Macroblock level QP adaptation using SSIM for H.264 Encoder

Kirti M. Sargar, Anita B. Wagh

***Abstract*— Video Compression has seen great advances in last two decades. With development in research and ever growing processing power/network bandwidth H.264/AVC has become a favorite codec for network video transmission. We propose a technique to enable efficient transmission of video. We will be using Structural Similarity (SSIM) to adapt Macroblock level QP to obtain optimal Rate Distortion Optimization. The scheme uses the fact that neighboring MBs are correlated, for each coded MB, SSIM is calculated and used for predicting the optimal QP for the neighboring MB. This gives better result than the default JM encoder rate control algorithm.**

***Index Terms*—H.264, NS2, Rate Distortion Analysis, SSIM,**

I. Introduction

Due to limited wireless bandwidth, video signals need to be compressed efficiently, with video compression standards like H.264/AVC[1][2][3] this has become possible. For efficient transmission content complexity is analyzed and the video data is encoded using this information in Intra-prediction or Inter-prediction mode[1]. The reconstructed video and the original video are available for quality evaluation at the encoder end. However, it is observed metrics like Sum of Square Error (SSE) or Sum of Absolute Differences (SAD) are used for quality evaluation[4]. Research has indicated[9] that Square Error is not an ideal way for determining the quality of the video. As SSE doesn't take into account the structural information for the image. Structural Similarity (SSIM) was proposed as new indicator to measure the similarity of two images.

The human visual system (HVS) can extract image structure information easily, SSIM between the two-image can be used as a standard to evaluate the image quality. The maximum value of SSIM is 1, which indicates the two images are one and the same. The SSIM is better than the MSE(Mean Squared Error) and the PSNR (peak signal to noise ratio) in the image similarity evaluation. This new similarity measure has led to a lot of research in Image processing as well as Video Compression domain.

Section II we will view some basics of Rate Distortion Optimization (RDO) and current techniques which use SSIM in H.264 encoder. Followed by the proposed algorithm in Section III, then we will provide the experimental results and conclude the paper.

***Manuscript received May, 2014*.**
Kirti M. Sargar, Dept. of Electronics and Telecommunication, Smt. Kashibai Navale, College of Engineering, Pune, India, Pune, India
***Anita Wagh**, Dept. of Electronics and Telecommunication, Smt. Kashibai Navale, College of Engineering, Pune, India, Pune, India*

II. Background

Rate Control is required to ensure constant bitrate output from the encoder for unpredictable channel conditions. Most of the techniques [20][21][22] focus on encoder complexity to reduce the bitrate and obtain optimum RDO. Conventional optimal encoding schemes reduce distortion by allocating optimum bits for encoding [23][24].

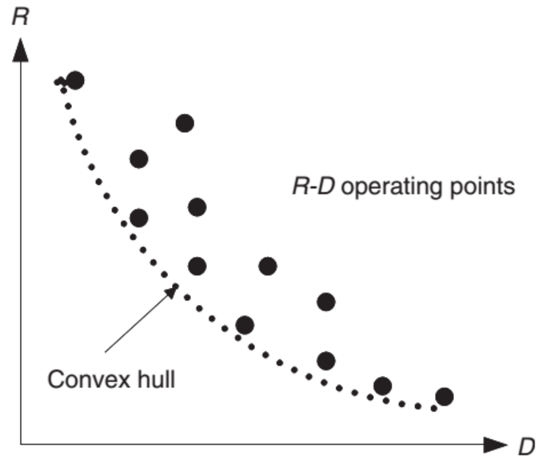

Fig 1. Rate Distortion curve for optimal operation points H.264/AVC

The Lagrangian formulation is provided [10] [24] which gives the balance between bitrate and quality as,

$$J = D_{SSE} + \lambda R \quad (1)$$

J is the cost function which needs to be minimized, it is a combination of distortion D and bitrate R, where λ is a lagrangian multiplier. It provides a trade-off between bitrate and distortion. Higher the value of λ more importance is provided to bitrate, while lower value of λ indicates higher importance to quality. For H.264 optimal value of λ obtained empirically[24] is

$$\lambda = 0.85 \times 2^{(QP-12)/3} \quad (2)$$

Thus higher QP value will reduce quality and improve bitrate and vice-versa. Although the main focus here is distortion D, which is traditionally calculated using SSE, SAD or Sum of Transformed Differences (SATD). As mentioned earlier traditional approaches are does not provide optimal insight in quality of reconstructed image. Over the past few years there have been various techniques have tried to incorporate SSIM as quality metric in H.264. Improved rate distortion algorithms are proposed[11][13] which incorporates SSIM for encoding quality evaluation. A RDO algorithm for I frame coding has been recommended[14], along with Lagrange multiplier adjustment to obtain a tradeoff between bitrate and quality. Similarly a RDO based on SSIM is proposed[18] which focuses on perpetual video quality. In [17] best intraprediction mode is calculated using SSIM. Another technique to decide of frame skipping has been proposed[15], an adaptive frame skipping mechanism is published which decides to skip frames based on SSIM metric obtained. The saved bits are thus used for encoding spatial quality for the following frames. A MB layer rate control scheme based on quadratic model has also been proposed[16]. The method also suggest use of quantization step by Lagrange multiplier method at MB level.

Although there are techniques available to adapt QP at MB level, it is necessary to have it more intuitive and use neighboring MB information which is already available from previously encoded MBs. In next section we will present our proposed algorithm, which uses SSIM for obtaining QP for a particular MB.

III. Proposed Algorithm

The proposed algorithm uses SSIM to obtain a correlation between the reconstructed image and the original video. SSIM[9] is given by following equation.

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3)$$

SSIM is Structural Similarity which takes value between $0 \leq SSIM \leq 1$. μ_x and μ_y are the mean value of image x and y. Similarly, σ_x^2 and σ_y^2 are variance of the images, and σ_{xy} is the cross correlation between the two images. c_1 and c_2 are constants with value << 1 to provide numerical stability to the above equation. Higher SSIM indicate better quality and correlation between the images.

SSIM is performed on block by block basis. We would be obtaining SSIM for each decoded MB in the adjacent neighborhood of the current MB which has to be encoded. Also, the adjacent MBs should contain from the same slice.

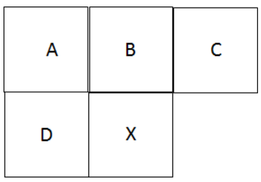

Fig 2. Current MB is denoted by X, neighboring Macroblocks SSIM value used for predicting the QP value.

$$QP_x = QP_{SLICE} + \frac{SSIM_A \times \delta_{QP}^A + SSIM_B \times \delta_{QP}^B + SSIM_C \times \delta_{QP}^C + SSIM_D \times \delta_{QP}^D}{SSIM_A + SSIM_B + SSIM_C + SSIM_D} \quad (4)$$

The modified QP formula is given in (4). The formula assumes the encoded MB has the same mode as the neighboring MB. So the best mode and optimum modes are selected using the information from neighboring MBs.

IV. Experiments and Results

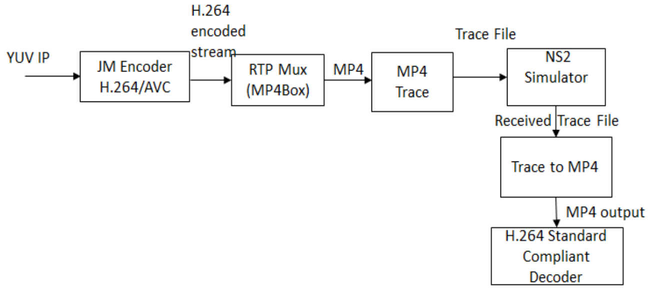

Fig 3. Evalvid video evaluation framework

Platform	X86 based
CPU	Intel Core 2 Duo 2.0 GHz
Memory	2GB
OS	Ubuntu 12.04
Compiler	GCC 4.8

Tab. 1. Hardware used for Simulation

The block diagram of Evalvid[8] video evaluation frame work is given in Fig.

1. We use uncompressed video sequences in CIF format (352 × 288) [5]. These video sequences are sampled in a videoconferencing format (format: 4:2:0).

2. After this, we compress the video files according to the H.264/MPEG-4 AVC standard. To this end, the videos sequences were encoded to 30 frames per second with a constant bit rate using the JM 18.6 software [4].

3. Then, the video frames are packaged using the RealTime Protocol (RTP). The MP4Box software is used to carry through this task [6]. The size of each packet is 1024 B.

4. Next, a trace file is generated in order to be used in the NS-2 network simulator. This file is created using the mp4trace software, which is part of the EvalVid project [7].

5. Finally, a traffic source is generated in the NS-2 network simulator. This goal is reached by establishing an UDP shipping agent and a receiving agent in the simulator [8].

The block diagram is shown in Fig 2. The specifications of hardware used is given in Fig. We observed the performance of the proposed scheme to be better than standard JM encoding. SSIM based MB level QP modifications gives better visual quality for the same bitrate. In the Fig. 3 and 4 a comparison of snapshots for JM sequence and the proposed algorithm is shown. The visual quality for the proposed algorithm is better than JM, it can be observed the detail regions of image are encoded better.

Fig 4. Subjective Quality comparison between original JM and SSIM based QP update algorithm.

Fig 5. Subjective Quality comparison between SSIM based QP update algorithm and original JM.

V. Conclusion

In this paper, we presented and evaluated a scheme using for getting optimum QP value for each MB, using information from neighboring Macroblocks. The main advantage is it gives better PSNR compared to original RDO algorithm used in JM 18.6 H.264 encoder. SSIM is a better way of comparing the quality of encoding compared to Mean Square Error (MSE) or Mean of Absolute Differences (MAD), as it takes into account the Structural information of image, rather than just using the pixel variation. The proposed method uses the correlation between neighboring MB to predict and optimum value of QP for the particular MB. An extension would be to perform multi-pass encoding to get better Rate Distortion Optimization.

Acknowledgment

References

[1] ITU-T, "Advanced video coding for generic audiovisual services," ITU-T Recommendation, H.264, Edition 8, April 2013

[2] S. Wenger, T. Stockhammer, and M. Hannuksela, "RTP payload Format for H.264 Video", draft-wenger-avt-rtp-h264-01.txt, Internet Draft, Work in Progress, March 2003

[3] S. Wenger, "H.264/AVC Over IP," *IEEE Transactions On Circuits and Systems for Video Technology*, vol. 13, no. 7, July 2003 pp 645-656.

[4] "JM v.18.6 software", Fraunhofer Heinrich-Hertz-Institut [Online] Available : http://iphome.hhi.de/suehring/tml/download/jm18.6.zip

[5] "Video Sequence" [Online] Available: http://www.tkn.tu-berlin.de/research/evalvid/cif.html.

[6] "MP4Box software" [Online]. Available: http://www.tkn.tu-berlin.de/research/evalvid/

[7] EvalVid, "A video quality evaluation tool" [Online]. Available: http://www.tkn.tuberlin.de/research/evalvid/

[8] S. McCanne, S. Floyd, and K. Fall, "The network simulator, ns-2" [Online]. Available: http://www.isi.edu/nsnam/ns

[9] Zhou Wang; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P., "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol.13, no.4, pp.600,612, April 2004

[10] Wiegand, T.; Girod, B., "Lagrange multiplier selection in hybrid video coder control," *International Conference on Image Processing,* 2001 , vol.3,no.,pp.542,545.

[11] Chen, H.H.; Yi-Hsin Huang; Po-Yen Su; Tao-Sheng Ou, "Improving video coding quality by perceptual rate-distortion optimization," *Multimedia and Expo (ICME), 2010 IEEE International Conference on* , vol., no., pp.1287,1292, 19-23 July 2010.

[12] An-Chao Tsai; Paul, A.; Jai-Ching Wang; Jhing-Fa Wang, "Efficient Intra Prediction in H.264 Based on Intensity Gradient Approach," *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on* , vol., no., pp.3952,3955, 27-30 May 2007.

[13] Chun-Ling Yang; Rong-Kun Leung; Lai-Man Po; Zhi-Yi Mai, "An SSIM-optimal H.264/AVC inter frame encoder," *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on* , vol.4, no., pp.291,295, 20-22 Nov. 2009

[14] Ziguan Cui; Xiuchang Zhu, "Subjective Quality Optimized Intra Mode Selection for H.264 I Frame Coding Based on SSIM," *Image and Graphics (ICIG), 2011 Sixth International Conference on* , vol., no., pp.157,162, 12-15 Aug. 2011

[15] Ziguan Cui; Xiuchang Zhu, "SSIM-based content adaptive frame skipping for low bit rate H.264 video coding," *Communication Technology (ICCT), 2010 12th IEEE International Conference on* , vol., no., pp.484,487, 11-14 Nov. 2010

[16] Ziguan Cui; Zongliang Gan; Xiuchang Zhu, "Structural similarity optimal MB layer rate control for H.264," *Wireless Communications and Signal Processing (WCSP), 2011 International Conference on* , vol., no., pp.1,5, 9-11 Nov. 2011

[17] Aswathappa, B.H.K.; Rao, K.R., "Rate-distortion optimization using structural information in H.264 strictly Intra-frame encoder," *System Theory (SSST), 2010 42nd Southeastern Symposium on* , vol., no., pp.367,370, 7-9 March 2010

[18] Chuohao Yeo; Hui Li Tan; Yih Han Tan, "On Rate Distortion Optimization Using SSIM," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.23, no.7, pp.1170,1181, July 2013

[19] Swaroop, K.V.S.; Rao, K.R., "Performance analysis and comparison of JM 15.1 and Intel IPP H.264 encoder and decoder," *System Theory (SSST), 2010 42nd Southeastern Symposium on* , vol., no., pp.371,375, 7-9 March 2010

[20] Akyol, E.; Mukherjee, D.; Yuxin Liu, "Complexity Control for Real-Time Video Coding,", *IEEE International Conference on Image Processing*, vol.1, no., pp.I - 77,I - 80, Sept. 16 2007-Oct. 19 2007.

[21] Y. Ivanov, C. Bleakley "Real-time H.264 video encoding in software with fast mode decision and dynamic complexity control," *ACM Trans. Multimedia Comput. Commun*. Appl. 6, 1, Article 5 (February 2010).
[22] Xiang Li; Wien, M.; Ohm, J.-R., "Rate-Complexity-Distortion Optimization for Hybrid Video Coding,", *IEEE Transactions on Circuits and Systems for Video Technology*, vol.21, no.7, pp.957,970, July 2011.
[23] Ortega, A., Ramchandran, K., and Vetterli, M. "Optimal trellis-based buffered compression and fast approximations" *IEEE Transactions on Image Processing* (1994), 3, 26– 40.
[24] Choi, J. and Park, D. "A stable feedback control of the buffer state using the controlled Lagrange multiplier method," *IEEE Transactions on Image Processing* (1994), 3(5), 546– 557.

Kirti M. Sargar

Has 3.5 years of teaching experience, she is currently pursuing M.E.in Signal Processing from Smt. Kashibai Navale College of Engineering Pune, India.

Anita B. Wagh

Has 5 years of teaching experience, currently she is working as Assistant Professor at Smt. Kashibai Navale College of Engineering Pune, She has done her M.E. in Electronics (Digital System)

