International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.6 (2015), pp.347-356
http://dx.doi.org/10.14257/ijmue.2015.10.6.33

Image Compression Based on Discrete Cosine Transform and Multistage Vector Quantization

Xiao Zhou*, Yunhao Bai and Chengyou Wang

School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, P. R. China
zhouxiao@sdu.edu.cn, 814401570@qq.com, wangchengyou@sdu.edu.cn

Abstract

In this paper, an image compression scheme is proposed, based on discrete cosine transform (DCT). This scheme is a hybrid method, which combines vector quantization (VQ) and differential pulse code modulation (DPCM). This scheme begins with transforming image from spatial domain to frequency domain using DCT. Then the block data is transformed into a vector according to zigzag order, and then truncated. After that, the vector is split into DC coefficient and AC coefficients. After scale quantization, DC coefficient is coded using DPCM. AC coefficients are coded using multistage vector quantization (MSVQ). Then, entropy encoding is performed on index-tables and DC part, separately. The experimental results show that, compared to conventional VQ and DCT-VQ schemes, proposed scheme has a better performance.

Keywords: *image compression; discrete cosine transform (DCT); multistage vector quantization (MSVQ)*

1. Introduction

Image compression has become an important research area for many years due to increasing demand on transfer and storage of data. The most used and successful image-encoding standard should be joint photographic experts group (JPEG) standard, also known as ITU-81 [1] published at 1991. Because JPEG encoding can be implemented quickly, around 80% of the images on the Internet observe JPEG standard. However, JPEG shows its limitation when it comes to high compression ratio. Users can recognize this defect, also known as blocking artifacts, easily. Lossy JPEG encoding process is shown in Figure 1.

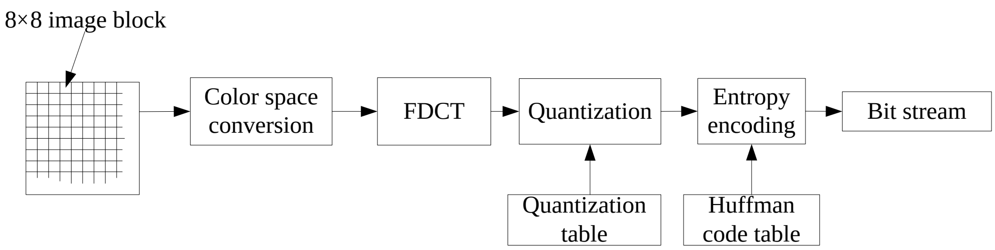


Figure 1. The Scheme of JPEG Encoding Process

Vector quantization (VQ) is one of the popular lossy image compression techniques because it has a simple decoding structure and can provide high compression ratio in image coding systems. Vector quantization can be regarded as the process of finding the nearest vector in the codebook to the waiting vector. When the nearest vector is given, only the index of the vector in the codebook is

ISSN: 1975-0080 IJMUE

transmitted. Moreover, compared to scale quantization (SQ), VQ is always superior even if the information source is non-memorable. If a high compression ratio is required, VQ can be a great choice.

The paper is organized as follows. After this introductory Section, we present a brief overview of the block based DCT and some improvement of VQ in Section 2. The proposed scheme for image compression is elaborated in Section 3. Experimental results are given in Section 4 to compare the performance of the proposed scheme with some other schemes based on VQ. Finally, conclusions and remarks on possible further work are given finally in Section 5.

2. DCT and Improvements of VQ

2.1. DCT

Discrete cosine transform is an orthogonal transform method proposed by N. Ahmed *et al*. [2] in 1974. DCT has been widely applied in image processing research since it was proposed. If a function is both real and even, then its Fourier series contains only cosine terms, then discrete cosine transform can be derived from discrete Fourier transform (DFT). Two-dimensional DCT can be defined as follows:

$$X(u,v)=\frac{2}{N}\sum_{n_2=0}^{N-1}\sum_{n_1=0}^{N-1}c(u)c(v)x(n_1,n_2)\cos\left[\frac{(2n_1+1)u\pi}{2N}\right]\cos\left[\frac{(2n_2+1)v\pi}{2N}\right] \quad (1)$$

where $n_1,n_2,u,v=0,1,\cdots,N-1$, $x(n_1,n_2)$ represent coefficients data in original block, $X(u,v)$ represent coefficients in the block after applying DCT. $c(u)=c(v)=\begin{cases}1/\sqrt{2}, & u=0,v=0,\\ 1, & \text{others.}\end{cases}$. Inverse discrete cosine transform (IDCT) can be defined as follows:

$$x(n_1,n_2)=\frac{2}{N}\sum_{v=0}^{N-1}\sum_{u=0}^{N-1}c(u)c(v)X(u,v)\cos\left[\frac{(2n_1+1)u\pi}{2N}\right]\cos\left[\frac{(2n_2+1)v\pi}{2N}\right] \quad (2)$$

Through DCT, correlations between each block can be better extracted, which is conducive to data compression. Thus, DCT is considered as a “quasi-optimal transform”. Due to some historical reasons, the size of each image block is 8×8 identically. Compared with DFT, DCT can be performed only using real number. Therefore, DCT is superior to DFT in computational complexity. Moreover, Fast algorithm for DCT, like algorithm proposed in [3] is also available to reduce computing cost.

After performing DCT, the block can be divided into 2 sub-bands: low frequency sub-band which contains most of the important visual parts of the image, and high frequency sub-band which contains details and textures of the image. Generally speaking, low frequency coefficients are more important than high frequency coefficients because the values of high frequency coefficients are usually closed to 0. Of all the DCT coefficients, $X(0,0)$ is called DC coefficient generally, and other coefficients is called AC coefficients. Due to non-importance of high frequency sub-band, in general, the high frequency sub-band is usually removed for compression purpose.

2.2. Improvements of VQ

Since LBG algorithm [4] was proposed, many kinds of vector quantizers have been proposed, such as tree-structured VQ (TSVQ) and classified VQ (CVQ). Due to the reliance on initial codebook and too many times of iteration, many methods are proposed to ameliorate classical LBG algorithm. To improve speed of codebook designing, fast code words search technique can be introduced. Moreover, other codebook-designing algorithm can be combined with LBG algorithm to improve quality of the final codebook. VQ can be also performed in the frequency domain. For example, DCT can be performed on the image block at first, then the vectors are truncated in order to discard high frequency sub-band in order to reduce the complexity of the quantizer. In recent years, many researches are carried out combining VQ and discrete cosine transform or wavelet transform. A. K. Pal *et al*. [5] and D. K. Mahapatra [6] proposed a hybrid DCT-VQ approach for efficient compression [5-6], which performs VQ in frequency domain using truncated vector directly. Proposed algorithm is shown in Figure 2.

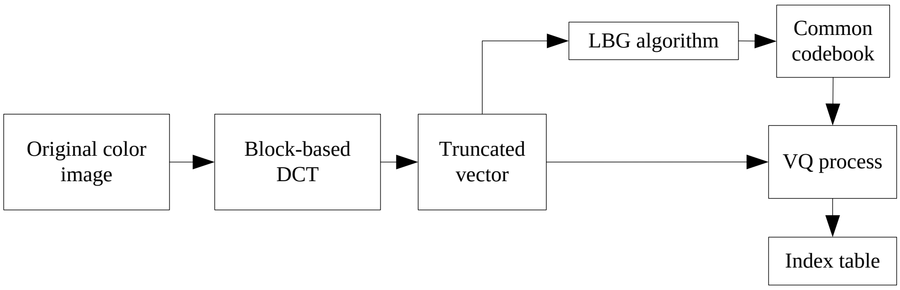


Figure 2. Block Diagram of Arup Kumar Pal's Algorithm

A. K. AI-Asmari [7] also made some improvement to conventional VQ compression method. In his paper, predictive locally adaptive vector quantization [7] is introduced, enable progressive image transmission. R.-J. Liou's method [8] is based on DCT sub-band decomposition, using self-organizing feature map (SOFM) to generate the codebook. H.-W. Tseng [9] proposed his algorithm on CVQ [9], dividing DCT blocks into shade blocks and edge blocks according to some criteria which is experimentally set. At bit rate of 0.16bpp, a peak signal to noise ratio (PSNR) value of 29.15dB could be achieved. H.-W. Tseng's algorithm is shown in Figure 3.

3. Proposed Algorithm

In this section, a new image-coding scheme is proposed using VQ and DCT, based on JPEG standard. Compared to other schemes that have been proposed, this algorithm can achieve a higher PSNR value yet preserve high quality of image at low-bit rate situation.

3.1. MSVQ

Multistage vector quantization (MSVQ) [10] can not only reduce the codebook size and computing time of full-search VQ, but also can enjoy small quantization distortion compared to tree-structured VQ. The main idea of MSVQ is that the whole VQ process is carried out in several successive stages. The first-stage vector quantizer only performs a relatively crude quantization of the input vector $\boldsymbol{X}$, then output the index i_1 in the codebook $\boldsymbol{C}_1$, and an error vector $\boldsymbol{e}_1 = \boldsymbol{X} - \boldsymbol{Y}_1$. Then, The second-stage quantizer operates on the error vector $\boldsymbol{e}_1$ outputted by the first-stage quantizer, providing an index in codebook and an error vector $\boldsymbol{e}_2 = \boldsymbol{e}_1 - \boldsymbol{Y}_2$. The third-stage quantizer may be performed on the error vector $\boldsymbol{e}_2$ and so on. The block diagram of MSVQ is illustrated in Figure 4.

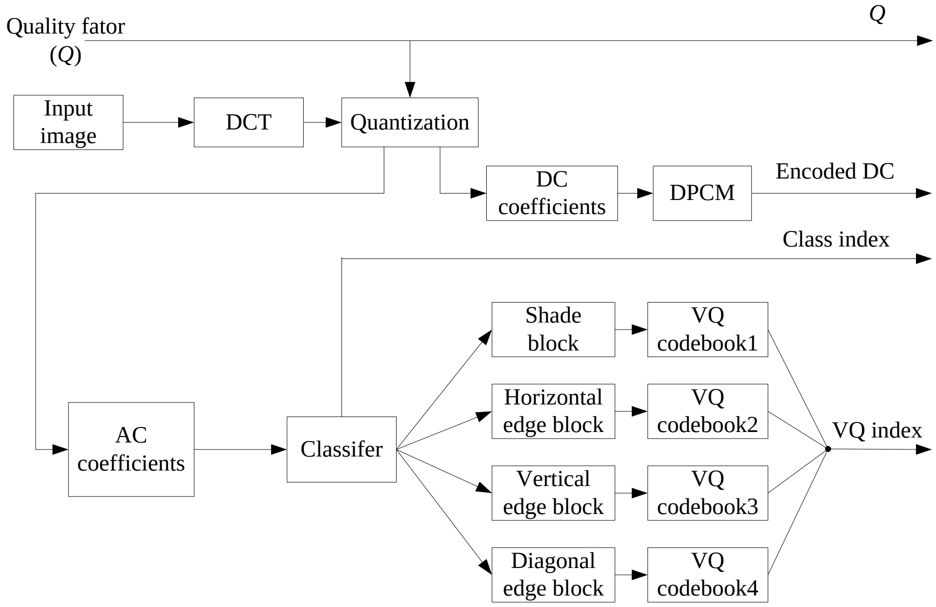


Figure 3. Block Diagram of H.-W. Tseng's Algorithm

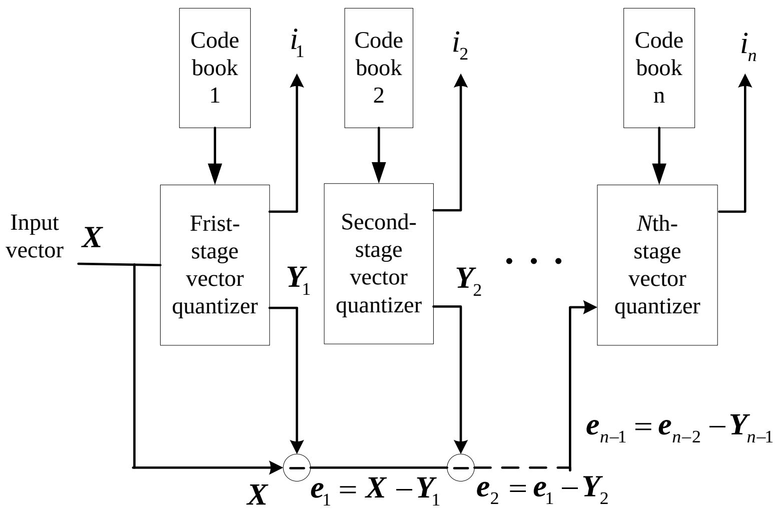


Figure 4. Multistage Vector Quantization System

On decoding side, after receiving all the codebooks and indices of each vector quantizer, the decoding result can be obtained by adding the entire reconstructed vectors together. Thus, the equivalent codebook can be generated from the Cartesian product. Compared to full-search VQ, MSVQ can reduce the data complexity from $N = N_1 \times N_2$ to $N_1 + N_2$. Moreover, though each stage of quantizer is connected in series, due to the reduction of codewords in each codebook, computing time of MSVQ is often less than full-search VQ.

3.2. DCT-MSVQ Scheme

Performing VQ in frequency domain is a feasible option to compress image. However, if full-time VQ is implemented directly, computing time and compression ratio can be both improved. In frequency domain, MSVQ can replace full-time VQ to get better performance. In addition, because the DC coefficients in low-sub-band represent the average brightness values of each blocks, these coefficients are often much larger and more correlative to their neighborhood. Differential pulse code modulation (DPCM) is introduced to perform coding task of DC coefficients after SQ is performed. As for the other AC coefficients of each block, high frequency AC coefficients are discarded at first, and then MSVQ is implemented after zigzag scanning.

LBG algorithm is the classic codebook-designing algorithm until now. However, its defects are also obvious. Because LBG algorithm excessively relies on its initial codebook, the quality of the final codebook may be damaged if the initial codebook is chosen poorly. Therefore, the performance of the whole system will be largely influenced. To solve this dilemma, the algorithm proposed in [11] is performed in the MSVQ process.

The main steps of the proposed image-encoding scheme are given below:

Step 1. Divide the input image into identically sized, non-overlapping $n \times n$ blocks, and then apply DCT to each block.

Step 2. To all the DCT coefficients, arrange them according to zigzag scanning order, then select first k coefficients, rest of the coefficients are discarded.

Step 3. Divide the vector into DC coefficient and a $k-1$ dimensional vector $\boldsymbol{X}$.

Step 4. Perform SQ and DPCM to DC coefficient.

Step 5. Initiate error vector $\boldsymbol{e}_0 = \boldsymbol{X}$.

Step 6. Use error vector $\boldsymbol{e}_n$ to train a common codebook $\boldsymbol{C}_n$ by using modified LBG algorithm proposed in [11].

Step 7. Using codebook $\boldsymbol{C}_n$ obtained from Step 6, perform VQ process then store index-table $\boldsymbol{Y}_n$ and error vector $\boldsymbol{e}_n$.

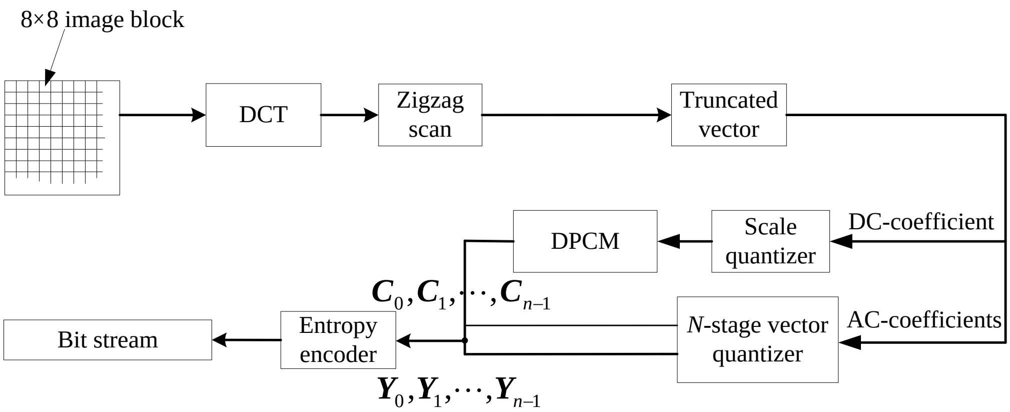
Step 8. If iterative time of VQ reaches N , which is iterative time required, go to Step 9; otherwise, return to Step 6 to perform next stage VQ.

Step 9. Export codebooks index-Tables $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_{n-1}$.

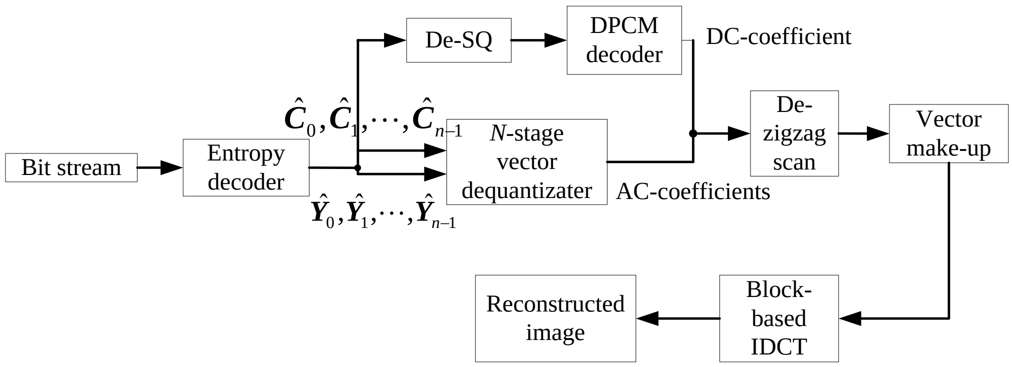
Step 10. Perform entropy-encoding process for codebooks $\boldsymbol{C}_0, \boldsymbol{C}_1, \cdots, \boldsymbol{C}_{n-1}$ index-tables $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_{n-1}$ obtain from Step 9, according to Huffman code rule. Transmit binary stream to the receiver.

On decoder side, the decoding process begins with the Huffman decoding, and it gives all the codebooks $\hat{\boldsymbol{C}}_0, \hat{\boldsymbol{C}}_1, \cdots, \hat{\boldsymbol{C}}_{n-1}$ and index-tables $\hat{\boldsymbol{Y}}_0, \hat{\boldsymbol{Y}}_1, \cdots, \hat{\boldsymbol{Y}}_{n-1}$ of MSVQ. All the error vectors can be reconstructed through the indices corresponding to the code words in codebooks. After that, reconstructed vector $\hat{x}$ can be obtained by summing word all the error vectors up. Then DC coefficient is combined to the reconstructed vector $\hat{x}$ with DPCM decoder then zigzag scan is implemented on the combined vector. After this process, zero-value is filled to all the other coefficients. Finally, IDCT is performed to get the final reconstructed image block data.

The block diagram of proposed algorithm is shown in Figure 5.



(a)



(b)

Figure 5. Scheme of Proposed Algorithm: (a) Encoding Process, (b) Decoding Process

4. Experimental Results

In this section, experimental results are presented to evaluate the performance of proposed DCT-MSVQ algorithm. Relative computations have been performed using MATLAB 2011 on the platform Core (TM) 2 i5 CPU 2.53GHz, 2048MB RAM, Microsoft Windows 7. The proposed technique is applied on a set of standard images in gray scale, which have identical size of 512×512 pixels.

Usually, peak signal to noise ratio (PSNR) in dB is used to measure the quality of the reconstructed image, which is calculated as follows:

$$\mathrm{PSNR} = 10\log_{10}\left[\frac{255^2}{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[I_{\mathrm{in}}(i,j)-I_{\mathrm{out}}(i,j)\right]^2}\right](\mathrm{dB}) \quad (3)$$

where I_{in} and I_{out} stand for the original image and the reconstructed image, respectively. M and N represent height and width of the test image.

Each test image is first divided into non-overlapped blocks of 8×8 size, DCT is applied to each blocks. Only 16 low frequency components are selected as a truncated vector. Each vector is divided into DC-coefficient and a 15-dimensional vector containing AC-coefficients. After that, SQ, which has a step-size of 16 and DPCM encoding, is performed on DC-coefficient, and a 2-Stage vector quantizer is implemented to encode the AC-coefficient part. In the process of MSVQ, all the truncated vectors and their error vectors are regarded as training vectors to train codebooks of each stage using PNN+LBG algorithm. Then, Huffman encoding is performed to generate bit stream for the index and DPCM code.

The performance of DCT+MSVQ method is compared with conventional VQ and DCT-VQ scheme on several test images. In conventional VQ and DCT-VQ, the size of blocks is also 8×8, and their codebooks' size is 512, identically. Two codebooks' sizes in MSVQ are 32 and 256, respectively. Reconstructed images of three schemes are shown in Figure 6, comparison of PSNR value and computing time are shown in Table 1 and Table 2.

Table 1. Performance Comparison of Conventional VQ, DCT-VQ and DCT+MSVQ (PSNR)

Test images	PSNR/dB		
	Conventional VQ	DCT-VQ	DCT+MSVQ
Lena	33.27	33.61	34.26
Peppers	32.53	32.61	32.90
Boat	30.15	30.44	30.67
Yacht	31.68	31.78	32.33
Zelda	33.65	33.83	34.70

Table 2. Performance Comparison of Conventional VQ, DCT-VQ and DCT+MSVQ (Computing Time)

Test images	Computing time/s		
	Conventional VQ	DCT-VQ	DCT+MSVQ
Lena	14.20	29.73	23.07
Peppers	16.05	22.17	20.72
Boat	17.91	22.80	19.37
Yacht	29.01	31.56	27.91
Zelda	30.49	39.13	33.96

(a) (b) (c)

Figure 6. Reconstructed Image Using DCT+MSVQ, Conventional VQ and DCT-VQ: (a) Proposed Algorithm, PSNR=34.26dB, (b) Conventional VQ, PSNR=33.27dB, (c) DCT-VQ, PSNR=33.61dB

The performance of the algorithm is also compared with performance of JPEG standard on Lena, which uses recommended quantization table and Huffman code table. Experimental result is shown in Table 3.

Table 3. Performance Comparison of DCT+MSVQ and JPEG on Lena

Bit rate/bpp	PSNR/dB	
	DCT+MSVQ	JPEG
0.49	34.26	34.60
0.47	33.77	34.12
0.40	33.58	33.74
0.38	33.23	33.70
0.34	33.22	32.90

Bit rate/bpp	PSNR/dB	
	DCT+MSVQ	JPEG
0.30	32.98	32.50

Figure 7 shows the original test image Lena, and Figure 8 shows the reconstructed image Lena using DCT+MSVQ scheme and JPEG, respectively, at bit rates of 0.32bpp and 0.40bpp.

Figure 7. Original Test Image Lena

Figure 8. Reconstructed Image Using DCT+MSVQ and JPEG: (a) Proposed Algorithm, Bit Rate=0.40bpp, PSNR=33.58dB, (b) JPEG, Bit Rate=0.40bpp, PSNR=33.74dB, (c) Proposed Algorithm, Bit Rate=0.30bpp, PSNR=32.98dB, (d) JPEG, Bit Rate=0.30bpp, PSNR=32.50dB

5. Conclusion

This paper proposes a new method of image compression code, using DCT, MSVQ and DPCM to meliorate the performance of conventional VQ algorithm and hybrid DCT-VQ technique. On aspects of compression ratio, consuming time and complexity, the proposed algorithm is superior to conventional VQ algorithm and hybrid DCT-VQ technique. Compared to JPEG scheme, the only complex operation in the algorithm is the codebook designing process, which is improved by using several small-sized codebooks instead of one codebook. As the experimental results show, the proposed algorithm has a better PSNR value compared to JPEG standard.

Acknowledgements

This work was supported by the Fundamental Research Funds of Shandong University (Grant No. 2014ZQXM008), the promotive research fund for excellent young and middle-aged scientists of Shandong Province, China (Grant No. BS2013DX022) and the National Natural Science Foundation of China (Grant No. 61201371). The authors would like to thank Xiaoyan Wang and Fanfan Yang for their kind help and valuable suggestions. The authors also thank the anonymous reviewers and the editor for their valuable comments to improve the presentation of the paper.

References

[1] ISO/IEC, "Information Technology—Digital Compression and Coding of Continuous-tone Still Images—Part 1: Requirements and Guidelines," ISO/IEC 10918-1 | ITU-T Rec. T.81 (First Edition), Feb. 17, **(1994)**.

[2] N. Ahmed, T. Natarajan and K. R. Rao, "Discrete cosine transform," IEEE Transactions on Computers, vol. 23 no. 1, **(1974)**, pp. 90-93.

[3] C. Loeffer, A. Ligtenberg and G. S. Moschytz, "Practical fast 1-D DCT algorithms with 11 multiplications," Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Glasgow, Scotland, vol. 2, May 23-26 **(1989)**, pp. 988-991.

[4] Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," IEEE Transactions on Communications, vol. 28 no. 1, **(1980)**, pp. 84-95.

[5] A. K. Pal, G. P. Biswas and S. Mukhopadhyay, "A hybrid DCT-VQ based approach for efficient compression of color images," Proceedings of the International Conference on Computer and Communication Technology, Allahabad, India, Sep. 17-19 **(2010)**, pp. 177-181.

[6] D. K. Mahapatra and U. R. Jena, "Partitional k-means clustering based hybrid DCT-vector quantization for image compression," Proceedings of the IEEE Conference on Information and Communication Technologies, Thuckalay, Tamil Nadu, India, April. 11-12 **(2013)**, pp. 1175-1179.

[7] A. K. AI-Asmari and A. S. Ahmed, "A low bit rate hybrid coding scheme for progressive image transmission," IEEE Transactions on Consumer Electronics, vol. 44 no. 1, **(1998)**, pp. 226-234.

[8] R. J. Liou and J. P. Wu, "Quadtree image compression using sub-band DCT features and Kohonen neural networks," Proceedings of the International Conference on Audio, Language and Image Processing, Shanghai, China, July. 7-9 **(2008)**, pp. 252-256.

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

[10] B. H. Juang and A. H. Gray, "Multiple stage vector quantization for speech coding," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Paris, France, vol. 7, May 3-5 **(1982)**, pp. 597-600.

[11] Y. X. Liang, J. W. Yang and Y. Li, "An improved PNN algorithm for designing LBG initial codebooks," Journal of Huazhong University of Science and Technology (Natural Science Edition), (in Chinese), vol. 38 no. 11, **(2010)**, pp. 25-28.

Authors

Xiao Zhou, she was born in Shandong province, China in 1982. She received her B.E. degree in automation from Nanjing University of Posts and Telecommunications, China, in 2003, her M.E. degree in information and communication engineering from Inha University, Korea in 2005, and her Ph.D. degree in information and communication engineering from Tsinghua University, China in 2013. Now she is a lecturer in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. Her current research interests include wireless communication technology, image processing and transmission technology.

Yunhao Bai, he was born in Shandong province, China in 1992. He received his B.E. degree in communication engineering from Shandong University, Weihai, China, in 2014. His research interests include image compression and transmission techniques.

Chengyou Wang, he was born in Shandong province, China in 1979. He received his B.E. degree in electronic information science and technology from Yantai University, China, in 2004 and his M.E. and Ph.D. degree in signal and information processing from Tianjin University, China, in 2007 and 2010 respectively. Now he is an associate professor in the School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, China. His current research interests include digital image and video processing, multidimensional signal and information processing, and smart grid technology.