Digital Video Compression Using Hybridized Technique: Discrete Cosine Transform and Vector Quantization

Awotunde, J.B.[1]*, Oladipo, I.D.[1], Tomori, A.R.[2], Abdulraheem, M.[2], Fatokun, O.M[3] & Fatoki, F.O.[2]
[1]Department of Computer Science, University of Ilorin, Ilorin, Nigeria.
[2]Directorate of Computer Services and Information Technology, University of Ilorin, Ilorin, Nigeria
[3]Department of Computer Science, Osun State University, Osogbo, Osun State, Nigeria
^**Correspondence:** jabonnetbylinks@gmail.com

ABSTRACT

The used of digital video application has progressively more useful and popular in mobile phones, mobile terminals, and personal digital assistants. But with limited bandwidth capacity and storage space, video data have to be compressed to reach its destination. Because video storing sequences and transmitting its raw data has become impracticable. The paper proposed a hybridized two video compression algorithms to compress video data (Discrete Cosine Transform and Vector Quantization). The main purpose of this is to efficiently utilize the advantages of both algorithms for better performance and to compute the compression rate on different video files using both algorithms. The two algorithms were chosen because both algorithms can be used to compress all types of video files. The result of the proposed system shows that the two used algorithms match each other in term of their weakness for better performance. The proposed methods have the potential of creating a better higher quality video stream at a maintained bit-rate without loss of video quality.

Keywords: Compression, Digital video, Correlation, Discrete cosine transform, Vector quantization

iSTEAMS Conference Proceedings Paper Citation Format
Awotunde, J.B., Oladipo, I.D., Tomori, A.R, Abdulraheem, M., Fatokun, O.M & Fatoki, F.O. (2018) Digital Video Compression Using Hybridized Technique: Discrete Cosine Transform and Vector Quantization. Proceedings of the 14th iSTEAMS International Multidisciplinary Conference, AlHikmah University, Ilorin, Nigeria, Vol. 14, Pp 59-66

1. INTRODUCTION

Digital video communication has been part of many application such as wires and wireless conversational services, broadcast services over satellite, terrestrial channels and digital video storage etc for many years. It formed as an integral part of the way we consume visual information, communication, and the creation of this communication (Abomhara et al., 2010; Muzhir & Talal 2011). It has become very difficult to store digital video on a storage device or to send it through the bandwidth of the transmission channel because they are finite in nature. Therefore, to store the full digital video, it has to be processed for easy transfer through the internet and to take less space on storage devices (Abomhara et al., 2010; Muzhir & Talal 2011). For example, to send about 167.96 Gigabytes digital video with a 720 x 480 pixel per frame and 30 frames per second that have a total 90 minutes full-color video, then the video has to be compressed for easy transfer far the internet and take less space on a storage device. Thus, several video compression algorithms had been developed to reduce the data quantity and provide the acceptable quality as possible as can. This paper starts with an explanation of the basic concepts of video compression algorithms and then introduces several video compression standards.

The compression of digital video become very important because of the correlation between one pixel and its neighbor pixels is very high, also the values of one pixel plus its adjacent pixels are very similar in nature. The correlation in a single frame of a video compression is called intraframe correlation and the correlation in the time direction is called interframe correlation. To reduce the storage quantity, the correlation between the pixels is reduced. A large number of still images is used to produced a video, and these images are taken at short time distance which makes their neighboring images to be similar. Therefore, there exists a high level of correlation between the frames or images in the time direction. If we can effectively reduce the interframe correlation, then video compression is successfully achieved.

Different methods have been proposed for video compression, and Moving Picture Experts Group (MPEG) is the most famous used technique, which is an ISO/TU standard for digital video compression. It similar to JPEG since it performs lossy compression for each frame by permanently remove pixels from the original images. Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are used by MPEG standard. The most difficult part in compressing color image lies in its intensity since the human eye is sensitive to variations in intensity. The video communication applications require bigger bandwidth to send data compared to audio communication. For example, a raw video of Quadrature Common Inter-Frame (QCIF) demands a bandwidth of 7 Mbps. But to meet the QoS requirement, the bandwidth specification can be reduced to as small as 5 Kbps (Axiotis & Xenikos, 2003, Ahmad et al., 2018) using any modern video compression codes designed for compress digital video contents for bandwidth narrowband systems/limited (Chang & Salim, 2003; Axiotis & Salkintzis, 2008; Ahmad et al., 2018) The paper used two compression algorithms to reduce the interframe and intraframe correlation of a digital video.

2. DISCRETE COSINE TRANSFORM (DCT)

The basis for almost all the image and video compression standard is Discrete Cosine Transform (DCT) based image coding. DCT is encountered very commonly in Digital Signal Processing and it is a derivative of Discrete Fourier Transform (DFT). The method transform space domain representation of an image to a spatial frequency domain known as DCT domain.

The equation of DCT is given below:

$$\mathrm{Y(k,l)} = \frac{\mathrm{C(k)C(l)}}{4} \sum\sum \mathrm{X_{ij}} \cos\left((2\mathrm{i}+1)\frac{\mathrm{k\pi}}{16}\right) \cos\left((2\mathrm{j}+1)\frac{\mathrm{k\pi}}{16}\right) \quad (1)$$

$$\boldsymbol{C(k)} = \left(\frac{1}{2}\right)^{\frac{1}{2}} \textbf{ if } \boldsymbol{k = 0} \quad (2)$$

This transformation can be viewed as the process to find each waveform, Y(k,l) represents its corresponding weight, and the scaled weight Y(k,l) yields the reconstructed version of the original 8 x 8 block from the sum of 64. DCT is commonly used since the information can be compressed to a very high degree and it energy compaction is among the highest next to the Karhunen-Loeve Transform. Also, DCT minimizes the blocking artifact present in many other transforms due to its favorable periodic in nature. Due to the finite word-lengths in a microprocessor, loss of information occurs due to rounding and truncating of calculated DCT values, but the loss of this information is irreversible. DCT is a lossless process which made it be a family of lossy compression.

DCT is useful in almost all standardized video coding algorithms. It is done on each 8 x 8 block (Xiang-Wei et al., 2008, 2009). The top left corner has the highest coefficients and the bottom right has the lowest using DCT (Ali, 1999). In a crisscross order from the top left to the bottom right, the coefficients are numbered to generate many small coefficients at the end. Its coefficients are then divided by the integer quantization value to reduce precision. At the end of the division, it is possible to lose the lower coefficients that are smaller than the quantization.

3. VECTOR QUANTIZATION

Vector quantization (VQ) is an algorithm used for compression of data that are already in digital form. It is a simplification of the scalar quantization to vector-valued inputs. The method has a well-designed mathematical explanation and is useful for coding of images, video, and speech. The quantization is performed at individual real-valued samples of the pixels waveform nearly in all coding schemes. For example, in transform methods, the transformed coefficients are scalar quantized prior to transmission. Better performance can be achievable coding with vectors quantization compare to scalars quantization even when the source is memoryless, Shannon's rate-distortion theory (Nasrabi & King, 1988).

A vector quantizer can be defined as a mapping Q of a k-dimensional Euclidean space onto a finite set of k-tuples Y.

$$Q : R^k \rightarrow Y \quad (3)$$

The technique can also be seen as an encoder/decoder pair that performs the functions Q and Q^{-1} respectively. Q(x) is the index in the codebook Y of the vector $\dot{x}$ that is closet (suitable distortion measure $d\big(x, Q(z)\big)$ is assumed) to x in some sense. The elements of Y are codewords or cded reproduction vectors. Also, Y is the cded reproduction alphabet. A vector quantizer is a subset of BK that is in practice applied to finite set S. VQ is referred to as Voronoi regions when it is used over set S, and it results in the partition of the set S into several disjoint subsets Si, It optimal is the one minimizes the expected distortion.

In the codebook, encoder needs to transmit only the index of the closest vector. The decoder that has an identical codebook uses the index to look up the appropriate vector. The codebook search and codebook design are the key elements of a good VQ scheme. Generalized Lloyd Algorithm (GLA) also called LBG algorithm is the popular vector quantizer design used to optimize codebook (Linde, Buzo & Gray, 1988). It's worth note that the decoding process could be as simple as table lookup, make VQ exceptionally pretty for low complexity decoder application.

4. SYSTEM IMPLEMENTATION OF VIDEO COMPRESSION APPROACHES

The implemented system use DCT and quantization on digital video, the different approach used for the compression technique are as follows:

Step 1: The digital video was filtered to generates de-nosing video
Step 2: Extraction of the frame to generates frames for a certain time.
Step 3: To generate effective frames, frame selection was performed
Step 4: The reordering of the frame to prepared frames
Step 5: Applying the DCT stage that is used to generates LL-band.
Step 6: Applying quantization to remove spatial frequencies content in an image
Step 7: Applying the VQ stage that is used to generates decoder
Step 8: Construct the final compressed video.

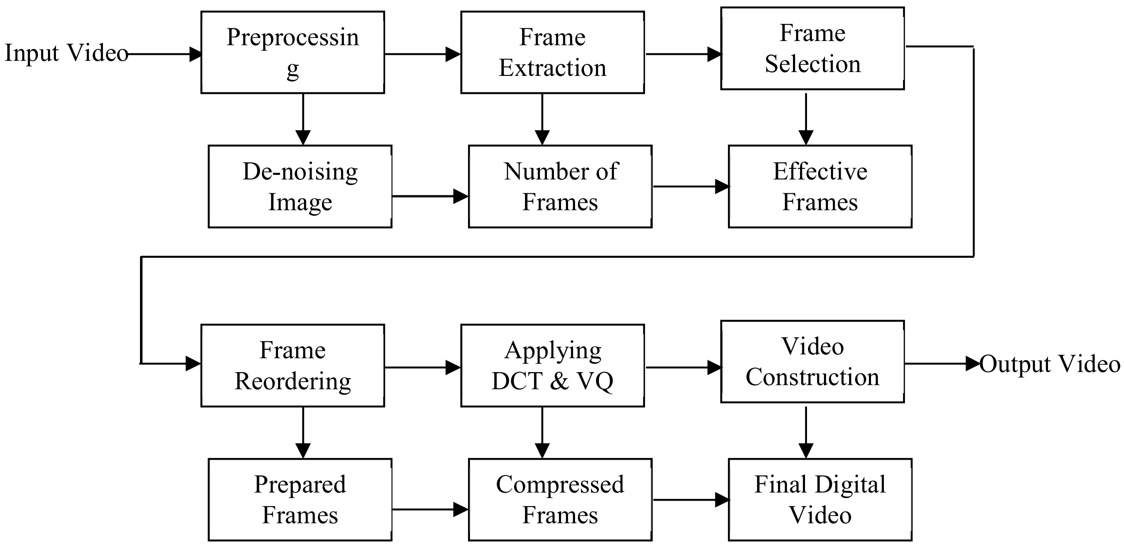

Figure 1: The Flowchart for the Proposed Video Compression

5. SYSTEM OVERVIEW

The paper defines a model by first take an input image and find out its color space of the image and work out its discrete cosine transform with its quantization, remain the image to be encoded and finally apply vector quantization algorithm and decode the file. The system was evaluated with other existing methods and using the PNSR method for the evaluation of the proposed system. The flowchart of the proposed system is shown in figure 2.

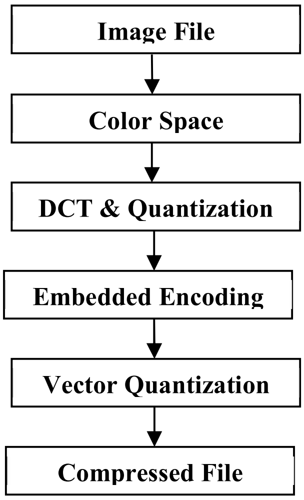

Figure 2: Flowchart of the Proposed System

5.1 Image Transformation

The step is to transform the image from spatial level representation into a different type of image using the transform method, then codes the transformed coefficients (values). The method is the best to provide a greater data compression when compared to predictive methods, only that in it expense of computational requirement.
Image compression using Discrete Cosine Transform

Through quantization, less important frequencies are discarded and important frequencies are used to retrieve the image decompression after the images are separated into parts of different frequencies by DCT. When DCT compared with another input dependent transforms the following advantages can be deduced: (i) DCT minimizes the block appearance called blocking artifact which results when boundaries between sub-images become visible (ii) It ability to pack most information in fewest coefficients and (iii) It can be implemented in a single integrated circuit (Boncelet, 1993; Deepak, Devansh, Kavya & Neha, 2014). 2D Discrete cosine transformation is as follows:

$$c(u,v) = D(u)D(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) cos\frac{(2x+1)u\pi}{2N}\ cos\frac{(2y+1)v\pi}{2N} \quad (4)$$

where u,v = 1,2,3, ... N-1

and

$$D(u), D(v) = \begin{cases} \sqrt{\frac{1}{N}}, u, v = 0 \\ \sqrt{\frac{2}{N}}, u, v = 1,2,3\ldots(N-1) \end{cases} \quad (5)$$

Its inverse transformation is given by:

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} D(u)D(v)c(u,v)cos\frac{(2x+1)u\pi}{2N}\ cos\frac{(2y+1)v\pi}{2N} \quad (6)$$

where u,v = 1,2,3, ... N-1

JPEG procedure

step 1: Digital image was divided into 4 by 4 or 8 x 8 blocks
step 2: DCT is designed to work on pixel values ranges from -128 to 127, though the pixel values of a block and white image range from 0-255. In order to use DCT on blocks, each block was modified to work in the range.
step 3: To calculate the DCT matrix equation 4 is used.
step 4: Equation 4 is applied to each block multiplying the modified block with DCT matrix on the left and transpose it matrix on the right.
step 5: Compressed each block through quantization
step 6: Entropy encoded by it quantized matrix
step 7: the Reverse process is used on the compressed image to reconstruct it.
step 8: Inverse DCT is used for decompression

5.2 Quantization

The compressed range of values to a single quantum value is quantization. The stream becomes more compressible when the number of discrete symbols in a given stream is reduced. Quantization matrix with a DCT coefficient matrix is used to carry out the transformation. Most of the compression takes place at quantization. Compression of the image is not always taking place at DCT since it is approximately lossless. The basic principle of quantization is the removal of spatial frequencies content in an image using the DCT operation followed the removal of the high-frequency content of the image. This operation is performed since the human eye is not responsive to the high-frequency content in an image taken, so removal of these spatial frequencies will not lead to any traceable loss in the quality of the image. There are recommended standard values of quantization table by JPEG standard used to deemphasize higher frequencies in the DCT image. Loss of information occurs during quantization, but the loss of this information is irretrievable.

5.3 Vector Quantization Algorithm

VQ undergo a structural decomposition, as a primary decomposition a vector quantizer is represented as:

$$Q(x) = \sum_{i=0}^{N} c_i S_i(x) \quad (7)$$

where the cell assignment function or selector function S_i defined as

$$S_i(x) = \begin{cases} 1 & if\ x \in R_i, \\ 0 & otherwise. \end{cases} \quad (8)$$

Without the restriction on the vector quantizer's partition geometry, further decomposition is not possible for the operation of the selector (Gersho, 1982). The restriction has to be applied to the partition geometry before further reduction is allowed.

To find the selector function easily the nearest neighbor condition vector have to be applied.

$$R_i = \{x : d(x,c_i) \le d(x,c_j)\}\ for\ all\ j \ne i. \quad (9)$$

To generate an output stream an encoder task is used, and the main tasks of an encoder are
Step 1: Read in a frame in the given format
Step 2: Used vector quantization to compress the frame
(a) Retrieve a codebook or generate it.
(b) Encode the generated frame
(c) Compress the bit-packing (indices)
Step 3: Write out the compressed frame in the appropriate format
Step 4: if the frames are left, then go to step 1.

5.4 Vector Quantizer Algorithm design (The Generalized Lloyd Algorithm (LBG Algorithm)

Step 1: Initialization: Given the codebook size N, a distortion threshold $\in \geq 0$, a distortion measure, an initial codebook C_0, and a training $\{i_j\}$, set an iteration index m = 1 and an initial distortion $D_0 = \infty$
Step 2: Partition the training sequence into region R_i, $1 \leq i \leq N$, using the nearest-neighbour condition.
Step 3: Computer the average distortion D_m.
If $\frac{D_{m-1} - D_m}{D_m} \leq \in then$
C_m is the final codebook
stop
else continue
Step 4: Find $C_{m+i} = \{\dot{C}_i\} where\ \dot{C}_i = Cent(R_i)$ is the centroid of R_i

$$m \leftarrow m + 1$$

Goto 2

6. RESULTS AND DISCUSSION

The paper tries to compress a digital video using a hybridize algorithms DCT and VQ the purpose is for the two algorithms to enhance each other for better performance. The input image is divided into 8 x 8 blocks, and 2D DCT is used to compute each block. The coefficients are quantized, coded and transmitted, JPEG file reader decodes the quantized DCT coefficients and computes the inverse of the 2D DCT of each block, and puts the blocks back together to form a single image. DCT coefficients have values close to zero, which means discarding them will not affect the quality of the reconstructed image. Then VQ reworks on the reconstructed image to get a better result digital image. The residual signal can be compressed by VQ, The 2D signal is divided into blocks of equal size and VQ encoder is used to compress the residual signal to it nearest-neighbor encoder It reference lookup table contains the centroids of the VQ partition. Each block is compared by the encoder in some predefined scanning order with each member of the lookup table to find the closest match in terms of the defined distortion measure which the mean squared error. The closest index matching code vector in the table is transmitted as the compressed representation of the corresponding vector.

To evaluate the performance of the algorithm, a measure is defined to compare the original video and the compressed video. Video compression systems are designed to minimize the mean square error (MSE) between two video sequences $\varphi 1$ and $\varphi 2$ is defined as

$$\mathrm{MSE} = \sigma_c^2 = \frac{1}{N} \sum_I \sum_{x,y} [\varphi_1(x,y,t) - \varphi_2(x,y,t)] \quad (10)$$

N is the total number of frames in either video sequences.

The peak-signal-to-noise ratio (PSNR) in decibel (dB) is most used as a quality measure in video coding is defined as:

$$\mathrm{PSNR} = 20 \log 10 \frac{255}{\mathrm{MSE}} \quad (11)$$

For the purpose of this paper, PSNR used to evaluate the proposed system. Table 1 shows the result of the performance of the proposed system. The proposed preprocessing scheme was tested on two different sequences for two level of the hybridize decomposition. PSNR and compression ratio was used for the quality of reconstructed frame and the performance of the proposed method. The used test approach was based on determining the effects of the involved quantized parameters on compression ratio (CR) and fidelity measure PSNR.

Table 1: Results for APSNR and ACR.

No. of level = 2 Frame# = 8		
A	**Average PSNR**	**Average Compression Ratio**
0.1	32	11.25
0.2	35	15.61
0.3	33	21
0.4	30	23.60
0.5	28	31
0.6	26	42.8
0.7	16	81
0.8	16	78

Table 1 shows the test results applied on Image of Mona Lisa sequence. The adopted fixed parameter value are Q_LL = 20, Q_HL, Q_LH, Q_HH = 120.

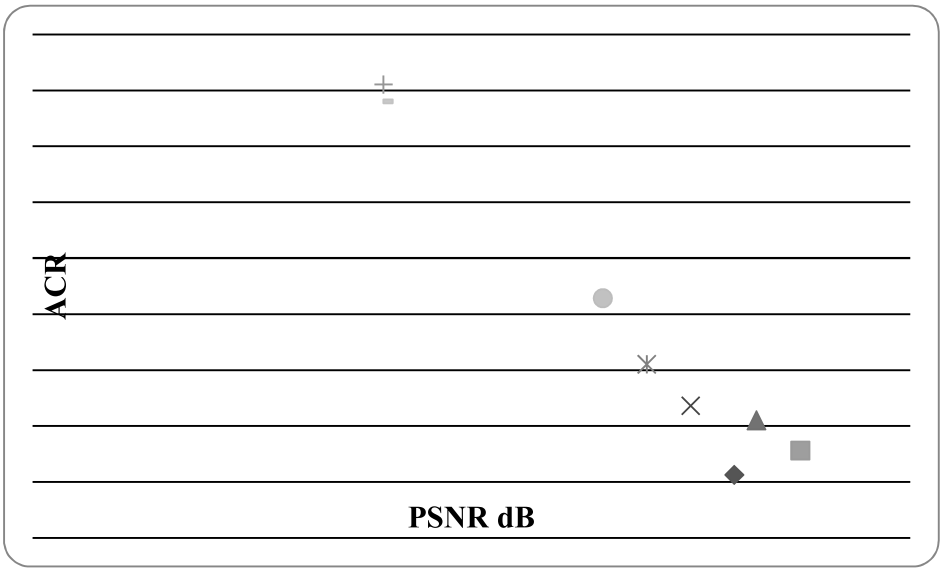

Figure 1: PSNR dB against ACR at 2nd Level

Figure 1 shows the relationship between the values of APSNR and ACR.

7. CONCLUSION

This paper introduced a hybridize DCT and QV concepts of video compression and used the characteristics of various video compression standards. The existing video compression can compress the video effectively, but there is still room for improvement. For instance, to reduced the temporal and spatial redundancy, DCT is exploited. Since DCT can reduce the temporal and spatial redundancy very well, but DCT can lead to the annoying unnecessary block artifact. VQ is therefore exploited to smooth the block object and to recovers the video as possible it can. The performance of the hybridize algorithms performed better but there is still room for improvement. The performance of the proposed technique is based on compression ratio and fidelity measure, and the tested results show that the quality of reconstructed frames is better in the second level.

REFERENCE

1. Abomhara M., Khalifa O.O., Zakaria O., Zaidan A.A., Zaidan B.B., and Rame A. (2010). Video Compression Techniques: An Overview, *Journal of Applied Sciences*, 10(16), 1834-1840.
2. Ali, S.T. (1999). Video Compression: MPEG-4 and beyond. http://www.cse.wustl.edu/-jain/cis788-99/compression/.
3. Xiang-Wei, L., Zhan-Mang, Z., Ming-Xin, Z., Zhe, W., and Guo-Quan, Z. (2008). A novel rough sets based key frame extraction algorithm in compressed-domain. *Information Technology Journal*, 7, 1043-1048.
4. Xiang-Wei, L., Zhan-Mang, Z., Ming-Xin, Z., Ya-Lln, Z., and Wei-Yi, W. (2009). Rapid shot boundary detection algorithm based on rough sets in video compressed domain. *Research Journal of Information Technology*, 1, 70-78.
5. Axiotis, D.I., and Xenikos, D.G. (2003) On the performance of TETRA short data service-transport layer, *Wireless personal communications*, Vol. 43 No. 4, December 2007, 1121-1135.
6. Chang, Y.C. and Salim B. M., (2003) MPEG-4 video error resilient tools performance evaluation and assessment, *In Communications, Computers and Signal Processing, PACRIM, 2003 IEEE Pacific Rim Conference on IEEE,* 70-72.
7. Axiotis D.I. and Salkintzis, A.K. (2008) Impact of voice traffic on the performance of packet data transmission in TETRA networks, *IEEE Transactions on Vehicular Technology,* 57(5), 3040-3052.
8. Ahmad, R.W., Rafi us, S., Atta ur, R.K., Kashif, .B., Junaid, .S., Shahzad, .S., Torki A.A., Sajjad, A.M. (2018) An Investigation of Video Communication over Bandwidth Limited Public Safety Network. *Malaysian Journal of Computer Science*. 31(2), 85-107
9. Muzhir S.A. and Talal A.H. (2011), Video Compression Algorithm Based on Frame Difference Approaches, *International Journal on Soft Computing*, 2(4), 67-79
10. Nasrabi, M. N. and King, R. A. (1988) Image coding using vector quantization: A Review, *IEEE Transaction Communication*, COM-36, 957-971
11. Linde A., Buzo Y. and Gray R. (1980) An algorithm for a vector quantization design, *IEEE Transaction Communication*, COM-28, 84-95
12. Boncelet, C. G. (1993) Rlock Arithmetic Coding for Source Compression, *IEEE Transactions on Information Theory*, IT-39, 1546-1554
13. Deepak K.J., Devansh G., Kavya G., Neha J. (2014) Image Compression using Discrete Cosine Transform and Adaptive Huffman Coding, *International Journal of Emerging Trends & Technology in Computer Science*, 3(1), 90-94
14. Gersho A., (1982) On the structure of vector quantizers, *IEEE Transaction on Information Theory*, 28, 157-166