EE398B Project Proposal: WZ Codec with Increasingly Refined Motion Estimation

Suchit Kaul, Michael Padilla, and Oscar Trejo-Huerta

Stanford University

mtp@stanford.edu, suchitk@stanford.edu, otrejo@stanford.edu

1. Introduction

Typical video codecs place the majority of the computational burden on the encoder, resulting in a relatively simpler decoder. However, in some applications such as cellular phones or simple ad-hoc sensor networks, it would be favorable to reverse this, having a simple encoder and complex decoder [1].

In the case of codecs using motion estimation, both the encoder and decoder have complete knowledge of both the image data as well as the associated motion vectors. To efficiently encode frames, the encoder needs to perform the computationally expensive operation of motion vector estimation. If we let the motion vector estimation be Y and the current frame be X, we know that we can theoretically transmit at a rate of

$$R = H(X|Y)$$

without loss. While intuitively one might suppose that without knowledge of Y at the encoder performance would suffer, Slepian and Wolf [2] showed that theoretically performance remains the same so long as knowledge of the joint statistics is retained at both the encoder and decoder. When distortion is allowed, Wyner and Ziv [3] proved that there is only a small performance loss compared to the one given by the distortion-rate function of conventional coding of X given Y.

This project aims at leveraging this theoretical result to develop a system that has a simple encoder with performance approaching that of typical video hybrid coders through the use of WZ Codec with Increasingly Refined Motion Estimation.

2. Proposed Codec

The system is based on the concepts behind a typical hybrid codec. However, while such a hybrid system would compute the motion vectors accurately and once for each block, our codec does so in an increasingly refined fashion only at the decoder. This refined motion compensation helps the decoder improve its knowledge of the side information -i.e. the displaced block in the previous frame that best matches the current block.

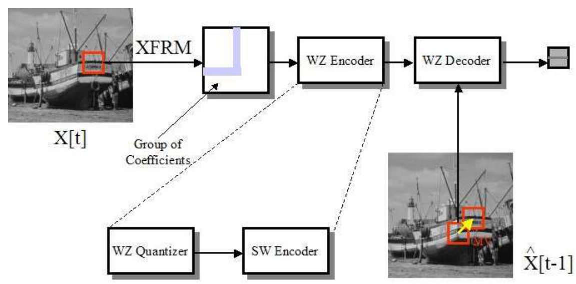

Figure 1: *WZ Codec with Increasingly Refined Motion Estimation.*

The basic operation of the codec is as follows. First, each frame is broken into blocks. Each block is transformed (typically with the DCT, optimal under high-rate assumptions [4]) and resulting coefficients are grouped into a predetermined order chosen by their ability to assist the decoder in accurate motion estimation. Each group of coefficients(GOC) is WZ coded, adapted to the joint statistics of the current GOC and the current reconstruction at the decoder as side information. Each coefficient in this GOC is scalar WZ coded, using optimal quantizers (uniform for high rate approximation [4]) followed by an ideal Slepian Wolf coder using the corresponding coefficient estimate at the decoder. Upon receiving each GOC, the decoder is able to further refine its motion estimates, resulting in increasingly accurate motion compensated reconstructions.

Overall, the block-wise operation of this codec may be broken down into stages:

1. First Stage: This notes the transmission to the decoder of the first GOC, no previous coefficients for the current frame having been received. Initially, side information (SI) is simply the co-located block in the previous frame. A coarse motion estimation is formed using the previously frame.

2. Second and Higher Stages: The next GOC is sent to the decoder, which combines this with the previous GOC(s) to refine its estimate of the current block. This improved reproduction of the current

block is then used with the reproduction of the previous frame to calculate a more refined motion estimation.

3. Goals

To evaluate our system, we will compare its R-D performance to that of state-of-the-art hybrid codecs. While we do not expect our system to outperform current generation codecs, we hope to be able to produce a system that produces comparable performance with the benefit of a considerably simpler encoder.

Practically and theoretically investigate a variety of implementation challenges and questions:

1. Transform coefficient groupings.
2. Optimal usage of previous MV estimates for current frame.
3. Changes necessary for operation with low bit rates.
4. Benefits of using vector side information (VQ) for low rates.
5. Adapting the encoder given changing side information at decoder.
6. Usage and design of hashes. Hashes are any small information sent by the encoder with the purpose of helping the decoder to perform motion compensation of the side information. This can be the same as the GOC, or additional information that is also sent.

4. References

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

[2] J.D. Slepian and J.K. Wolf, *Noiseless Coding of Correlated Information Sources*, IEEE. Trans. on Inf. Theory, Vol. IT-19, pp. 471-480, July 1973.

[3] A. D. Wyner and J. Ziv, *The Rate-Distortion Function for Source Coding with Side Information at the Decoder*, IEEE Trans. Inform. Theory, Vol. IT-22, Jan. 1976.

[4] D. Rebollo-Monedero, A. Aaron, and B. Girod, *Transforms For High Rate Distributed Source Coding*, Asilomar Conf. On Signals, Systems and Computers, 2003.

[5] D. Rebollo-Monedero, R. Zhang, and B. Girod,*Design of Optimal Quantizers for Distributed Source Coding*, DCC 2003.