Scalable Feature Extraction for Remotely Motion Estimation and Texture Compensation for video

Yao-Chung Lin, Yun-Hsuan Sung, and Shu-Ping Yeh
{yclin79, yhsung, shuping}@stanford.edu

Conventional MC-DCT video coding explores the spatial correlation as well as temporal correlations at encoder by transformations and block-based motion estimations. Recently, some researches [1][2] revisited the theory of coding with side information developed by Slepian & Wolf [3] and Wyner & Ziv [4], and developed new coding schemes which discover the correlations at decoder.

The motion estimation at decoder is dual against the motion estimation at encoder. Thus, many existing fast motion estimation methods provide different directions to select the "hash code". Luo, *et al* [5] proposed *All Binary Motion Estimation (ABME)* which constructs binary pyramid structure for current frame and anchor frame, and matches them through bit-wise operation. Wang, *et al* [6] proposed an N-Queen decimated lattice pattern for block matching instead of full block comparison with less quality loss. The hash codes should not only provide the motion information but also texture information. In [7], portions of quantized DCT coefficients are used as texture hints and yield a significant saving of Wynzer-Ziv bits.

Our goal is to implement and compare different hash codes that help the decoder to do both the motion estimation and texture compensation. First, we will try the N-queen decimation pattern [6] and encode the quantized samples. Secondly, we will try to binarize current frame according to ABME [5] and send coded binary sequence to the decoder. Furthermore, we are going to develop other novel pattern decimation methods based on the previous results. In decoder, we will use several criterions, such as sum of squared difference (SSD), sum of absolute difference (SAD), Hamming distance, etc., to match image block from previous frame. Since the decoder can be unlimited powerful, the decoder can make reference to not just the previous frame, but also the frames before that. For application purpose, we may try to make the rate of hash codes scalable such that the partial decoded qualities are optimum for arbitrary truncated rates.

Reference

[1] B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero, "Distributed Video Coding", in Proc. IEEE, Special Issue on Advances in Video Coding and Delivery, 2003

[2] R. Puri and K. Ramchandran, “PRISM: A ‘reversed’ multimedia coding paradigm,” in Proc. International Conference on Image Processing, Barcelona, Spain, Sept. 2003.

[3] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,“ IEEE Transactions on Information theory, vol. IT-19, no. 4, pp. 471-480, July 1973.

[4] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” IEEE Transcations on Information Theory, vol. IT-22, no. 1, pp. 1-10, Jan. 1976

[5] Jeng-Hung Luo; Chung-Neng Wang; Tihao Chiang, “A novel all binary motion estimation (ABME)”, In Proc. IEEE, International Symposium on Circuits and Systems, 2002.

[6] Chung-Neng Wang, Shin-Wei Yang, Chi-Min Liu, and Tihao Chiang, “A Hierarchical N-Queen Decimation Lattice and Hardware Architecture for Motion Estimation”, IEEE Transaction on Circuits and Systems for Video Technology, Vol. 14 No. 4 April 2004.

[7] Shantanu Rane, "Hash-Aided Motion Estimation and Rate Control for Distributed Video Coding," EE 392J Digital Video Processing Course Project, Winter 2004