

Analysis of Video Quality Variation with Different Bit Rates of H.264 Compression

Yen-Jen Chen¹, Yang-Jen Lin¹, Sheau-Ling Hsieh^{2*}

¹Department of Electronic Engineering, Ming Chi University of Technology, Taiwan

²Information Technology Service Center, National Chiao Tung University, Taiwan

Email: yjchen@mail.mcut.edu.tw, M00158011@mail2.mcut.edu.tw, *sl_hsieh@cc.nctu.edu.tw

Received 16 March 2016; accepted 19 May 2016; published 26 May 2016

Abstract

The study applied a charge-coupled device (CCD) camera to send video signals to 4 DaVinci™ development boards (TMS320DM6446) of Texas Instruments (TI) to carry out H.264 Baseline Profile video coding. One of the development boards coded in the Variable Bit Rate (VBR) mode, and the other three development boards coded in the Constant Bit Rate (CBR) mode. In addition, the constant rates are 2 Mbps, 1.5 Mbps and 1 Mbps respectively. The H.264 video compression files produced by the boards were analyzed via video analysis software (CodecVisa) in the study. This software can analyze and present the compression data characteristics of the video files under each video frame, *i.e.*, bits/MB, QP, and PSNR. In this research, the characteristics of data of each frame under four different compression conditions were compared. Their differences were calculated and averaged, and the standard deviation was evaluated. It was further connected with the values of quality characteristics and the peak signal to noise ratio (PSNR) of each frame to analyze the relation among the frame quality, the compression rate of CBR, as well as the quantitative granularity. The preliminary conclusion of the study is that the compression behaviors of CBRs in different coding sources are adjusted in a specific proportion in order to cope with the change in frame complexity. The frame will be severely damaged by a critical value during the process of network transmission while the source rate is less than the value of the characteristic.

Keywords

H.264 Video Coding, CodecVisa, PSNR, VBR, CBR, QP

1. Introduction

H.264/AVC is a widely used international video coding standard, and its coding technique consists of VCL (Video Coding Layer) and NAL (Network Abstraction Layer). It also considered the coding efficiency, the network adaptability, and the transmission efficiency. It has made an advanced improvement comparing to commonly being used coding technique like MPEG-4 [1].

Network traffic is unpredictable when the H.264 video signal transmits streaming video under a constant and

*Corresponding author.

limited network bandwidth when lots of network packets are needed to be transmitted at the same time. The drastic change in a video scene will result in unstable network traffic. As for the video bit stream generated by a video signal compressed through the H.264 algorithm, its source rate consists of two modes: VBR (Variable Bit Rate) and CBR (Constant Bit Rate) [2]. As shown in **Figure 1** and **Figure 2** [3], VBR is to fix QP (Quantization Parameter) after video compression, so the data compressed source rate is low when the change of frames in a video segment is not high, and vice versa. The video compressed source rate varies with the complexity and the change of frames. CBR uses the rate control to dynamically adjust QP according to the complexity and the change to make a video segment have a thicker quantitative scale if its frames are complex or changing frequently, or have a thinner quantitative scale if its frames are static or not changing frequently. As a result, the entire video compressed source rate almost maintains constant. No matter in which mode the network camera is set, its source rate is generated according to the frame content of video rather than the currently available network bandwidth. Therefore, when a video passes through a shared-bandwidth network, such as the Internet, its frequency and frame sharpness are unpredictable regardless of its H.264 source rate mode, either VBR or CBR.

This study outputs the source video signal at a controlled source rate in either VBR or CBR mode. It sets the resolution and the coding time length to generate the coding packets with the H.264 Baseline Profile by using the TI DM6446 development board. Through using CodecVisa Ver. 4.33, a paid video analysis software, it analyzes the H.264 image frame by frame and then analyze down to the macroblock (MB) pattern information from a single frame. In this research, the characteristics of data of each frame under different compression conditions were compared. Their differences were calculated and averaged, and the standard deviation was evaluated. It was further connected with the values of quality characteristics and the peak signal to noise ratio (PSNR) of each frame to analyze the relation among the frame quality, the compression rate of CBR, as well as the quantitative granularity. The preliminary conclusion of the study is that the compression behaviors of CBRs in different coding sources are adjusted in a specific proportion in order to cope with the change in frame complexity. The frame will be severely damaged by a critical value during the process of network transmission while the source rate is less than the value of the characteristic. The rest of this paper is as follows. Section 2 describes the background of H.264 video coding and related works. Section 3 shows the research method. The experiment results and corresponding analysis are shown in Section 4. The final section concludes this paper.

2. Background and Related Works

The H.264 video coding method is an advanced video coding standard established on the basis of MPEG-4, which is officially named as H.264/AVC [4]. Its coding flow mainly includes the following 6 parts: inter-prediction, intra-prediction, transform, quantization, loop filter, and entropy coding [5]. The video coding mechanism of H.264 is block-based, *i.e.* the entire video is cut into many rectangular areas which are called the macroblocks (MB). These macroblocks are then coded. Either the intra-prediction or the inter-prediction is used to remove the similarity between video signals to obtain the so-called residual. The visual redundancy is removed

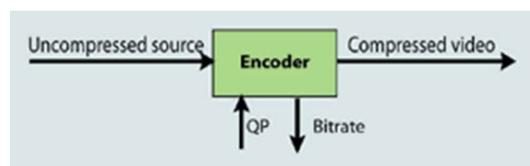


Figure 1. VBR fixes the QP value [3].

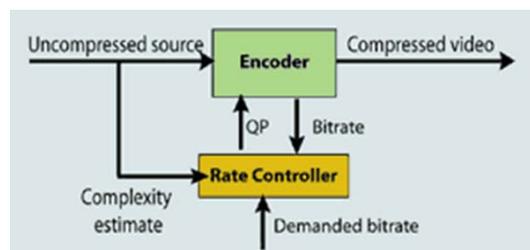


Figure 2. CBR dynamically adjusts the QP value [3].

from the residual through spatial transform and quantization. The VCL (Video Coding Layer) entropy coding will output the coded bit stream, which is then packed into the NAL-unit packet of NAL (Network Abstract Layer) and sent to a far-end device or stored into a storage medium via the network.

Let's take the most basic H.264 Baseline Profile coding specification as an example. The code of a video uses GOP (Group of Pictures) as the coding unit. As shown in **Figure 3**, each GOP consists of one I frame and 29 P frames, and each frame consists of a single set of slices or multi-set of slices. Each set of slices consists of several MBs. The VCL code calculates the residual value of each MB and the MB of the previous frame. The residual value is then sent to NAL for packing.

I-frame is a complete frame and also the first frame of a GOP, which will not carry out predictive coding with the previous frame or the next frame, and it has the highest and important information content in coding and decoding. P-frame is a coding frame predicted by the previous I-frame or P-frame. The coder needs not to record MB which is not changed in the last P-frame of the two frames. Thus, the MB information content of the P-frame is lower than that of the I-frame.

There are quite a few documents about the H.264 video compression, but these research techniques focus on optimizing VBR coding and controlling images to maintain their quality [6] or analyzing contributions of I/P/B frame under different GOP lengths [7]. The reason why the analysis software CodecVisa Ver. 4.33 used in this study is because this software complies with the demand of the study. It can capture data of the H.264 video in all decoding phases from the video sequence of the outmost layer to the macroblock of the image layer. Detailed statistic information is that bits/MB, QP value, and PSNR (Peak Signal-to-Noise Ratio) are taken out from each frame to discuss the QP variation of VBR and CBR in each frame and the change between bits/MB to find out the geometric proportion. This is the biggest difference from other literatures.

3. Research Method

The DaVinci™ development boards (TMS320DM6446) used in this study requires 3 system environments, which are Linux, MontaVista and EVM respectively. A computer, having Linux installed, is used to establish a NFS (Network File System) space to let DM6446 access the directory and execute the program code, and download the dvsdk suite and the H.264 Baseline Profile decoding package from the TI website. They will be stored in the directory/home/user. The second is MontaVista, the system core of DM6446, which is an embedded operating system based on Linux kernel and established in the directory/opt of the NFS. The third is EVM, the controlled environment of DM6446, which uses the UART cable to connect between the DM6446 development board and the Linux computer, transmit data in RS-232 mode, open the terminal console to set parameter between the development and the Linux computer, *i.e.*, baud rate, development board IP address, NFS Server IP address, uImage file path, and memory space allocation. After connecting, a person can give instructions to the development board to execute a program by the terminal console.

In this study, the Linux computer is installed with RedHat 4, which is a virtual computer established by VMware in a real computer installed with Windows 7. The virtual computer network is set as bridged to let the Linux virtual computer connect with the DM6446 development board. The H.264 coding and decoding package can be downloaded from the TI website. The H.264 coding and decoding functions of this package are Baseline Profile. Its mode is VBR coding mode by default, and can be set to the CBR coding mode to control the video compressed source rate. The two coding modes are being used in this study.

After the CCD camera captures a video signal, it will simultaneously send the video signal to 4 DM6446

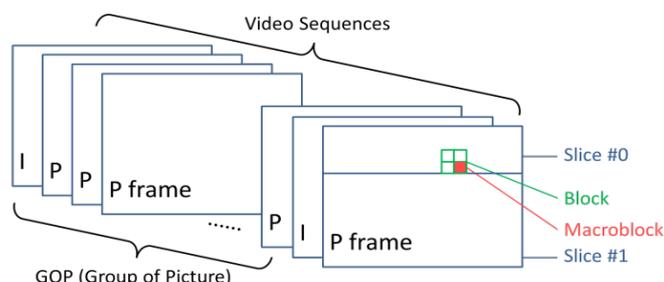


Figure 3. H.264 GOP and the coding mechanism for Baseline profile [5].

development boards to do H.264 video coding through a 1-to-4 Video AMP/Splitter, as shown in **Figure 4**. The 4 development boards are coded by the default VBR coding mode and three CBR coding modes. The CBR source rates are controlled at 2 Mbps, 1.5 Mbps and 1 Mbps, respectively. The coded video files are sent to the NFS folder of the Linux PC (RedHat 4). The coded file of the first development board (board 1) is named video_def.264; that of the second development board (board 2) is named video_2M.264; that of the third development board (board 3) is named video_1.5M.264, and that of the fourth development board (board 4) is named video_1M.264. Finally, the four H.264 files are copied from the NFS folder to the Windows 7 physical machine. The video analysis software (CodecVisa Ver. 4.33) captures contents of the four H.264 files and then shows the compression characteristics of each video frame. These values of characteristics are input into the Microsoft Excel to establish statistics charts for analyses.

The experiment is carried out in a manner that the Linux PC starts a terminal to execute the description file code.sh, to use the command sshpass to connect to the four DM6446 development boards and then to order the boards to simultaneously execute the description files having been written in the boards. The four description files are named EncBoard1.sh, EncBoard2.sh, EncBoard3.sh and EncBoard4.sh. The four boards almost execute video compression at the same time so that they have the common time basis when comparing frame quality subsequently. The contents of code.sh and EncBoard1.sh, EncBoard2.sh, EncBoard3.sh and EncBoard4.sh are listed as follows:

- Command content of encode.sh:

```
sshpass ssh root@192.168.1.200 sh EncBoard1.sh &
sshpass ssh root@192.168.1.201 sh EncBoard2.sh &
sshpass ssh root@192.168.1.202 sh EncBoard3.sh &
sshpass ssh root@192.168.1.203 sh EncBoard4.sh &
```

- Function content of EncBoard1.sh

Execute H.264 default VBR coding mode video for 5 seconds, and the image resolution is 720 × 480 (NTSC format)

- Function content of EncBoard2.sh

Execute H.264's CBR coding mode video for 5 seconds, the bit rate is set as 2 Mbps, and the image resolution is 720 × 480 (NTSC format)

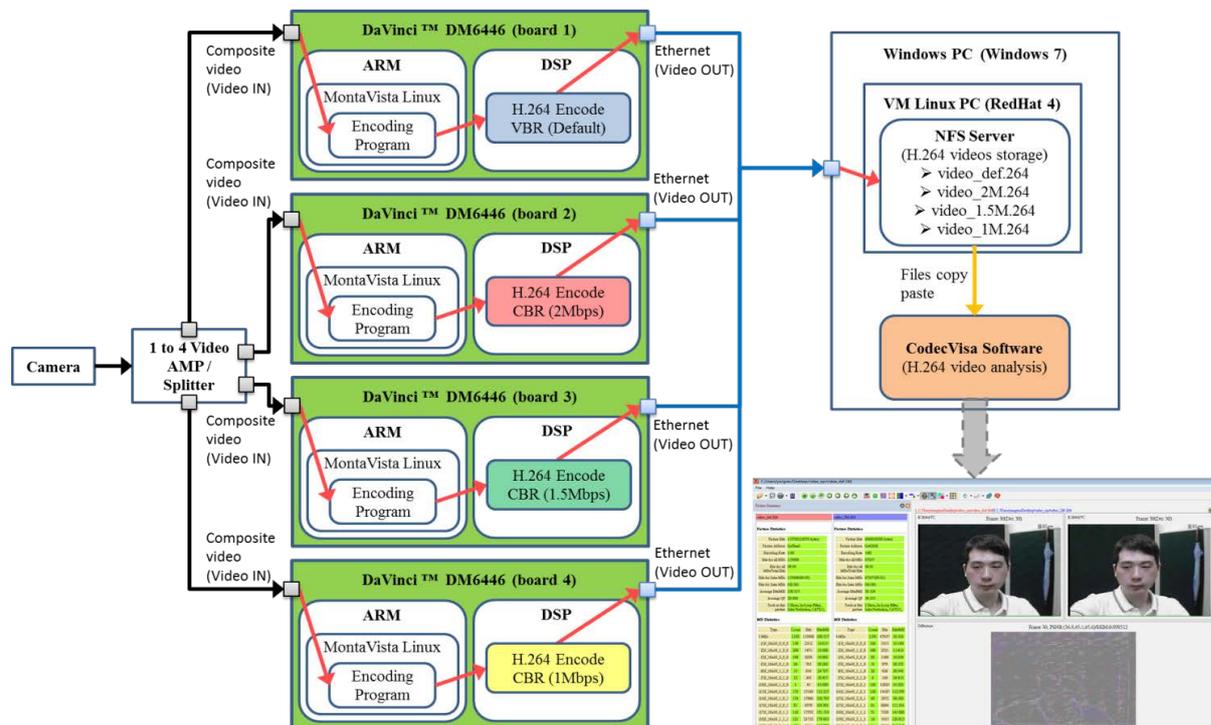


Figure 4. Experimental system framework.

- Function content of EncBoard3.sh

Execute H.264's CBR coding mode video for 5 seconds, the bit rate is set as 1.5 Mbps, and the image resolution is 720×480 (NTSC format)

- Function content of EncBoard4.s

Execute H.264's CBR coding mode video for 5 seconds, the bit rate is set as 1 Mbps, and the image resolution is 720×480 (NTSC format)

The software CodecVisa is applied to make the analysis, the analyzed outcomes of video_def.264 and video_2M.264 into the "comparison group 1", the analyzed outcomes of video_def.264 and video_1.5M.264 into the "comparison group 2", and the analyzed outcomes of video_def.264 and video_1M.264 into the "comparison group 3", which are then imported to Microsoft Excel and the chart data being used to verify the PSNR characteristics under different coding modes. The analyzed outcomes of video_1.5M.264 and video_2M.264 are made into the "comparison group A", and the analyzed outcomes of video_1M.264 and video_1.5M.264 are made into the "comparison group B", which are then imported to Microsoft Excel and the chart data being used to verify the QP values under different coding modes as well as the geometric proportion between Average bits/MB.

4. Experiment Results

In this study, it emphasizes the geometric proportion relation among the PSNR, bits/MB, and QP value of H.264 video coding under different coding modes and source rates through applying analyzing different groups of experiment films generated at 2 Mbps, 1.5 Mbps and 1 Mbps by the default VBR coding and CBR coding. The analyses focus on the parameter changes in continuous coding video frames. Since the early QP value of CBR is under an unstable state, it eliminates the first I-frame and the first P-frame in order to not let the initial value affect the entire statistic.

The experiment result is expressed in a manner that the video analysis software CodecVisa is used to extract the four video files compressed by the development boards, *i.e.*, video_def.264, video_2M.264, video_1.5M.264 and video_1M.264. The image information of each frame is read, and then the data is imported to Microsoft Excel for analysis. The analysis process is measured in mean and deviation, and the result is evaluated in standard deviation [8].

As illustrated in **Figure 5**, it shows the number of MBs of a single frame, and there are totally 1350 (45×30) MBs. The red block indicates this MB is generated by intra-prediction, and the green block indicates this MB is generated by inter-prediction. As shown in **Figure 6**, it calculates the bits/MB average of each frame of the four video files. For the experiment with the compression result file video_def.264, where the mode is VBR and the QP value is fixed at 28, the number of bits per MB (bits/MB) will highly increase when the video frame complexity increases, and vice versa. As for the 68th to the 74th frame, it has 7 frames continuously exceed 100 bits/MB. In the other side, the source rate is constant in the CBR coding mode, thereby the difference between the peak and the valley is not high, and it will have no continuous frames with a large number of bits. Since the early QP value of CBR is under an unstable state, it eliminates the first two frames, *i.e.* the number zero and one frames, called the 0th and 1st frames in this paper. Thus, for the CBR experiment at the source rate of 2 Mbps, the value of the 120th I-frame is 92.113 bits/MB for the peak, and the value of the 99th P-frame is 28.825 bits/MB for the valley.

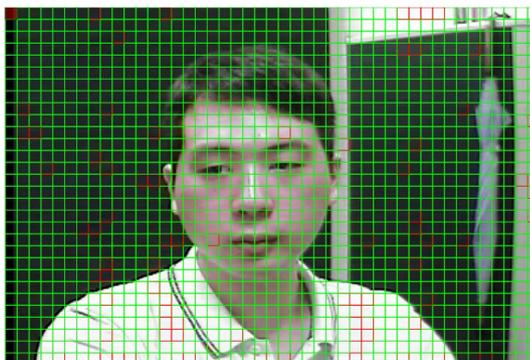


Figure 5. The all macroblocks in a frame.

As for the 2nd frame to the 150th frame, the average of the 149 frames is 49.766 bits/MB. The value of the 60th I-frame is 73.087 bits/MB for the peak at 1.5 Mbps, and the value of the 99th P-frame is 21.393 bits/MB for the valley. The average of the 149 frames is 36.893 bits/MB. The value of the 120th I-frame is 52.152 bits/MB for the peak at 1 Mbps, and the value of the 99th P-frame is 13.459 bits/MB for the valley. The average of the 149 frames is 24.45 bits/MB.

The QP average of the 1350 (45 × 30) MBs of each frame is shown in **Figure 7**. This figure presents the QP average of each frame. The QP default is 28 in the VBR coding mode; however, the QP average varies with the video complexity to keep the source rate constantly in the CBR mode. From this figure, the QP averages at 2 Mbps, 1.5 Mbps and 1 Mbps have been adjusted at a specific rate. In order to determine the characteristic of the rate adjustment, the QP difference is obtained in the way of subtraction of the QP averages of each frame. The difference, from that the QP average of the green line at 1.5 Mbps is used to subtract that of the red line at 2 Mbps, is defined as the “comparison group A”. The difference, from that the QP average of the blue line at 1 Mbps is used to subtract that of the green line at 1.5 Mbps, is defined as the “comparison group B”. The results are as shown in **Figure 8**.

As mentioned above, it omits the 0th and 1st frames in the early coding stage, and the 2nd to 150th frames are counted for the following statistical formulae. Thus, the QP difference, average, deviation and standard deviation are determined from the 149 frames.

- QP differences of the two comparison groups are expressed as below:
 Given comparison group A of QP difference is DA, and comparison group B is DB
 Given \overline{QP}_i (2 Mbps) is $V1_i$, \overline{QP}_i (1.5 Mbps) is $V2_i$ and \overline{QP}_i (1 Mbps) is $V3_i$;
 Given $i = 2 \dots n, n = 150$

$$DA = (V2_2 - V1_2, \dots, V2_i - V1_i, \dots, V2_n - V1_n) \tag{1}$$

$$DB = (V3_2 - V2_2, \dots, V3_i - V2_i, \dots, V3_n - V2_n) \tag{2}$$

As shown in **Figure 8**, it indicates the “QP difference” after subtraction of the QP averages of each frame at 3 different source rates in the CBR coding mode. The 2nd frame to the 150th frame is counted. The QP difference of the comparison group A ranges from 0.2666 to 5.305. The “average” of the 149 frames is 3.423. The QP

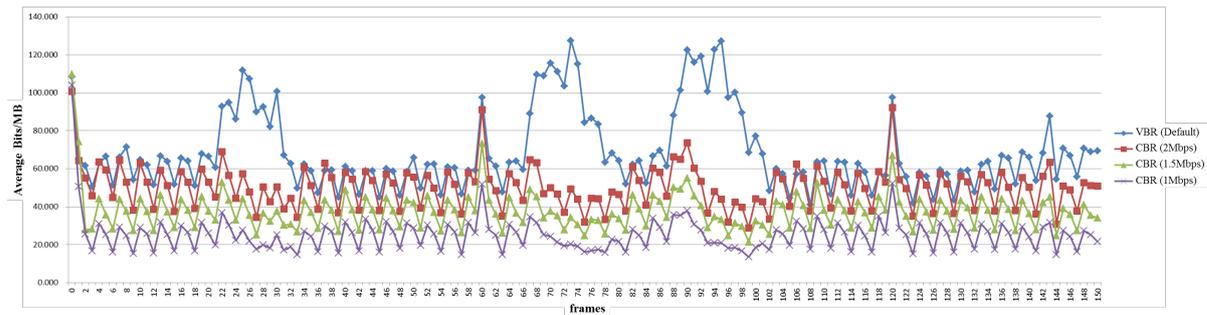


Figure 6. The average bits/MB line charts.

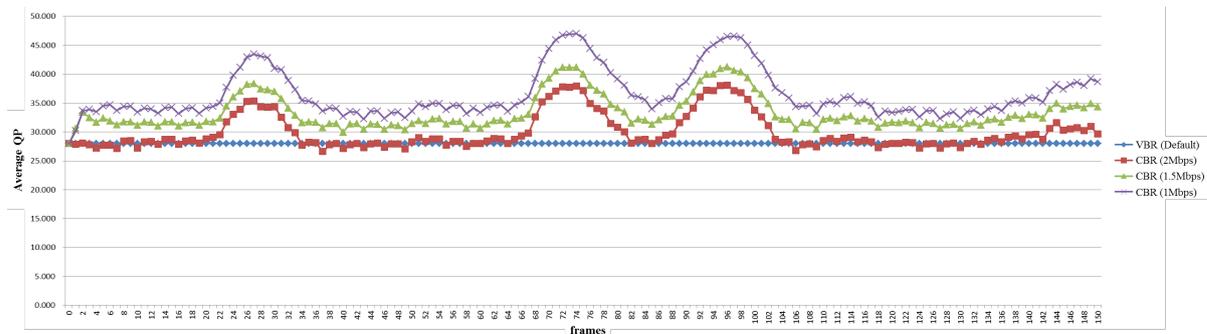


Figure 7. The average QP line charts.

difference of the comparison group B ranges from 0.278 to 6.371. The “average” of the 149 frames is 3.235.

- Averages of the two comparison groups are expressed ad below:

Given the average of DA is \bar{X}_{DA} , and that of DB is \bar{X}_{DB}

Given $i = 2, n = 150$

$$\bar{X}_{DA} = \frac{1}{n-1} \sum_{i=2}^n DA_i = \frac{1}{149} \sum_{i=2}^{150} DA_i = 3.423 \quad (3)$$

$$\bar{X}_{DB} = \frac{1}{n-1} \sum_{i=2}^n DB_i = \frac{1}{149} \sum_{i=2}^{150} DB_i = 3.235 \quad (4)$$

The averages of the two comparison groups are respectively 3.423 and 3.235 based on **Figure 8**. The average of each comparison group subtracts the value of each frame to obtain the “deviation” of the comparison group. The “absolute deviation” of the comparison group as shown in **Figure 9** is used to count the 2nd frame to the 150th frame. The absolute deviation of the comparison group A ranges from 0.001 to 1.88. The mean of 149 frames is 0.301. The absolute deviation of the comparison group B ranges from 0.021 to 3.136. The mean of 149 frames is 1.031. The results of the two comparison groups show that the mean absolute deviation of the comparison group A is 0.301, and it changes largely compared to the comparison B (with the mean absolute deviation, 1.031) when the frame complexity increases. Therefore, it is close to the proportion adjustment when the source rate is changed from 1.5 Mbps to 2 Mbps.

- Absolute deviation of the two comparison groups is expressed below:

Given the absolute deviation of DA is $|dA|$, and that of DB is $|dB|$

Given $i = 2 \dots n, n = 150$

$$|dA| = \left(\left| DA_2 - \bar{X}_{DA} \right|, \dots, \left| DA_i - \bar{X}_{DA} \right|, \dots, \left| DA_n - \bar{X}_{DA} \right| \right) \quad (5)$$

$$|dB| = \left(\left| DB_2 - \bar{X}_{DB} \right|, \dots, \left| DB_i - \bar{X}_{DB} \right|, \dots, \left| DB_n - \bar{X}_{DB} \right| \right) \quad (6)$$

The “standard deviation” of the comparison group is obtained that the deviation values are squared, totalized

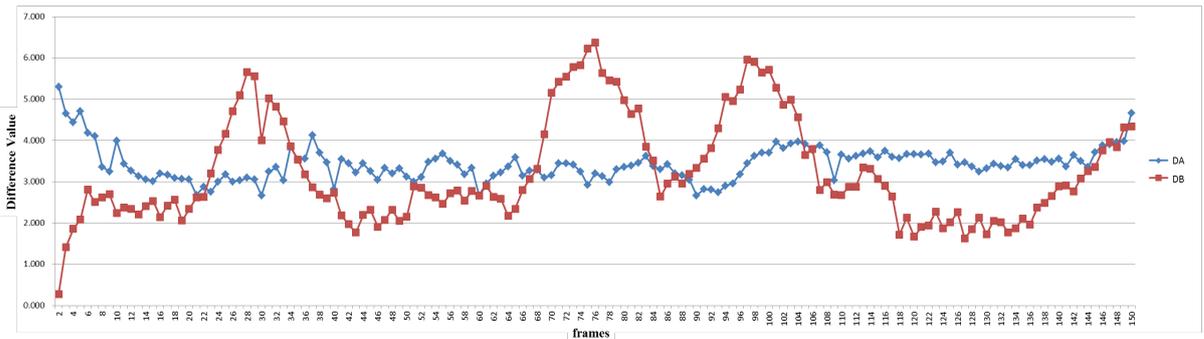


Figure 8. The diff. value of the two comparison groups.

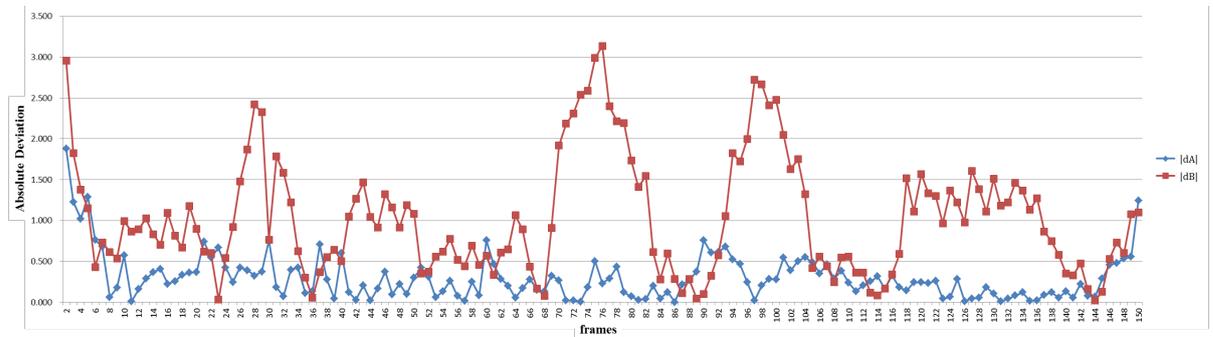


Figure 9. The abs. deviation of the two comp. groups.

and then divided by the number of values respectively, and then its square root is calculated. The standard deviation reflects the statistical distribution of the two comparison groups. When it is closer to 0, it complies more with the low variability and means that proportion exists in the QP adjustment between two different source rates. The above-mentioned result shows that when the source rate is 1 Mbps, the adjustment of the QP value will drastically change and lose its proportion to 1.5 Mbps due to the increase of the frame complexity compared to the QP values of 2 Mbps and 1.5 Mbps.

- Standard deviations of the two comparison group are expressed as below:

Given the standard deviation of DA is σ_{DA} , and that of DB is σ_{DB}

Given $i = 2, n = 150$

$$\sigma_{DA} = \sqrt{\frac{\sum_i^n dA_i^2}{n-1}} = \sqrt{\frac{\sum_2^{150} dA_i^2}{150-1}} = 0.410 \tag{7}$$

$$\sigma_{DB} = \sqrt{\frac{\sum_i^n dB_i^2}{n-1}} = \sqrt{\frac{\sum_2^{150} dB_i^2}{150-1}} = 1.253 \tag{8}$$

Figure 10 shows the actual image frames. When the complexity of a H.264 Baseline profile image file increases in the scene of hand vibration, the frame for the 1 Mbps experiment is severely damaged, for example the 73rd frame.

In image analysis and research, the most important data is PSNR (Peak Signal-to-Noise Ratio). Its analysis basis is to measure the quality between the original frame and the compressed frame, which ranges from 25 dB to 50 dB. If this value is higher, it represents the quality is better. In this study, the video_def.264 replaces the original image file to make a comparison with other 3 image files in the CBR coding mode for PSNR analysis. This is because video_def.264 is compressed by VBR. Its distortion rate is very low after extracting; thereby it is used to replace the original image file. As shown in **Figure 11**, it uses the analysis data of video_def.264 and video_2M.264 as the “comparison group 1”, that of video_def.264 and video_1.5M.264 as the “comparison group 2”, and that of video_def.264 and video_1M.264 as the “comparison group 3”. From the 3 comparison groups, the quality of each frame is measured according to PSNR. **Figure 11** shows PSNR is affected when the frame complexity increases. The frame quality has severely damaged when its corresponding PSNR value is 26.2 dB (**Figure 10** and **Table 1**) for the 1 Mbps experiment. As a result, the PSNR value of an image from an H.264 video with the CBR source rate of 1 Mbps will reduce to 28 dB below when the frame complexity increases; therefore, the frame becomes hardly identified.

- Information about experiment image file is as below:

Video Resolution: 720×480 (NTSC format)

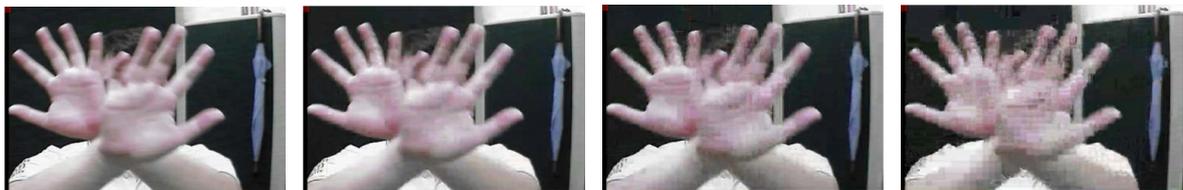


Figure 10. The first 73 frame (default, 2 Mbps, 1.5 Mbps, 1 Mbps).

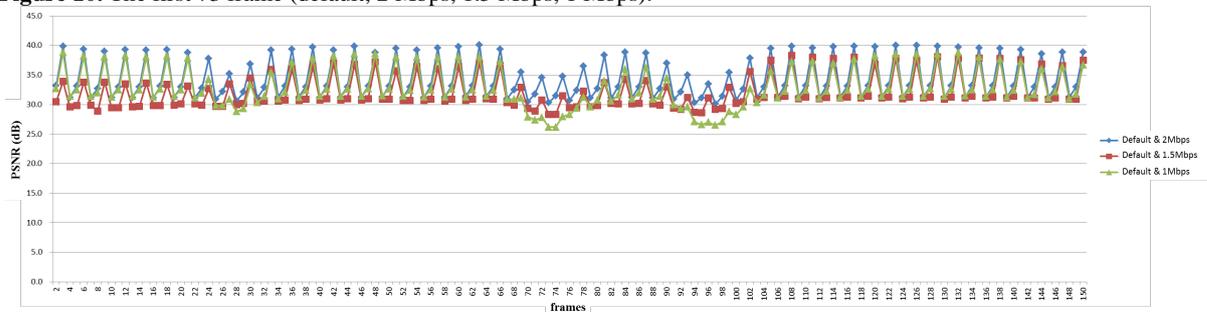


Figure 11. The compare PSNR line charts.

Table 1. PSNR all values of the result.

Value \ Bitrate	2 Mbps	1.5 Mbps	1 Mbps
Max	40.10	38.30	38.80
Minimum	30.10	28.30	26.20
Average	34.19	32.08	32.70

Coding Format: H.264 Baseline profile.

GOP (Group of Pictures): 30 frames.

FPS: 30 frames/sec.

Total frames: 151.

Video Times: 5 second.

5. Conclusions

The study used 4 DaVinci™ development boards (TMS320DM6446) to encode images captured by a CCD camera by following H.264 Baseline Profile. After encoding by VBR and CBR at 2 Mbps, 1.5 Mbps and 1 Mbps, the generated video files are imported to the video analysis software CodecVisa to read the characteristics of video compression data (e.g. bits/MB, QP, PSNR). These data is imported to Microsoft Excel to create statistics charts. From the statistical formulae, the standard deviations σ_{DA} and σ_{DB} are equal to 0.410 and 1.253, respectively, of the two QP-difference comparison groups for 2 Mbps-to-1.5 Mbps and 1.5 Mbps-to-1 Mbps. The above results with the observation of PSNR and actual frame image show that image identification is difficult from an H.264 video with the CBR source rate of 1 Mbps when the frame complexity increases. The conclusion is drawn from the above-mentioned experiment results that it is close to achieve the proportion relation between the two QP values for each frame compressed with CBR 1.5 Mbps and 2 Mbps respectively. In addition, when the complexity of a frame is large, it will be damaged or lost if its CBR source rate is lower than 1 Mbps.

For the sake of resource, this study only used 4 TMS320DM6446 development boards; therefore, more CBR source rates cannot be analyzed, e.g., 3 Mbps or 0.5 Mbps. However, the study verified that the corresponding QP values can be controlled to get close to the proportion when each frame of the H.264 image is adjusted at 2 Mbps and 1.5 Mbps.

References

- [1] Zhou, J.H., Shi, X.G. and Song, W. (2006) A New Frame Skipping Strategy in Rate Control Scheme Based on H.264 the Motion Complexity. *Journal of Zhejiang University of Technology*, **34**, 672-675.
- [2] Seeling, P. and Reisslein, M. (2012) Video Transport Evaluation with H.264 Video Traces. *IEEE Communications Surveys & Tutorials*, **14**, 1142-1165. <http://dx.doi.org/10.1109/SURV.2011.082911.00067>
- [3] Rate Control and H.264, PixelTools Corporation. http://www.pixeltools.com/rate_control_paper.html#
- [4] H.264/AVC Wikipedia. http://en.wikipedia.org/wiki/H.264/MPEG-4_AVC
- [5] Wiegand, T., Sullivan, G.J., Bjontegaard, G. and Luthra, A. (2003) Overview of the H.264/AVC Video Coding Standard. *IEEE Trans. Circuits Syst. Video Technol.*, **13**, 560-576. <http://dx.doi.org/10.1109/TCSVT.2003.815165>
- [6] Lee, H. and Sull, S. (2012) A VBR Video Coding for Locally Consistent Picture Quality with Small Buffering Delay under Limited Bandwidth. *IEEE Trans. on Broadcasting*, **58**, 47-56. <http://dx.doi.org/10.1109/TBC.2011.2164308>
- [7] Koumaras, H., Skianis, C., Gardikis, G. and Kourtis, A. (2005) Analysis of H.264 Video Coded Traffic. *INC 2005 Fifth International Network Conference*, Samos Island, July 2005, 441-448.
- [8] Chiou, H.J., Lin, P.F., Hsu, P.C. and Chen, Y.Y. (2012) Statistics: Principle and Application. Wu-Nan Book Inc., 71-85.