimage}{No.imagesreturned} \tag{3}$$

$$recall = \frac{No.relevantimages}{100} \tag{4}$$

4. **Experimental results and discussion.** We perform simulations on a standard database [16] in the experiment that is carried out on a Pentium IV computer. This database includes 1000 images of size 384×256 or 256×384, which are classified into ten classes, each class including 100 images ($Class1\_people$, $Class2\_beach$, $Class3\_buliding$, $Class4\_bus$, $Class5\_dinosaur$, $Class6\_elephant$, $Class7\_flower$, $Class8\_horse$, $Class9\_mountain$ and

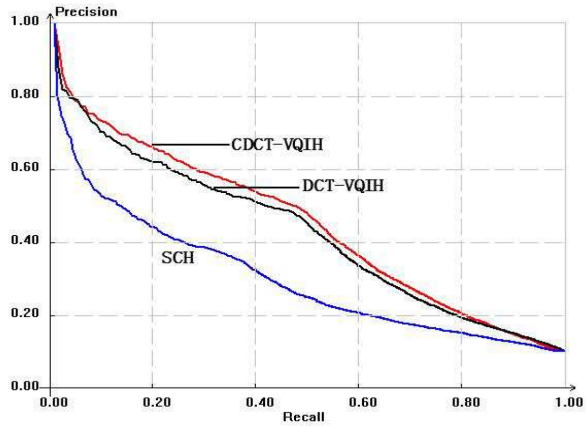


FIGURE 3. The comparisons of the average P-R curves of SCH, DCT-VQIH and CDCT-VQIH with the same dimension of 192

$Class10\_food$). We randomly select five images from each class, i.e., in total 50 images, as the test query images. For each test query image, we perform the retrieval process based on each kind of features. For each number of returned images (from 1 to 1000), we average the recall and precision value over 50 test query images.

In our simulation experiments, we compared our features with the DCT-VQIH-based feature [8] and the traditional spatial-domain colour-histogram-based (SCH) feature [17] based on the same YCbCr colour space and the same feature dimension of 192. On the basis of the same training images, the dimension of the classification feature is 4 and the dimension of the DC coefficient based index histogram is 16, thus the cascaded dimension is $4\times16 = 64$. So the dimension of the VQ index histogram is 128. Thus, our feature based on CDCT-VQIH is also with the dimension of 192. For SCH based feature, we extract three 64-bin colour histograms from each image based on the YCbCr colour space for comparisons, i.e. the SCH feature is also with the dimension of 192. We also generate 12 codebooks for DCT-VQIH feature extraction, and the feature dimension for DCT-VQIH is $12 \times 16 = 192$. To show the comparison more clearly, we adopt the precision-recall (P-R) curves to evaluate the performance and effectiveness of retrieval methods. The comparison curves of the simulation results are shown in Fig. 3. As we can see from the curves in Fig. 3, more steep the curve is, the higher efficiency of image retrieval algorithm is.

To see the performance of our proposed algorithm more directly, we give a further comparison. Here, we show the example of retrieved images from the natural-scene image database[16] for the SCH and traditional DCT-VQ based schemes and the proposed CDCT-VQ based scheme in Table 1. All of the images in the database are sorted according to the order from the most similar one to the least similar one. We show only the first 16 retrieved images with the highest similarities due to the limitation of the space and rank the average distance (from $Dis.1$ to $Dis.16$) between the query image and the retrieved images from low to high in the image database[16].

In Fig. 3, we can see that the recall performance of proposed CDCT-VQ based scheme is higher than the SCH and traditional VQ based schemes; moreover the error rates which are the values of dis shown in Table 1 are much lower. That is to say, when we provide an example image, our proposed scheme can obtain more similar results.

5. **Conclusion.** In this paper, we propose a new CDCT-VQ based feature for image retrieval, which can largely improve the image retrieval efficiency in the DCT domain. A

| Distance | SCH | DCT-VQIH | CDCT-VQIH |
|---|---|---|---|
| Dis.1 | 0 | 0 | 0 |
| Dis.2 | 0.0321567 | 0.0287770 | 0.0096657 |
| Dis.3 | 0.0361357 | 0.0291109 | 0.0119752 |
| Dis.4 | 0.0413837 | 0.0302656 | 0.0125537 |
| Dis.5 | 0.0422838 | 0.0306834 | 0.0134898 |
| Dis.6 | 0.0426499 | 0.0360775 | 0.0135104 |
| Dis.7 | 0.043003 | 0.0371170 | 0.0147509 |
| Dis.8 | 0.0447642 | 0.0376319 | 0.0171733 |
| Dis.9 | 0.0454774 | 0.0402941 | 0.0172487 |
| Dis.10 | 0.0466318 | 0.0414964 | 0.0174780 |
| Dis.11 | 0.04664434 | 0.0436088 | 0.0182120 |
| Dis.12 | 0.0483499 | 0.0442917 | 0.0187861 |
| Dis.13 | 0.0495111 | 0.0450602 | 0.0189779 |
| Dis.14 | 0.0499458 | 0.0451644 | 0.0205198 |
| Dis.15 | 0.0507784 | 0.0466417 | 0.0209840 |
| Dis.16 | 0.0508733 | 0.0499752 | 0.0216015 |

TABLE 1. Comparison of the average distance between query image and the retrieved images with the three schemes.

detailed classifier is designed to classify the blocks in the DCT domain into four different categories, which is used to further classify visually important edge-dense blocks. In addition, our feature combines the classification histogram and energy histogram with the VQ index histogram, being also a kind of multi-feature including the edge classification information, energy information and texture information. As the comparison results shown in Fig. 3 and Table 1, the proposed method has better retrieval accuracy when compared to the SCH based scheme and the existing DCT-VQ based scheme. In the future research, we will focus on exploring more efficient and high-performance features.

**Acknowledgments.** This work was supported partially by the Nature Science Foundation of Guangdong Province (2015A030310172), the Science and Technology Plan Projects of Shenzhen (JCYJ20150324140036830, GJHZ20160226202520268), and the Zhejiang Provincial Natural Science Foundation of China under grant No. LZ17F030008.

## REFERENCES

[1] S. Dhawan, A review of image compression and comparison of its algorithms,*International Journal of electronics and Communication technology* ,vol. 2, no.1, pp.22–26, 2011.

[2] N. Ahmed, T. Natarajan, and K. R. Rao, Discrete cosine transform, *IEEE transactions on Computers*, vol.100, no.1,pp.90–93, 1974.

[3] M. J. Shensa, The discrete wavelet transform: wedding the a trous and Mallat algorithms,*IEEE Transactions on signal processing*, vol.40, no.10,pp. 2464–2482, 1992.

[4] E. Delp, and O.Mitchell, Image compression using block truncation coding,*IEEE transactions on Communications*, vol.27, no.9, 1979.

[5] A. Belal,A. Elmasry, Verification of minimum-redundancy prefix codes, *IEEE Transactions on Information Theory*, vol.52, no.4,2006.

[6] P.C. Cosman, K.L. Oehler,E.A. Riskin , Using vector quantization for image processing *Proceedings of the IEEE* vol.81, no.9, 2014.

[7] H. H. Chen, J, J, Ding, and Sheu, H. T, Image retrieval based on quadtree classified vector quantization, *Multimedia tools and applications*, vol.72, no.2, pp.1961–1984, 2014.

[8] Z. M. Lu and H. Burkhardt, Colour image retrieval based on DCT-domain vector quantisation index histograms,*Electronics Letters*, vol.41, no.17, pp.956–957,2005.

[9] F. X. Yu, H. Luo and Z. M. Lu, Colour image retrieval using pattern co-occurrence matrices based on BTC and VQ, *Electronics letters*, vol.47, no.2, pp.100–101, 2011.

[10] Z. M. Lu and Y. P. Feng, Image retrieval based on histograms of EOPs and VQ indices, *Electronics Letters*, vol.52, no.20, pp.1683–1684,2016.

[11] H. W. Tseng and C. C. Chang, A very low bit rate image compressor using transformed classified vector quantization,*Informatica*, vol.29, no.3, pp.335–341,2005.

[12] Y. C. Lin and S. C. Tai, A fast Linde-Buzo-Gray algorithm in image vector quantization, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol.45, no.3, pp. 432–435, 1998.

[13] Y. C. Lin and S. C. Tai, Image compression based on discrete cosine transform and multistage vector quantization, *International Journal of Multimedia and Ubiquitous Engineering*, vol. 10, no.6, pp. 347–356, 2015.

[14] H. H. Chen and J. J. Ding,Image retrieval based on classified vector quantization using color local thresholding classifier. *Image Processing (ICIP)*, 2012 19th IEEE International Conference on. IEEE , 2012.

[15] D. S. Kim and S. U. Lee, Image retrieval based on quadtree classified vector quantization, *Multimedia tools and applications*, vol.72, no.2, pp.1961-1984, 2014.

[16] J. Li, Photography image database. http://www.stat.psu.edu/ jiali/index.download.html.

[17] R. Chaudhari , A. M. Patil, Content based image retrieval using color and shape features, *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering*, vol.1, no.5,2013.

[18] R. Atta, M. Ghanbari, An efficient face recognition system based on embedded DCT pyramid, *IEEE Transactions on Consumer Electronics*, vol. 58, no.4, pp.1285-1293, 2013.

[19] Z. M. Lu, J. S. Pan and S. H. Sun, Image Coding Based on Classified Side-match Vector Quantization, *IEICE Transactions on Information and Systems* , vol. E83-D, no.12, pp.2189-2192, 2000.