Performance Evaluation of H.264 AVC Using CABAC Entropy Coding For Image Compression

Mr.P.S.Jagadeesh Kumar
Associate Professor,
Department of Electrical Engineering,
DIT University, Dehradun, Uttarakhand
E-mail: jagadeesh_dit@yahoo.com

Dr.Gagan Singh
Professor & Head,
Department of Electrical Engineering,
DIT University, Dehradun, Uttarakhand
E-mail: gaganus@gmail.com
www.dit.edu.in

Abstract

The proposed paper aims in evaluating the performance of H.264/MPEG-4 part 10, "Advanced Video Coding" standard using CABAC entropy coding for compression of different types of images. The performance evaluation is based on the metrics such as compression ratio and PSNR at different bits per pixel (Bpp). The different types of images used to test the performance of the proposed system are compound image, text image, natural image and computer generated image. The proposed system has been simulated in matlab simulink environment. It is observed that the H.264AVC using CABAC entropy coding is very efficient in compressing compound images with higher compression ratio and PSNR compared to other types of images at high bitrates.

Keywords: H.264/AVC, CABAC, Compound image, Daubechies Wavelet Transform.

I. INTRODUCTION

The latest video compression standard, *H.264 (also known as MPEG-4 Part 10/AVC for Advanced Video Coding)*, is expected to become the video standard of choice in the coming years. H.264 is an open, licensed standard that supports the most efficient video compression techniques available today. Without compromising image quality, an H.264 encoder can reduce the size of a digital video file by more than 80% compared with the Motion JPEG format and as much as 50% more than with the MPEG-4 Part 2 standard. This means that much less network bandwidth and storage space are required for a video file. Or seen another way, much higher video quality can be achieved for a given bit rate. Jointly defined by standardization organizations in the telecommunications and IT industries, H.264 is expected to be more widely adopted than previous standards. H.264 has already been introduced in new electronic gadget such as mobile phones and digital video players, and has gained fast acceptance by end

© 2013. The authors - Published by Atlantis Press

users. H.264 also has the flexibility to support a wide variety of applications with very different bit rate requirements. H.264/ MPEG-4 AVC can be used in compressing images [1]. Many researches pertaining to image compression have been noticed but the need for compound image compression patrols the need of the hour. *Compound images* occupy larger size compared to natural images and have images, text and graphics in them. Thus it provides a greater challenge to compress them effectively. Moreover H.264 is a well known video coding standard and for the first time it's being used to compress compound images by the use of context adaptive base arithmetic coding (CABAC). There is the memory constraint of storage of the compound images which has given the idea for compressing them and using it for later use.

II. WHAT IS CABAC?

The Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG are finalizing a new standard for the coding (compression) of natural video images. The new standard will be known as H.264 and also MPEG-4 Part 10, "AVC". The standard specifies two types of entropy coding: *Context-based Adaptive Binary Arithmetic Coding* (CABAC) and Variable-Length Coding (VLC). Context adaptive binary arithmetic coding is a form of entropy coding used in H.264/MPEG4 AVC video encoding. It is notable for providing much better compression than most other encoding algorithms and is one of the primary advantages of the H.264 AVC encoding scheme [2]. It is a lossless compression technique. The design of CABAC is in the spirit of our work. To circumvent the drawbacks of the known entropy coding schemes for image compression, we combine an adaptive binary arithmetic coding technique with a well-designed set of context models. Guided by the principle of alphabet reduction, an additional binarization stage is employed for all non-binary valued symbols.

III. CABAC FRAMEWORK

The CABAC encoding process consists of at most, four elementary steps:

1. Binarization: CABAC uses Binary Arithmetic Coding which means that only binary decisions (1 or 0) are encoded. A non-binary-valued symbol (e.g. a transform coefficient or motion vector) is "binarized" or converted into a binary code prior to arithmetic coding. This process is similar to the process of converting a data symbol into a variable length code but the binary code is further encoded (by the arithmetic coder) prior to transmission. Stages 2, 3 and 4 are repeated for each bit (or "bin") of the binarized symbol.

2. Context model selection: A "context model" is a probability model for one or more bins of the binarized symbol. This model may be chosen from a selection of available models depending on the statistics of recently-coded data symbols. The context model stores the probability of each bin being "1" or "0".

3. Arithmetic encoding: An arithmetic coder encodes each bin according to the selected probability model. Note that there are just two sub-ranges for each bin (corresponding to "0" and "1").

4. Probability update: The selected context model is updated based on the actual coded value (e.g. if the bin value was "1", the frequency count of "1"s is increased). The data flow diagram of the proposed system is shown in Fig.1.

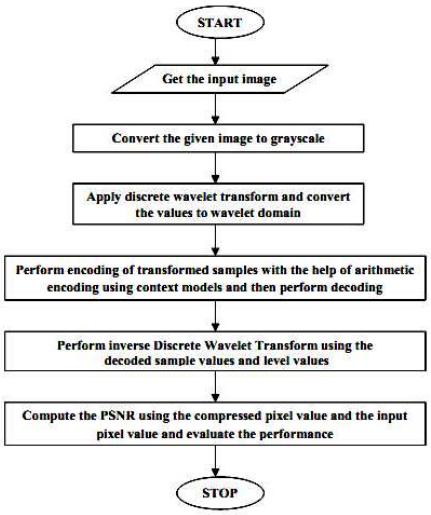


Fig.1. Data flow diagram

IV. IMAGE COMPRESSION SCHEME

The input image is first converted to grayscale and processed. Discrete Wavelet Transform is applied to the grayscale image and the values are transformed to wavelet domain. After transform, Context Adaptive Base Arithmetic Coding (CABAC) is applied and it is encoded and decoded suitably. The decoded pixel array values are now compared with the input pixel array values and then the difference is computed as the Mean Scalar Error

(MSE) and the peak signal to noise ratio (PSNR) is evaluated for various bits per pixel values. Finally, based on the compression ratio obtained we evaluate the performance and conclude on the efficiency of this scheme. The entire block diagram of the image compression scheme is shown in Fig.2. Programming and developing the proposed compression scheme is faster with MATLAB than with traditional languages because MATLAB supports interactive development without the need to perform low-level administrative tasks, such as declaring variables and allocating memory.

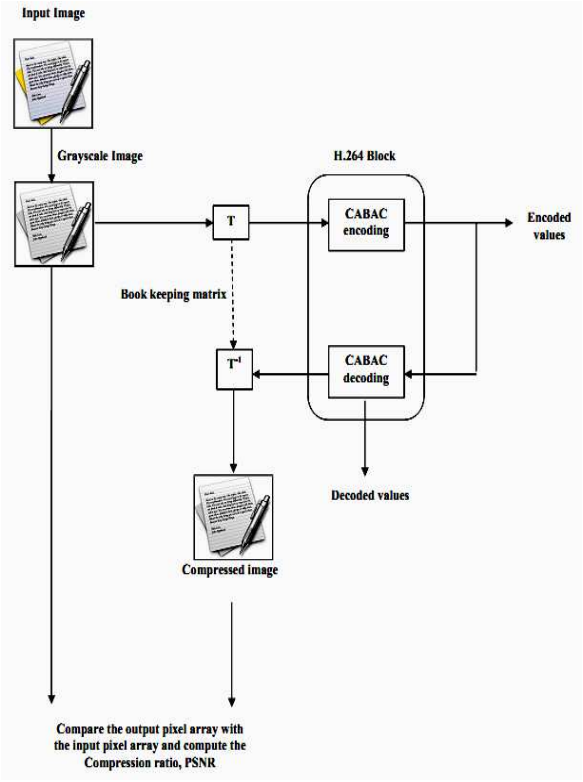


Fig.2. Image compression scheme

V. DISCRETE WAVELET TRANSFORM

The Discrete wavelet transform is a powerful tool for processing signals. *Daubechies wavelet* is a wavelet used to convolve image data. The wavelets can be orthogonal, when the scaling functions have the same number of coefficients as the wavelet functions, or biorthogonal, when the number of coefficients differs. The JPEG 2000 compression standard uses the biorthogonal Daubechies 5/3 wavelet (also called the LeGall 5/3 wavelet) for lossless compression and the Daubechies 9/7 (also known as the Cohen-Daubechies-Fauraue 9/7 or the "CDF 9/7") for lossy compression [3]. Daubechies 9/7 is used in our proposed compression scheme. In general Daubechies wavelet has extremal phase and highest number of vanishing moments for defined support width. The wavelet is also easy to put into practice with minimum-phase filters. This wavelet is often also called the CDF 9/7 wavelet (where 9 and 7 denote the number of filter tap s). There are several ways wavelet transforms can decompose a signal into various sub bands. These include uniform, octave-band and adaptive or wavelet packet decomposition. Out of these, octave-band decomposition is the most widely used [4]. The decomposition of the signal into different frequency bands is simply obtained by successive high pass and low pass filtering of the time domain signal as shown in Fig.3. This filter pair is called the analysis filter pair . First, the low pass filter is applied for each row of data, thereby getting the low frequency components of the row. But since the low pass filter is a half band filter, the output data contains frequencies only in the first half of the original frequency range. They are down-sampled by two, so that the output data contains only half the original number of samples. Now, the high pass filter is applied for the same row of data, and similarly the high pass components are separated [5].

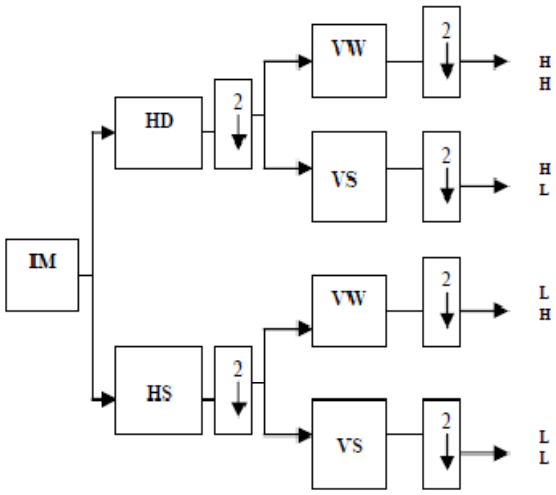


Fig.3.Decomposing signals into sub-bands

IM-Image, HD-Horizontal Decomposition, HS- Horizontal Scaling, VW- Vertical Wavelet, VS-Vertical Scaling

The output obtained from DWT is the decomposed image vector(C) and the book-keeping matrix(S).

$$C = [a(n)\ \ h(n)\ \ v(n)\ \ d(n)\ \ h(n-1)\ \ \ldots\ \ v(1)\ \ d(1)]$$

where a, h, v and d are columnwise vectors containing approximation, horizontal, vertical and diagonal coefficient matrices, respectively. C has 3n+1 sections where n is the number of wavelet decompositions. S is an (n+2)x2 bookkeeping matrix:

$$S = [sa(n,:)\ \ sd(n,:)\ \ sd(n-1,:)\ \ \ldots\ \ sd(1,:)\ \ sx]$$

Sa-approximate size entry, sd-detailed size entry

The equations for Daubechies wavelet is given by,

$$
\begin{aligned}
h(z) &= h_3(z^3+z^{-3})+h_2(z^2+z^{-2})+h_1(z+z^{-1})+h_0\\
g(z) &= g_4(z^4+z^{-4})+g_3(z^3+z^{-3})+g_2(z^2+z^{-2})+g_1(z+z^{-1})+g_0\\
\tilde{h}(z) &= \tilde{h}_4(z^4+z^{-4})+\tilde{h}_3(z^3+z^{-3})+\tilde{h}_2(z^2+z^{-2})+\tilde{h}_1(z+z^{-1})+\tilde{h}_0\\
\tilde{g}(z) &= \tilde{g}_3(z^3+z^{-3})+\tilde{g}_2(z^2+z^{-2})+\tilde{g}_1(z+z^{-1})+\tilde{g}_0
\end{aligned}
$$

$$
\begin{aligned}
h_0 &\doteq 0.788\,485\,616\,614, & g_0 &\doteq 0.852\,698\,679\,009, & \tilde{h}_0 &= g_0,\ \tilde{g}_0 = h_0,\\
h_1 &\doteq 0.418\,092\,273\,333, & g_1 &\doteq -0.377\,402\,855\,613, & \tilde{h}_1 &= -g_1,\ \tilde{g}_1 = -h_1,\\
h_2 &\doteq -0.040\,689\,417\,620, & g_2 &\doteq -0.110\,624\,404\,418, & \tilde{h}_2 &= g_2,\ \tilde{g}_2 = h_2,\\
h_3 &\doteq -0.064\,538\,882\,646, & g_3 &\doteq 0.023\,849\,465\,019, & \tilde{h}_3 &= -g_3\ \tilde{g}_3 = -h_3,\\
 & & g_4 &\doteq 0.037\,828\,455\,507 & \tilde{h}_4 &= g_4
\end{aligned}
$$

The filters for lifting scheme implementation are,

$$
\begin{aligned}
p_1(z) &= \alpha(z+1)\\
u_1(z) &= \beta(1+z^{-1})\\
p_2(z) &= \gamma(z+1)\\
u_2(z) &= \delta(1+z^{-1})
\end{aligned}
$$

with scale factors $K_s = \zeta$ and $K_d = 1/\zeta$ where

$$
\begin{aligned}
r_0 &= \tilde{h}_0 - 2\tilde{h}_4\cdot\tilde{h}_1/\tilde{h}_3 & \alpha &= \tilde{h}_4/\tilde{h}_3 \doteq -1.586\,134\,342\,060\\
r_1 &= \tilde{h}_2 - \tilde{h}_4 - \tilde{h}_4\cdot\tilde{h}_1/\tilde{h}_3 & \beta &= \tilde{h}_3/r_1 \doteq -0.052\,980\,118\,573\\
s_0 &= \tilde{h}_1 - \tilde{h}_3 - \tilde{h}_3\cdot r_0/r_1 & \gamma &= r_1/s_0 \doteq 0.882\,911\,075\,531\\
t_0 &= r_0 - 2r_1 & \delta &= s_0/t_0 \doteq 0.443\,506\,852\,044\\
 & & \zeta &= t_0 \doteq 1.149\,604\,398\,860
\end{aligned}
$$

VI. EXPERIMENTAL RESULTS

A.PERFORMANCE METRICS

The compression ratio can be measured as the ratio of the number of bits required to represent the image before compression to the number of bits required to represent the same image after compression and is given by the formula,

Compression ratio = size of original image/ size of compressed image

From the above equation, it is obvious that as the compression ratio increases the compression technique employed is more effective [2]. PSNR is often used as a quality measurement between the original and a compressed image. The higher the PSNR, the better the quality of the compressed or reconstructed image. To compute the PSNR, the *mean square error (MSE)* is calculated. Most image compression systems are designed to minimize the mean square error between two image sequences Ψ1 and Ψ2, which is defined as

$$MSE = \sigma_e^{\,2} = \frac{1}{N}\sum_{t}\sum_{x,y}[\Psi_1(x,y,t)-\Psi_2(x,y,t)]^2$$

The *peak-signal-to-noise ratio (PSNR)* in decibel (dB) is more often used as a quality measure in image coding, which is defined as

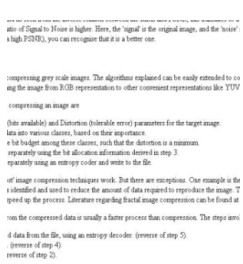
$$PSNR = 20\log_{10}\frac{255}{MSE}$$

After calculating the PSNR and compression ratio for different types of images, a performance analysis is done to effectively compute the process and substantially explain the insight behind the approach.

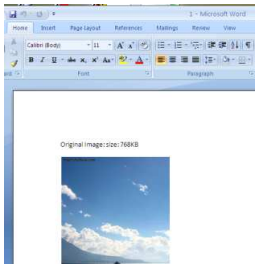
B.TEST IMAGES

The different types of images such as Natural image, Text image, Desktop image and Compound image as shown in Fig.4 were used to test the performance of the proposed system based on the metrics compression ratio and PSNR at different *bits per pixel* (Bpp).

i) Natural image



ii) Text image



iii) Desktop image

iv) Compound image

Fig.4 Test Images

C. NUMERICAL RESULTS

The Table.1 given below shows the comparison of compression ratio and PSNR for different types of images for the proposed system at different bits per pixel (Bpp).

S.No	Input Image	*Bpp=0.8		*Bpp=1.0	
		Compression ratio	PSNR (dB)	Compression ratio	PSNR (dB)
1.	Natural Image	3.46:1	36.75	3.39:1	37.17
2.	Compound Image	4.85:1	42.86	4.923:1	46.58
3.	Text Image	1.65:1	21.85	1.622:1	22.80
4.	Desktop Image	2.14:1	41.75	2.193:1	41.90

***Bpp – Bits per pixel**

Table.1 Inference table depicting the performance of the proposed system for different types of images

D. CONCLUSION

The above inference table clearly depicts that the compression ratio and PSNR are very effective for compound images when compared to the other types of images. The visual quality seems to very good for desktop and natural images though their compression ratio is less compared to compound images. Thus the proposed system is very effective for compound, desktop and natural images with respect to visual quality and provides high compression ratio for compound images and an acceptable compression ratio for desktop and natural images. The proposed system is not very effective in compressing text images both in terms of compression ratio as well as visual quality. The proposed system shows higher compression ratio and PSNR for compound images when the bits per pixel are increased. From the above observations, it is concluded that H.264/MPEG-4 part 10, "Advanced Video Coding" standard using CABAC entropy coding is very effective in compressing compound images both in terms of compression ratio and PSNR at different bits per pixel. The pictorial results are show in Fig.5 and Fig.6 for different bits per pixel Bpp=1 and Bpp=0.8 respectively. The pictorial results clearly show that the visual quality of text images is very poor for the proposed system and the visual quality of other types of test images seems to be good.

REFERENCES

[1] Cuiling lan, Guangming Shi and Feng wu, "Compress Compound Images in H.264/MPEG-4 AVC by exploiting Spatial Correlation" IEEE Transactions on Image Processing, vol.19 no.4, pp. 946-957, April 2010.

[2] Wenpeng ding, Yan lu and Feng wu, "Enable Efficient Compound Image Compression in H.264/AVC Intra Coding" IEEE Transactions on Image Processing, Vol.10 no.3, pp. 337-340, Sep.2009.

[3] Florinabel D J, Juliet S E, Dr Sadasivam V, "Efficient Coding of Computer Screen Images with Precise Block Classification using Wavelet Transform". Vol.91no.5, pp.856-562, May 2010.

[4] Jagannath D.J and Shanthini Pandiaraj, "Lossless Compression of a Desktop Image for Transmission" International Journal of Recent Trends in Engineering, Vol.2 no.3, pp. 27-29, Nov.2009.

[5] B.-f Wu, C.-C Chiu and Y. –L Chen "Algorithms for compressing compound document images with large text/background overlap", IEEE Proc.Vis. Image signal Process, Vol. 151 no. 6, pp.453-459, December 2008.

ACKNOWLEDGEMENT

I am very thankful to my management and colleagues of DIT, Dehradun to provide me such a good platform to represent my research work.

BIOGRAPHIES

Mr.P.S.Jagadeesh Kumar, Associate Professor, Electrical Engineering Department, DIT University, Dehradun has 13 years of teaching experience. He received his B.E degree from University of Madras in EEE discipline in the year 1999. He obtained his M.E degree in 2004 with specialization in CSE from Annamalai University, presently pursuing PhD in Anna University, Chennai.

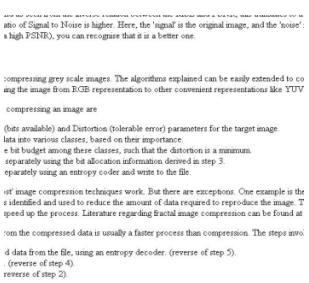
Dr.Gagan Singh, Professor & Head of Electrical Engineering Department, DIT University, Dehradun has 13 years of teaching experience. He received his doctorate from Uttarakhand Technical University. He also received the Young Scientist Award in first Uttaranchal State Science Congress by Uttaranchal Council of science & Technology, Government of Uttaranchal, in Nov'2006.

5.1 Original natural image

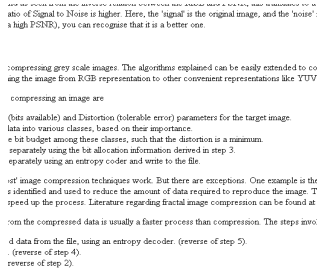
5.2 Compressed natural image

5.3 Original compound image

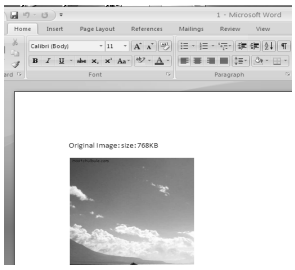
5.4 Compressed compound image



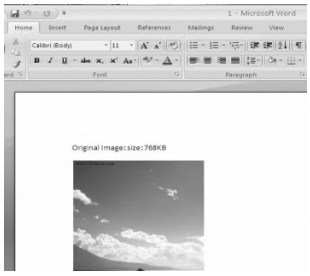
5.5 Original text image



5.6 Compressed text image



5.7 Original desktop image



5.8 Compressed desktop image

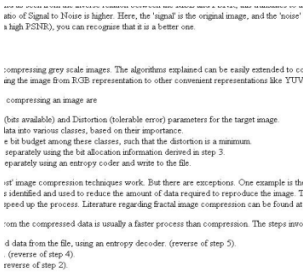
Fig.5 Visual quality of the different types of test images under the proposed system for Bpp=1

6.1 Original natural image

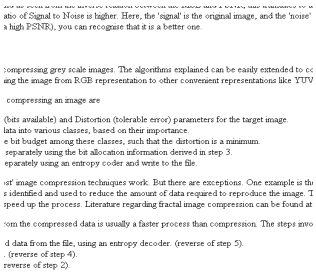
6.2 Compressed natural image

6.3 Original compound image

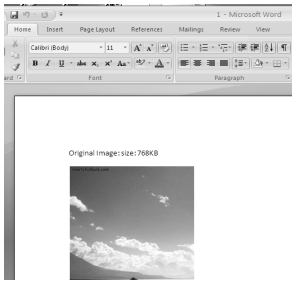
6.4 Compressed compound image



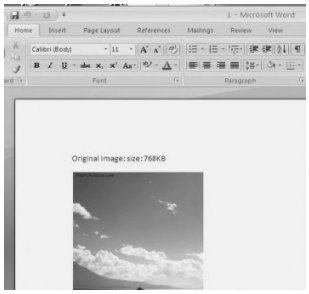
6.5 Original text image



6.6 Compressed text image



6.7 Original desktop image



6.8 Compressed desktop image

Fig.6 Visual quality of the different types of test images under the proposed system for Bpp=0.8