Distributed source coding of video

Forchhammer, Søren; Van Luong, Huynh

Published in:
Proceedings of 8th Workshop on Information Theoretic Methods in Science and Engineering

Publication date:
2015

Document Version
Peer reviewed version

Citation (APA):
Forchhammer, S., & Van Luong, H. (2015). Distributed source coding of video. In *Proceedings of 8th Workshop on Information Theoretic Methods in Science and Engineering* (pp. 11-14). IEEE.

DISTRIBUTED SOURCE CODING OF VIDEO

Søren Forchhammer[1] *and Huynh Van Luong*[2]

[1]DTU Fotonik, Technical University of Denmark
2800 Kgs. Lyngby, sofo@fotonik.dtu.dk
[2]Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg
91058 Erlangen, huynh.luong@fau.de

ABSTRACT

A foundation for distributed source coding was established in the classic papers of Slepian-Wolf (SW) [1] and Wyner-Ziv (WZ) [2]. This has provided a starting point for work on Distributed Video Coding (DVC), which exploits the source statistics at the decoder side offering shifting processing steps, conventionally performed at the video encoder side, to the decoder side. Emerging applications such as wireless visual sensor networks and wireless video surveillance all require lightweight video encoding with high coding efficiency and error-resilience. The video data of DVC schemes differ from the assumptions of SW and WZ distributed coding, e.g. by being correlated in time and non-stationary. Improving the efficiency of DVC coding is challenging. This paper presents some selected techniques to address the DVC challenges. Focus is put on pin-pointing how the decoder steps are modified to provide adaptive decoding in distributed coding.

1. INTRODUCTION

Conventional video coding employs temporal prediction of frames to be coded. The apparent motion is represented by displacement vectors of blocks from previously coded data. This provides efficient coding, but also puts a heavy processing load on the encoder. In DVC an important issue is to use distributed techniques to encode the video frames individually, but utilize the temporal correlation on the decoder side for efficient video coding.

The Slepian-Wolf and Wyner-Ziv theorems addresses distributed coding in a set-up with two sequences, X and Y, each independent and identically distributed (iid), but jointly statistically dependent. The Slepian-Wolf theorem states that X can be independently encoded but decoded given the side-information (SI) Y at the same rate, $H(X|Y)$, as an optimal encoder having access to Y, under certain conditions. The Wyner-Ziv theorem extends this to the lossy case in a rate-distortion setting again under certain conditions.

We shall take this mind set but investigate it for real data in DVC were the assumptions of iid sequences do not hold. We shall use the term Side Information Generation to the processing of decoded data at a given point to provide estimates of the data, X, to be decoded. A prominent approach to DVC is Transform domain Wyner-Ziv (TDWZ) video coding [3], where a feedback channel is employed to let the decoder control the rate by requests. In the basic setting (called GOP2) every other frame (called Key Frames) is coded using intra-coding and the frames in between are coded using distributed techniques and decoded using the two surrounding frames as side information and called WZ frames. The feedback introduced serves to adapt the bit-rate as the required number of bits is varying and not known.

The TDWZ DVC coding architecture employs a DCT like transform on 4 x 4 blocks. While providing some decorrelation, there is still significant correlation in the transformed data. The coding efficiency has been improved considerably by a number of techniques.

In Sec. 2, we present a basic TDWZ DVC architecture as in [3] and improved in the DISCOVER codec [4]. In Sec. 3, improvements by making the decoder adaptive based on reestimations are presented. First to capture crossband correlations [5] and extended in the side information and noise learning (SING) codec [6] introducing an optical flow technique for motion estimation to compensate the weaknesses of the block based SI generation and in the motion and reconstruction reestimation (MORE) [7] codec, where the updated information is used to iteratively reestimate the motion and reconstruction. Finally, an adaptive mode decision (AMD) is investigated to take advantage of skip and intra mode in DVC by deciding the coding modes based on the quality of key frames and rate of WZ frames. Benchmark results of the resulting MORE-AMD [8] and the other techniques are briefly presented in Sec. 4. In Sec. 5, the SW coding based on rate-adaptive error-correcting techniques is revisited [9].

2. DISTRIBUTED VIDEO CODING

The architecture of a TDWZ video codec [4] is depicted in Fig. 1. In this codec, the sequence of frames is split into key frames and Wyner-Ziv (WZ) frames. Key frames are intra coded using conventional video coding techniques such as H.264/AVC intra coding. The Wyner-Ziv frames are transformed (4×4 DCT), quantized and decomposed into bitplanes.

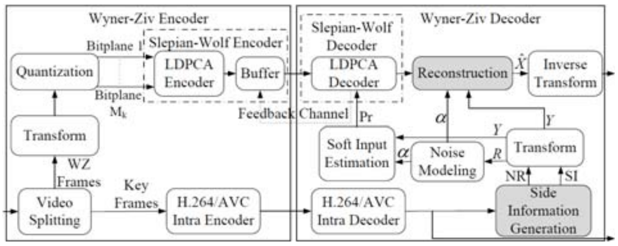

Figure 1. The architecture of TDWZ video codec.

Each bitplane is fed to a rate-compatible LDPC Accumulate (LDPCA) encoder [10] from most significant bitplane to least significant bitplane. The corresponding error correcting information is stored in a buffer and requested by the decoder through a feedback channel. The WZ frame is predicted at the decoder side by using decoded frames as references. The predicted frame, called the SI frame, is an estimate of the original WZ frame. Given the available SI, soft-input information (conditional probabilities *Pr* for each bit) within each bitplane is estimated using a noise model. Thereafter the LDPCA decoder starts to decode the bitplanes selected by the quantizer, ordered from most to least significant bitplane, to correct bit errors. The decoder requests bits from the buffer until the bitplane is decoded. Thereafter CRC check bits are sent for confirmation. After all bitplanes are successfully decoded, the WZ frame can be decoded through combined de-quantization and reconstruction followed by an inverse transform.

3. ADAPTIVE DVC USING REESTIMATIONS

Adaptive coding in distributed source coding is enabled by the feedback based request of parity bits for rate-adaptation. After a successful decoding of a code block, the decoder can update the side information and thereby the soft-information for decoding the next block. Thus the side-information used in the decoding may generally be seen as a mapping of the causal data and the frame level approach presented in Sec. 2 may be extended to (sub)band level and bit-plane level, where a code block in this section is given by the information required to decode a bit-plane of one coefficient subband. This adaptation using decoded blocks may also be applied in the motion estimation step using partially decoded frames. We shall first focus on how the decoder may introduce adaptive coding, while the encoder remains the same.

3.1. Crossband correlations in DVC

The Crossband DVC scheme [5] enhances the DISCOVER architecture using previously decoded subbands in the noise modeling for the next subbband. Specifically after decoding a subband, a classification is performed. When modeling a new subband one or two previous subbands are used to predict the classification. This subband level adaptive processing is combined with a bit-plane level updating of estimates for each coefficient.

3.2. Multi-hypotheses decoding using optical flow

In the SING codec [6] multi-hypothesis decoding was used integrate a number of decoder-side adaptive techniques.

3.2.1. Dense motion fields using optical flow

The motion field in the side information generation is backward adaptive in the DVC scheme, thus the motion (vectors) are not coded. This allows using a dense motion field. In the SING codec [6], global optical flow (OF) was used to calculate dense motion fields to supplement a more conventional overlapping block motion compensation (OBMC).

3.2.2. Multi-hypotheses decoding

The rate-adaptive LDPCA decoder may be fed with multiple sets of soft-information, and terminate and ‘selecting’ the set first to decode (subject to a CRC). This provides a generic approach to decoder side adaptation in distributed coding. SI based on both optical flow and block-based OBMC can e.g. be combined to provide multiple hypotheses [6].

3.2.3. Adaptive noise modeling

Different noise modeling may also be adaptively selected using the multi-hypothesis approach. In SING, techniques based on clustering of DCT blocks, calculating feature vectors and updating and refining these was applied. Distributions of the residuals from previous frames were also used and the number of clusters adapted.

3.3. Re-estimation of motion

A challenge in DVC, incl. the scheme presented so far is the prediction of the motion at the decoder side for the WZ frame, which is not available as opposed to conventional video coding, especially for sequences with high or complex motion. As the WZ frame is being decoded, also the motion may be reestimated. Two instances of this were introduced in the MORE codec [7]. The partially reconstructed frames were used to reestimate motion for both the optical-flow (after each band) and the block-based (OBMC after each bit-plane) techniques. This can improve the prediction of the values. To also improve the estimate of the distribution, the residue may also be motion compensated using an updated estimate of the residue of the previous WZ frame and the current motion estimation to calculate a motion compensated residue. These techniques were integrated in the SING [6] codec (Figure 2).

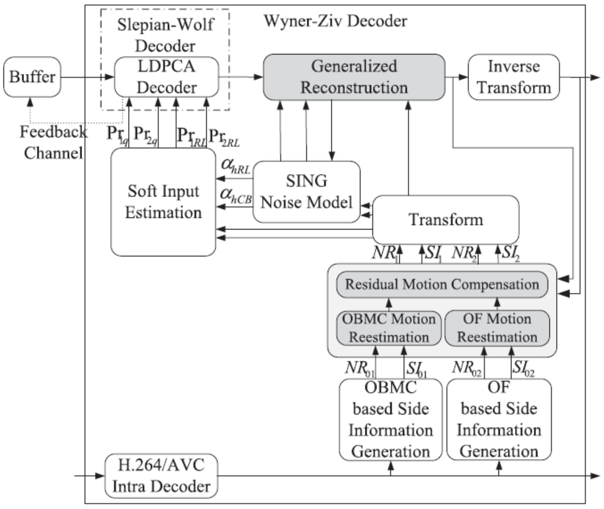

Figure 2. TDWZ decoder with the motion and reconstruction reestimation (MORE).

The combination of initial side information and reestimated side information based on reestimated motion in MORE is based on an adaptive selection process, trying to estimate influence on rate and distortion. The rate is estimated by the *ideal code length* (ICL), which after decoding a bit-plane may be calculated by summing minus log of the conditional probability assigned by the soft-input to each bit. A Lagrangian based rate-distortion cost function is used to adaptively selecting one block-based and one OF based estimate, SI and residual, NR, for the further processing to form the multiple inputs to the LDPCA decoder (Figure 2).

3.4. Decoder side driven adaptive mode decision

In video coding, skip coding and intra coding are used as additional modes. When advantageous in an operational rate-distortion sense [11] applying these modes improves performance. Introducing this adaptive mode decision (AMD) in DVC does require a change of the encoder to switch between modes. The decision can be encoder based introducing extra encoder processing steps and/or fed back from the decoder. Initial experimental results of integrating AMD in the MORE codec were reported in [8].

4. NUMERICAL RESULTS

The methods presented in Secs. 2 and 3 were tested on the four standard test sequences: *Foreman*, *Hall*, *Soccer* and *Coast* for a number of different bit-rates. Operational rate-distortion performance were calculated, expressing the quality by PSNR values [7,8]. The weighted average improvements (measured by Bjöntegaard differences [12]) over DISCOVER [4] are given in Tables 1 and 2. The resulting MORE [7] codec achieved an average improvement in PSNR is 2.5 dB on the WZ frames (for GOP2) and gained 1.2 dB measured over all frames. The performance of Crossband, SING, and MORE(AMD) are also given for comparison.

Table 1. Bjøntegaard PSNR improvement (dB) over DISCOVER for WZ frames

Codec	Crossband	SING	MORE	MORE(AMD)
Foreman	0.65	1.52	**3.00**	2.93
Hall	0.39	0.99	1.42	**1.95**
Soccer	1.33	2.70	**4.19**	4.182
Coast	0.36	0.41	0.65	**0.85**
Average	0.64	1.49	2.47	**2.58**

Table 2. Bjøntegaard PSNR improvement (dB) over DISCOVER for all frames

Codec	Crossband	SING	MORE	MORE(AMD)
Foreman	0.33	0.75	**1.43**	1.41
Hall	0.19	0.40	0.58	**0.61**
Soccer	0.73	1.51	**2.26**	2.23
Coast	0.19	0.22	0.27	**0.34**
Average	0.33	0.76	1.22	**1.22**

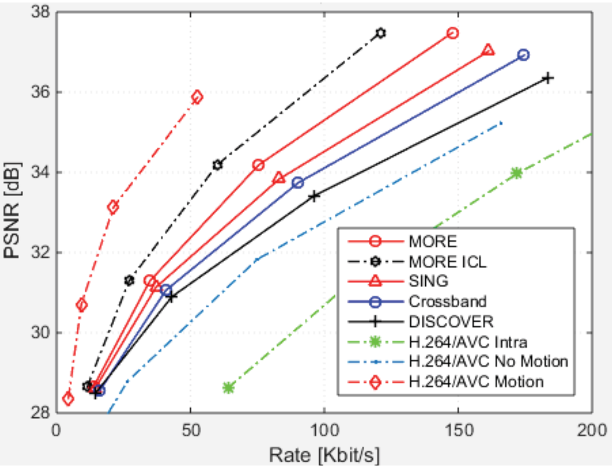

Figure 3. PSNR vs. rate for selected codecs for WZ frames (QCIF, 15Hz) for *Coast*.

The RD performance of the MORE, SING, and Crossband as well as H.264/AVC coding is also depicted in Figure 3 for the *Coast* sequence for WZ frames. In addition, the MORE ICL, which is obtained by replacing LDPCA coding with a calculation of the ideal code length (ICL) over all the decoded bitplanes, is also given. We calculate the ideal code length, ICL [6][7], at the decoder side based on the soft-input values used when decoding the information bits.

5. SW CODING REVISITED

Comparing the ideal code lengths (ICL) and the actual code lengths in DVC provides an evaluation of the loss in distributing the coding applying error-correcting coding, instead of e.g. arithmetic coding based on the conditional probabilities, see Figure 2. Investigations of the results obtained using the LDPCA code, widely used in DVC, show that especially for low conditional probabilities, there is a relatively high loss, which may be an issue in DSC in general. As an alternative rate-adaptive BCH [9] for coding with feed-back was studied. The feedback provides the capability to adapt to unknown statistics and also to reduce the coding loss FEC codes endures when small code blocks are used. Linear block codes with extensible parity matrix, H, may readily be used for rate-adaptive coding, extending the matrix and sending new syndromes when more information is requested.

In [9], a rate-adaptive BCH (RA-BCH) code was introduced and analyzed for the case of bounded distance decoding and assuming iid error probability with known error probability between the side information, Y, and the information data, X. The scheme also involved using syndromes for checking and making the number of syndromes used to confirm a decoding adaptive to the number of syndromes received thus far. For error probability, $p = 0.01$, $H(X|Y) \sim 0.08$ (which also gives the average ICL). Based on simulations with this set-up at a bit-error-rate of 10^{-5}, the average code lengths where ~ 0.10 for RA-BCH for length 1023 and ~ 0.144 for LDPCA of length 1584. For fixed rate coding, a bound of ~ 0.146 was calculated and for both fixed rate LDPCA and BCH of the lengths considered the rate would be above 0.2, thus showing the clear benefit of using feed-back for these short code block lengths. In these comparisons BCH was clearly better that LDPCA. The challenge towards using RA-BCH in DVC is to generalize to soft-input decoding.

6. CONCLUDING REMARKS

We have given a brief overview of elements of a state-of-the-art DVC scheme with focus on aspects which may be of general interest when applying DSC to real data, especially video data. This included ways to make a DSC scheme with feed-back adaptive on the decoder side. Also it was pointed out that as DVC and DSC schemes improve performance, the loss in current error-correcting techniques applied become an issue towards achieving distributed coding without out performance loss as suggested by the classic Slepian-Wolf and Wyner-Ziv papers.

7. REFERENCES

[1] D. Slepian and J. K. Wolf "Noiseless coding of correlated information sources", *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471-480, 1973.

[2] A. D. Wyner and J. Ziv "The rate-distortion function for source coding with side information at the decoder", *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1-10, 1976.

[3] B. Girod, A. M. Aaron, S. Rane and D. Rebollo-Monedero "Distributed video coding", *Proc. IEEE*, vol. 93, no. 1, pp. 71-83, 2005.

[4] (2007, Dec.). Discover Project [Online]. Available: http://www.discoverdvc.org/

[5] X. Huang and S. Forchhammer "Cross-band noise model refinement for transform domain Wyner-Ziv video coding", *Image Commun.*, vol. 27, no. 1, pp. 16-30, 2012.

[6] H. V. Luong, L. L. Rakêt, X. Huang, and S. Forchhammer, “Side Information and Noise Learning for Distributed Video Coding using Optical Flow and Clustering,” *IEEE Trans. Image Proc.*, vol. 21, no. 12, pp. 4782-4796, 2012.

[7] H. V. Luong, L. L. Rakêt, and S. Forchhammer, "Re-estimation of Motion and Reconstruction for Distributed Video Coding," IEEE *Trans. Image Proc.*, vol. 23, no. 7, pp. 2804-2819, 2014.

[8] H. V. Luong, S. Forchhammer, J. Slowack, J. De Cock, and R. Van de Walle, "Adaptive Mode Decision with Residual Motion Compensation for Distributed Video Coding," *APSIPA Transactions on Signal and Information Processing*, vol. 4, no. e1, pp. 1-11, 2015. (DOI: http://dx.doi.org/10.1017/ATSIP.2014.21).

[9] M. Salmistraro, K. Larsen and S. Forchhammer, “Rate-adaptive BCH codes for distributed source coding,” *EURASIP Signal Process. J.*, 2013.

[10] D. Varodayan, A. Aaron, and B. Girod, “Rate-adaptive codes for distributed source coding,” *EURASIP Signal Process. J.*, vol. 86, no. 11, pp. 3123–3130, 2006.

[11] J. Slowack, S. Mys, J. Skorupa, N. Deligiannis, P. Lambert, A. Munteanu, and R. Van de Walle, “Rate-distortion driven decoder-side bitplane mode decision for distributed video coding,” *Signal Process.: Image Commun.*, 25 (9) (2010), 660–673.

[12] G. Bjøntegaard, “Calculation of average PSNR differences between RD curves,” in *Proc. VCEG Contrib.*, Apr. 2001.