Video Coding Technique using MPEG Compression Standards

A. J. Falade* and F. J. Olanrewaju

Department of Electrical & Electronics Engineering, University of Ilorin, Ilorin, Nigeria

ABSTRACT: Digital video compression technologies have become part of life, in the way visual information is created, communicated and consumed. Some application areas of video compression focused on the problem of optimizing storage space and transmission bandwidth (BW). The two dimensional discrete cosine transform (2-D DCT) is an integral part of video and image compression, which is used in Moving Picture Expert Group (MPEG) encoding standards. Thus, several video compression algorithms had been developed to reduce the data quantity and provide the acceptable quality standard. In the proposed study, the Matlab Simulink Model (MSM) has been used for video coding/compression. The approach is more modern and reduces error resilience image distortion.

KEYWORDS: Compression, Simulink, Video and coding

[Received December 4 2011; Revised November 7 2013&November23 2013; Accepted November 28 2013]

I. Introduction

Video coding is a key technology for all areas of multimedia communication and storage. Technologies such as mobile video, mobile TV, DTV, DVD players, digital cameras, video telephony and multimedia messaging have all been enabled through the availability of efficient compression algorithms. Photographs, printed text, and other hard copy media are now routinely converted into digital form, and the direct acquisition of digital images is becoming more common as sensors and associated electronics of very high quality are becoming available now (Sivanandam et al., 2009 ;Choi and Park, 1989). MPEG-1 systems and MPEG-2 video have been developed collaboratively with the International Telecommunications Union- (ITU-T). The DVB selected MPEG-2 added specifications for modulation over satellite, cable, and terrestrial channels. The Advanced Television System Committee selected parts 1 and 2 of MPEG-2 for terrestrial broadcasting of digital HDTV also added its modulation for terrestrial broadcasting (Sivanandam et al., 2009).

Dumitra and Haskell (1990) proposed a content-based video coding method by texture replacement. The limitations are due to partially unreliable texture detection. Ndjiki-Nya et al. 2012 and Huffman, 1951, have shown that the mean squared error (MSE) criterion, typically used in hybrid video codecs like H.264/AVC may not be an adequate coding distortion measure for high frequency regions displayed with limited spatial resolution. Girod Bernd et al. (1996) based their research work on scalable video. Scalable video codecs offer substantial new features which address the special demand of today's multimedia communication systems, like computational limited decoding and video transmissions over heterogeneous packet networks or wireless channels. More promising approaches can be built upon spatio-temporal resolution pyramids, proposed by Uz et al. (2004). Open-loop as well as closed-loop pyramid Scalability can also be utilized for transmission over heterogeneous packet networks since it offers an easy way to reduce the bit rate of transmitted video data in case of congestion. Coders can be used for their efficient encoding and multi-scale motion compensation can be easily included. However, the emerging MPEG-4 standard is supposed to provide contents based functionalities in addition to efficient video coding (Girod Bernd et al. 1996). A solution for the optimum trade-off by applying rate-distortion theory has been presented recently for region-based motion compensation. The regions are optimized with respect to a Lagrangian cost function by variation of the region contours (Turag et al., 2005).

Compression is needed to simply reduce the amount of space that video would otherwise take to store. There are many factors to consider when choosing a compression technique. These include real-time/ non-real-time, symmetrical/ asymmetrical, compression ratio, lossless/ lossy compression, inter-frame intra-frame. Real-time/ non-real-time refers to capturing, compressing, decompressing and playing back all in real time with no delays. Symmetrical implies capturing, storing, and playback at the same rate, while asymmetrical uses more time to compress and hence may have an advantage for playback speed. Inter-frame/ intra-frame is a more powerful method which uses a predictivetechnique.

The method employed was majorly based on software analysis which was achieved by the use of MATLAB. The MPEG standard is designed to be generic. This work is able to compress a video data and compare the distortion in each frame of the original video and the compressed video. A very promising technique for achieving very low data rates in video coding is model-based coding. In a model-based system the coder analyses the scene using three dimensional models of

*Corresponding author’s e-mail address:falade.alaba@yahoo.com

the objects. Yun Q.Shi and Huifang Sun, (2008) proposed a rate control algorithm for low data rate video.

II. Video Coding Analysis

In video systems other than television, non-interlaced video or 'progressively scanned' or 'sequentially scanned' video is common and all the lines of a frame are sampled at the same instant in time. The red, green and blue (RGB) signals coming from a colour television camera can be equivalently expressed as luminance (Y) and chrominance (UV) components. The chrominance bandwidth may be reduced relative to the luminance without significantly affecting the picture quality. For standard definition video, CCIR recommendation 601,defines how the component (YUV) video signals can be sampled and digitized to form discrete pixels. The active region of a digital television frame is 720 pixels by 576 lines for a frame rate of 25 Hz. Using 8 bits for each Y, U or V pixel, the uncompressed bit rates for 4:2:2 and 4:2:0 signals are therefore:

4:2:2:720 x 576 x 25 x 8 + 360 x 576 x 25 x (8 + 8) = 166 Mbit/s
4:2:0: 720 x 576 x 25 x 8 + 360 x 288 x 25 x (8 + 8) = 124 Mbit/s

MPEG is capable of compressing the bit rate of standard-definition 4:2:0 video down to about 3-15 Mbit/s. For digital terrestrial television broadcasting of standard-definition video, a bit rate of approximately 6 Mbit/s is thought to be a good compromise between picture quality and transmission bandwidth efficiency. In MPEG-2, three 'picture types' are defined they are; intra pictures, predictive pictures and bi-directionally predictive. The picture type with prediction modes may be used to code each block. Many compression techniques have been developed in the last decades. They can differ on the way the information is processed, the algorithms chosen or the coding methods used, but they shared a general block-diagram. This is constructed as shown in Figure1, which represent the idea of most video encoders.

III. Spatial Redundancy

When analysing any series of consecutive frames from a video sequence. The image information's differ from one frame to the other and is reduced to some small changes due to motion or illumination alterations in the content. Although the temporal correlation is removed, the spatial redundancy still exists. For analysis of 2-D signals such as images, we require a 2-D version of the two dimensional discrete cosine transform (DCT). For an $N \times N$ matrix, the 2-D DCT is computed in a simple way, i.e. the 1-D DCT is applied to each row of the matrix and then to the column of the result. Therefore, MPEG exploits 2-D DCT relation, given below to remove the spatial correlation.

$$F(u,v) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (1)$$

For $u = 0,.....,N-1$ and $v = 0,.....,N-1$ where $N = 8$ and $C(k) = 1/\sqrt{2}$ for $k = 0$ or 1 otherwise

-

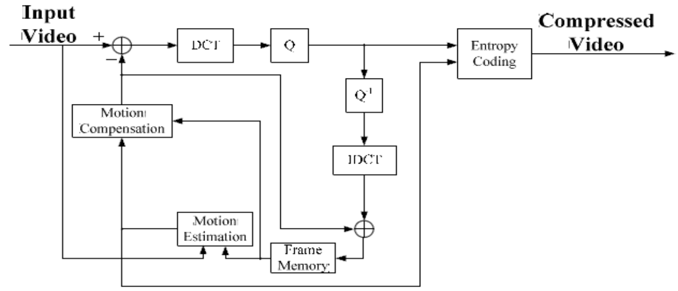

Figure 1: Generalized encoder structure of MPEG

The function of the coder is to transmit the DCT block to the decoder, in a bit rate efficient manner, so that it can perform the inverse transform to reconstruct the image. Quantization is used to reduce the number of possible values to be transmitted, reducing the required number of bits. Quantization noise introduced by the coder is not reversible in the decoder; hence this makes coding and decoding process 'lossy'. The references between the different types of frames are realized by a process called motion estimation or motion compensation. The correlation between two frames in terms of motion is represented by a motion vector. The resulting frame correlation, and therefore the pixel arithmetic difference, strongly depends on how good the motion estimation algorithm is implemented.

IV. MSM for Video Compression

MSM is a guide software used in simulating signals in more complicated designs. Figs 1 to 5 are the blocks used to achieve the compression. The blocks get the media file in any format, in this case, used an AVI (Audio Video Interleave) video format. The video is being loaded from the file, the loaded video file is being fed into a splitter, the function of the splitter is to select or reorder the input video, which is a multidimensional signal, and it output the video in RGB (Red Green Blue) format. The RGB is fed to the color space conversion block, which converts the RGB into Y (Luminance) and UV (Chrominance).

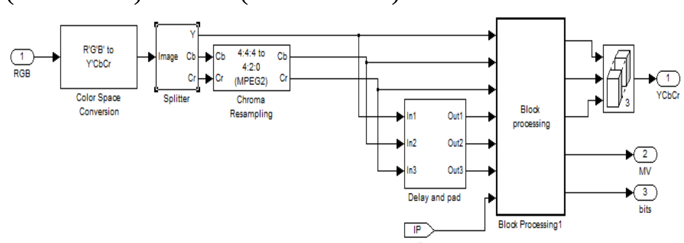

Figure 2: Block Process for the Video Compression (Encoder Block)

The spited video is being fed into another splitter which split the image into Y (Luminance) and Cb& Cr (Crominance). These conversions from RGB to YCbCr specify the colour spacing between the frames. The chroma resembling is to reduce the space required for the storage of the video file, and also to reduce bandwidth consumption, follow by the delay and padding of the splited image/video in the next block. The process block extracts the submatrices which has being specified and sends each submatrix to another subsystem for processing. The splited images/video was reassembled and the block is given in Figure 3. The decoder processes is similar to the encoder block but in an opposite direction. In the sense that in the decoder block, the frames of the video is being assembled back, the assembled frames results to the compressed video. The motion compensation block is shown in Figure4, the splited frame.

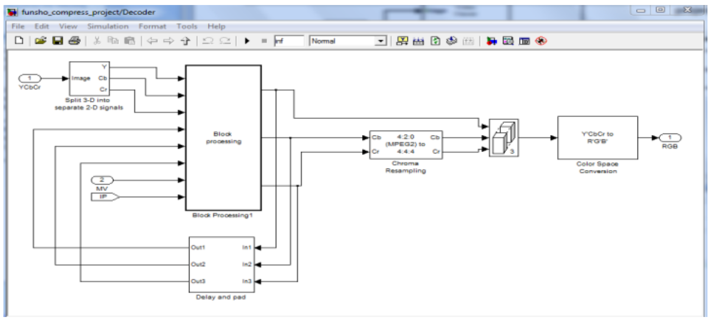

Figure 3: Block Process for the Video Compression (Decoder Block)

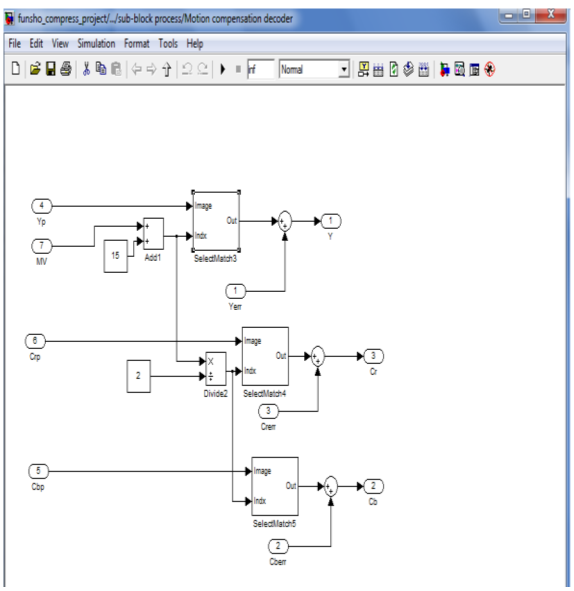

Figure 4: Motion Compensation/Estimation Block.

V. DISCUSSIONS

Bit-stream of MPEG encoder is sequence information, group of pictures information, picture information, slice information and macroblock information.Figure 6 gives the breakdown of the MPEG video bitstream, hence here are some tested data that were obtained in the compression of a video data.

Figure 6 shows the finished simulink model for the video compression. The blocks comprises of all the steps that were explained above. Some of the frames, extracted using matlab GUI (graphical user interface) integrated development environment, are displayed below. Analyzing the above

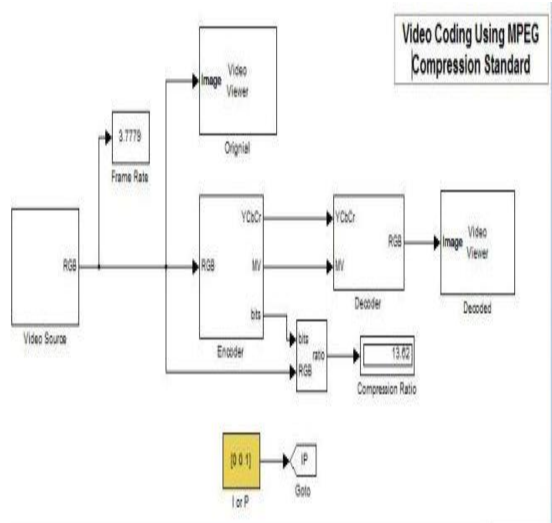

Figure 5: Final Simulink Model for the Video Compression

frames, it was observed that there was no much distortion in the original frame of the video and its corresponding compressed frame. Although all compression was done using matlabsimulink, the codes will display the frames and also does the compression. Observing Figure 6, it is clearly noted that frame 50 of the compressed data is equivalent to the 100th frame of the original data, hence it can be clearly stated that some part of the video has been chopped off during the compression which result in reduction in the quality of the data and its corresponding size.

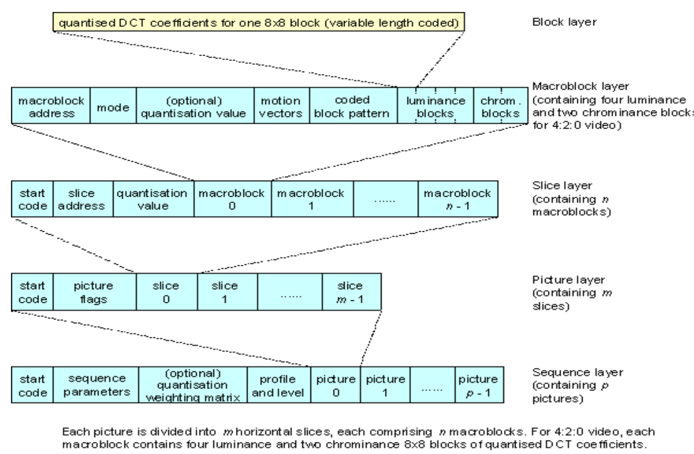

Figure 6: Encoder Bit-stream of MPEG

VI. CONCLUSION

In this work , the coding/compression of the video was done using MSM which gives a more modern approach to video compression with a less error resilience image distortion. MPEG compression which consists of multiple conversion and image/video compression algorithms using discrete cosine and MSM were developed. This method competes favourably with the compression standards, to achieve an optimal compression with best possible quality.

However, the picture quality obtained through an MPEG codec depends strongly upon the

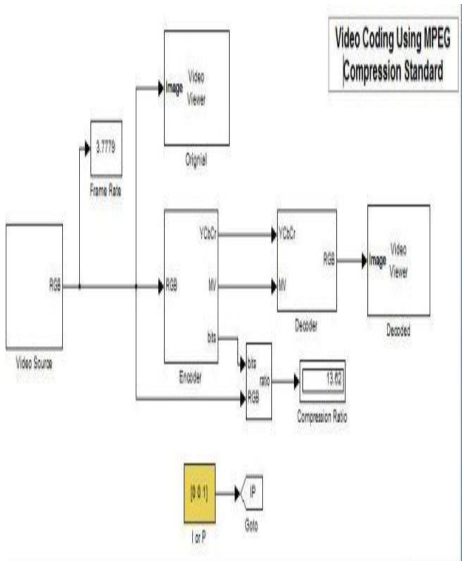

Figure 7: MSM for the Compression

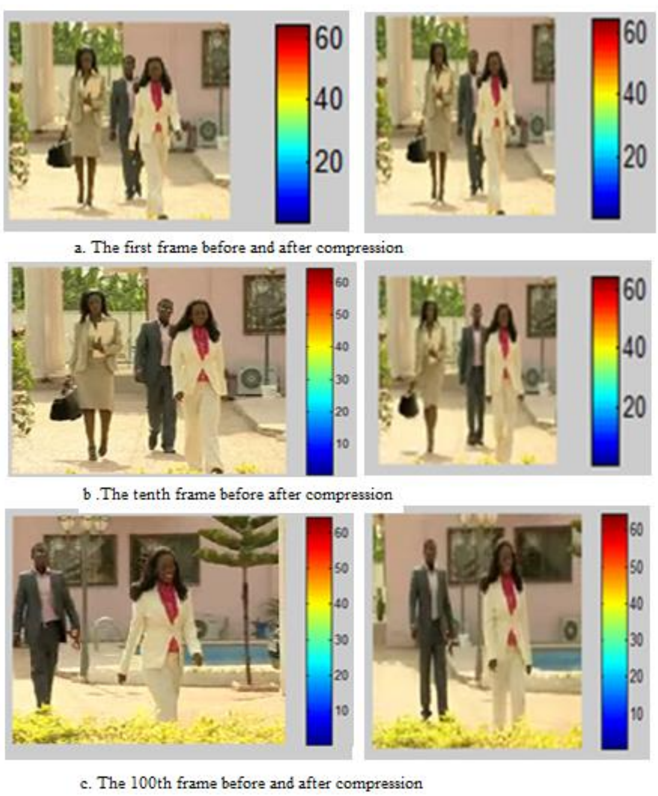

Figure 8: Sample image before and after compression

picture content.

REFERENCES

Brunco, D., Villiari, M., Sata, A. and Pulialito (2003). A. QoS management for MPEG-4 Flows in wireless environment. J. of microprocessors and Microsystems 27: 85-92.

CCIR 601-The origins of the 4:2:2 DTV Standard. Available online at: https:/tech.ebu-ch/docs/techreview/trev_304_rec601. Accessed on Feb. 20, 2011.

Chiariglione, L. (1996). MPEG and multimedia communications. Available online at: [http://www.cselt.it/ufv/leonardo/paper/isce96.html]. Accessed on Feb 7, 2011.

Choi, W. Y. and Park, R. H. (1989).Motion vector coding with conditional transmission.Signal Processing, 18 : 259-267.

Dumitra and Haskell (1990).Content-Based Video Coding Method by Texture Replacement.vailable online at: @ieeexpolre.iee.org/iel5/4284552/4284553. Accessed on Feb. 7, 2011.

Girod B., K. Ben Younes, R. Bernstein, P. Eisert, N. Farber et al. (1996). Recent advances in Video Compression. Available online at: www.linkedin. Com/pub/khaled-ben-younes/26/54/385.Doi=10.1.56.4665. Accessed on March 8, 2011.

Huffman, D. A. (1951).A method for the construction of minimum redundancy codes. In the Proceedings of the Institute of Radio Engineers 40 : 1098-110.

Ndjiki P., D. Doshkov, H. Kaprykowsky, F. Zhang, D. Bull and T. Wiegand (2012). Perception-oriented video coding based on image analysis and completion: A review. Signal proc. Image comm., 27 (6) : 579-594.

Netravali, A. N. and Limb J.O (1980). Picture coding: A review, Proc. IEEE, Trans. On Acoustics, speech and signals, 68 (3): 7-12.

Priya T.V., Srvaya M., Bhavana B., Pravallika K., Gayathri K. (2013). Scalability Techniques of MPEG-2 Standard for Video Compression.Vol. 3, Issue1, ISSN:2277 128X, Available online at: www.Ijarcsse.com. Accessed on Nov. 2, 2013.

Richardson, I.E. (2003). H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia. John Wiley & Sons Ltd.ISBN : 0-470-84837-5.

Sivanandam, S.N., A.P. Pandian and P. Rani. (2009). Lossy still image compression standards: JPEG and JIPEG2000. Int. J. the computer, the internet and management, 17 (2): 69-84.

Turag D.S., Schaar M. V. and Pesquet-popescu (2005).Complexity Scalable motion Compensated wavelet video encoding. IEEE Trans. on Circuits and Systems for Video Tech., 15 (8): 982-993.

Uz K. M., Vetterli, M, and Le Gall D. J. (2004). Interpolative multiresolution coding of advanced television with compatible subchannels. IEEE Trans. Circ. Syst. 283-292.

Vetterli M. and Herley C. (1992). Wavelets and Filter Banks : Theory and Design'', IEEE Trans. on signal processing, 40 (9): 2207-2232.

Yun Q.Shi and Huifang Sun, (2008).Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standard.Taylor& Francis group. 2nd Edition.

ZeNian Li and Drew M.S., (2004). Fundamental of multimedia. Pearson Education, Inc., Upper Saddle River, NJ, 07458.