Improved Reconstruction for Distributed Video Coding

Ralph Hänsel, Erika Müller
Institute of Communications Engineering
University of Rostock
Rostock, Germany
$\{$*ralph.haensel, erika.mueller*$\}$*@uni-rostock.de*

***Abstract*—Distributed video coding (DVC) gains more and more interest in the recent decade. State-of-the-art Wyner-Ziv decoding is performed by a mandatory feedback channel. It's mandatory for rate allocation and high reconstruction quality. The feedback channel is a handicap for real application. If no feedback channel is available the RD-performance of the system is decreased.**

We propose a soft reconstruction algorithm, which significantly increases the reconstruction quality if no feedback channel is available. If a feedback channel is available, the proposed soft reconstruction shows the similar performance as the best state-of-the-art reconstruction algorithm (MMSE).

***Keywords*-Distributed Source Coding (DSC), Distributed Video Coding (DVC), Reconstruction**

I. Introduction

Distributed video coding (DVC) becomes more and more important in the recent years. It is based on the theories of D. Slepian, J. Wolf [1] and A. D. Wyner, J. Ziv [2]. In contrast to conventional coding systems (MPEG-4, H.264/AVC) distributed video coding gives the ability to develop low complexity encoders. Error-robust transmission and multiview video compression are further application fields.

The known DVC systems (e.g. [3], [4]) use a feedback channel. It's necessary for accurate rate allocation and thus successful decoding. But the use of a feedback channel is a problem in practical applications e.g. storage or realtime decoding.

Currently, most of the codecs use channel coding techniques (e.g. Turbo Codes, LDPC - Low Density Parity Check) to implement the Slepian-Wolf (SW) codec. They transform the quantization symbols to a binary representation before bit plane by bit plane Slepian-Wolf encoding. On the receiver side each bit plane is decoded using the Slepian-Wolf decoder. It is typically implemented by a channel code decoder. The hard output of the decoder (no confidence information) is used for reconstruction of the pixel values (see sec. III).

The hard output decoding and MMSE (minimum mean square error) reconstruction is the best known decoding method, if a feedback channel is available.

If only a reduced feedback channel is available. For example, the exact sum rate is known but the rate for each bit plane is unknown. Thus the feedback channel is only used to request higher sum rate but not for each bit plane. In this case the conventional bit plane by bit plane decoding fails. This problem is solved by inter bit plane decoding as proposed in [5]. Because bit planes are decoded in parallel and information is exchanged.

Another case of a reduced feedback channel setup is the fixed sum rate scenario. Here the feedback channel is used to control the rate for the significant bit planes. Due to the limited sum data rate and reduced feedback channel usage, the SW decoding of the last bit plane fails. Therefore the reconstruction quality is low. In this paper we propose a soft reconstruction algorithm, which improves the quality in this case.

If the setup provides no feedback channel the common solution in the literature is to estimate the necessary rate at the encoder [6]. But if the estimation is not accurate conventional decoding and reconstruction will fail. Therefore we propose the use of inter bit plane decoding in conjunction with soft reconstruction to significantly improve the reconstruction quality for this case.

In section II the basic distributed video coding system with the proposed soft reconstruction is described. The widely applied two state-of-the-art reconstruction algorithms are discussed in section III. The proposed soft reconstruction algorithm is presented in section IV, whereas simulation results are presented in section V and the conclusions are given at the end of this paper (sec. VI).

II. System Setup

A. Encoder

The basic pixel domain DVC codec used in this paper is shown in figure 1. Every pixel X_i of the grayscale input frame X is scaled and quantized obtaining the quantizer symbol q_i. A uniform 2^M step quantizer is used. The quantizer symbol is passed to the Slepian-Wolf encoder, which performs a convolutional encoding for the separated bit planes $q_i^{(b)}$ of q_i. The resulting parity bits are highly punctured and transmitted to the decoder on request.

ICUMT'09, October 12-14, 2009, St.-Petersburg, Russia.

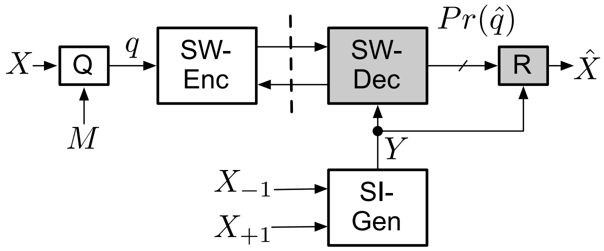


Figure 1. Basic DVC codec scheme

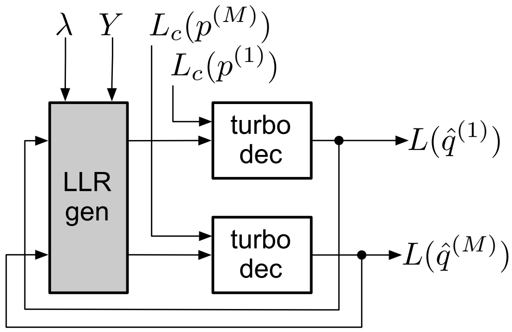


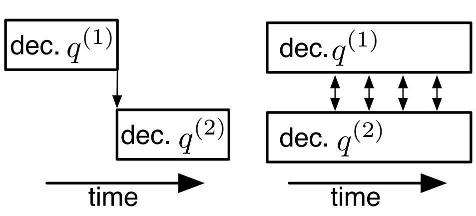
Figure 2. Slepian-Wolf decoder

B. Decoder

The decoder generates the side information Y by a temporal interpolation process using the previous X_{-1} and next X_{+1} frame. The side information is used in the reconstruction process and in the inter bit plane Slepian-Wolf decoder (fig. 2) to compute the Log-Likelihood-Ratio (LLR-value) for the turbo decoding. This is done for every bit plane b of a symbol q_i. The b-th bit plane of q_i is $q_i^{(b)}$. The Slepian-Wolf decoder can do the conventional plane by plane decoding (fig. 3(a)) and inter bit plane decoding (fig. 3(b), [5]). Whereas the inter bit plane mode decodes all bit planes in parallel and exchanges information in a bidirectional way. between the decoders.

In contrast to a conventional DVC decoder, the reconstruction algorithm takes the probability of each quantization symbol $Pr(\hat{q})$ and the side information Y to estimate $\hat{X}$ (reconstructed frame) of the original frame X.

The usage of the probability of every quantisation symbol



(a) plane by plane (b) inter bit plane

Figure 3. Decoding processes for $M = 2$, (a) plane by plane decoding, (b) inter bit plane decoding

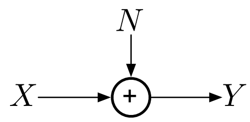


Figure 4. Correlation noise model

will result in an improved reconstruction quality compared to state-of-the-art reconstruction, in cases where the hard output decoding does not work (insufficient data rate).

C. Noise Model

The correlation between the original frame X and the side information Y is described by the noise model. This model is used to estimated the confidence information for SW decoding (LLR gen., figure 2). Furthermore the MMSE reconstruction and the proposed soft reconstruction incorporate the noise model for improved performance.

In this noise model the side information Y is a noisy version of the original frame X (figure 4). The noise N is describe by a Laplacian distribution (eq. 1).

$$p_N(n_i) \quad = \quad \frac{1}{2}\lambda e^{-\lambda|n_i|} \tag{1}$$

III. State-of-the-Art Reconstruction

Two reconstruction algorithms are widely used in distributed video coding. The reconstruction process converts the decoded quantization symbol $\hat{q}_i$ to an estimate $\hat{x}_i$ of the original value x_i. In distributed video coding the additional side information is used to increase the reconstruction quality (fig. 5(a)).

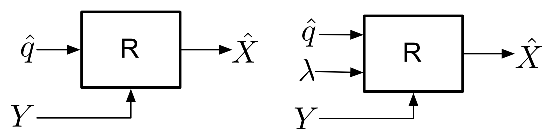
A. Conventional Reconstruction

The reconstruction algorithm presented in [7] (fig. 5(a)) is based on the decoded quantization symbol (quantization bin) and the side information Y. In case the side information is inside the reconstruction bin, the reconstructed value $\hat{x}_i$ is equal to the side information value y_i. Otherwise, the value will be clipped at the bin borders (z_j, z_{j+1}).

$$\hat{x}_i \quad = \quad \begin{cases} z_j & : \quad y_i < z_j \\ y_i & : \quad z_j < y_i < z_{j+1} \\ z_{j+1} & : \quad z_{j+1} < y_i \end{cases} \tag{2}$$

B. Minimum MSE Reconstruction

The enhanced version of the conventional reconstruction algorithm takes the correlation noise parameter λ (fig. 5(b)) into account, to generate an improved reconstruction quality. A minimum MSE (mean squared error) reconstruction algorithm is presented in [8]. The correlation between the side information Y and the original frame X is modeled as adaptive noise $Y = X + N$, where the noise N is Laplacian distributed with parameter λ. The optimal reconstruction



(a) (b)

Figure 5. State-of-the-art reconstruction ((a) `hard` and (b) `hard mse`)

value $\hat{x}_{opt}$ (minimum MSE) is given by the expectation value:

$$\hat{x}_{i,opt} = \mathrm{E}[x_i | x_i \in [z_j, z_{j+1}), y_i] \tag{3}$$

The advantage of the side information clipping algorithm [7] is a very low computational complexity. On the other hand the minimum MSE reconstruction algorithm [8] increases the PSNR but the computational complexity is slightly higher. In section V we compare our proposed reconstruction algorithm with the two state-of-the-art algorithms presented in this section.

IV. Proposed Soft Reconstruction

The Slepian-Wolf decoder described in the section II provides Log-Likelihood-Values (LLR) for each bit plane instead of only one bit (hard decision). The LLR values additionally include a confidence information. In contrast to the state-of-the-art reconstruction algorithms described in section III the proposed algorithm will also use the confidence information from the Slepian-Wolf decoder. The soft information is used to calculate the probability for each quantization symbol $\hat{q}$.

The confidence information can also be used in case of unsuccessful SW decoding. Thus it will provide increased reconstruction quality also in case of unsuccessful decoding, which occurs while semi and non feedback channel usage.

Furthermore, the reconstruction process takes the side information Y and the correlation noise parameter λ (fig. 6).

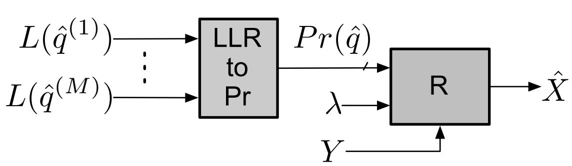


Figure 6. Soft reconstruction algorithm (`soft`)

At first the probability for each quantization symbol $\hat{q}$ is calculated from the LLR value (provided by SW decoder) for each bit plane by equation 4 and 5. The probabilities for the separate planes ($\Pr(\hat{q}_i^{(b)})$) are combined to the probability for the symbole ($\Pr(\hat{q}_i)$) as shown in equation 6.

$$\Pr(\hat{q}_i^{(b)} = 0) = \frac{e^{L(\hat{q}_i^{(b)})}}{1 + e^{L(\hat{q}_i^{(b)})}} \tag{4}$$

$$\Pr(\hat{q}_i^{(b)} = 1) = 1 - \Pr(\hat{q}_i^{(b)} = 0) \tag{5}$$

$$\Pr(\hat{q}_i = q) = \prod_b \Pr(\hat{q}_i^{(b)} = q^{(b)}) \tag{6}$$

As noted in section III the optimal reconstruction value $\hat{x}_{i,opt}$ in terms of minimum MSE is reached, when it is equal to the expectation value given in equation 7. The expectation value is constrained by the probability for the quantization symbols and the value of the side information y_i.

$$\hat{x}_{i,opt} = \mathrm{E}[x_i | \Pr(\hat{q}_i), y_i] \tag{7}$$

For this reason the pdf $p'(x_i)$ is computed by considering the quantization symbol probability. The pdf $p(x_i)$ (eq. 8) is independent of the quantization symbol probability.

$$p(x_i) = p_N(x_i - y_i) \tag{8}$$

The pdf $p(x_i)$ is weighted by the probability of each quantization symbol obtaining $p'(x_i)$. This is described in the equations 9 and 10, where $\mathrm{Q}(x_i)$ is the quantizer operation and gives the corresponding quantization symbol q_i for x_i.

$$p'(x_i) = w(x_i) p(x_i) \tag{9}$$

$$w(x_i) = \Pr(\hat{q}_i = q_i) \mid q_i = \mathrm{Q}(x_i) \tag{10}$$

The process of weighting the pdf for a two plane setup is shown in figure 7.

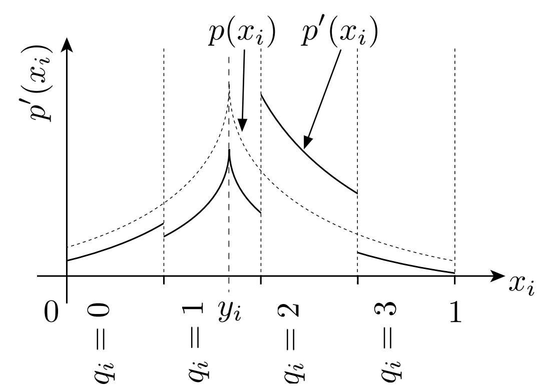


Figure 7. Weighted pdf $p'(x_i)$, ($M = 2$, quantizer parameter)

The pdf $p'(x)$ is used to calculate the optimal reconstruction value by equation 11.

$$\hat{x}_{i,opt} = \mathrm{E}[x_i | \Pr(\hat{q}_i), y_i] = \frac{\int_0^1 x_i \, p'(x_i) \, \mathrm{d}x_i}{\int_0^1 p'(x_i) \, \mathrm{d}x_i} \tag{11}$$

In the state-of-the-art reconstruction algorithms (sec. III), one quantization bin is selected and a minimum MSE reconstruction or a side information clipping reconstruction is performed. The weighting of the bins is therefore an extension of the bin selection (MMSE). Soft reconstruction converges to the MMSE reconstruction if one bin has a probability of 1. For this reason the proposed algorithm and the minimum MSE reconstruction will produce the same reconstruction values in case of very high LLR values (probability of one bin very close to 1).

Our proposed soft reconstruction algorithm uses the probability of each quantization symbol $\Pr(\hat{q}_i)$ to achieve an improved reconstruction quality. In case of weak or false decoding the distortion introduced by a false quantization symbol will be reduced. The reason for this behavior is based on the probability of a false decoded bit plane of a quantization symbol. It is close to $0,5$ which is similar to ignoring this plane.

V. Simulation Results

The simulation results are obtained by processing the sequences `coastguard` and `foreman` with QCIF resolution and 30fps. Furthermore, a keyframe distance of two is used (KWK), where lossless coded keyframes are assumed. The laplace parameter λ is estimated in an offline process.

Three potential scenarios for soft reconstruction are identified as described in the introduction (sec. I):

A) full feedback channel usage - hard bit plane by bit plane decoding, conventional reconstruction (state-of-the-art decoding, e.g. [3], [4])

B) semi feedback channel usage - last bit plane soft decoding, soft reconstruction (fixed sum rate)

C) non feedback channel usage - inter bit plane decoding, soft reconstruction (fixed rate per bit plane)

We compare the proposed soft reconstruction algorithm (eq. 7) with the two state-of-the-art reconstruction algorithms hard (side information clipping, eq. 2) and hard mse (minimum MSE, eq. 3).

A. Full Feedback Channel Usage

If the feedback channel is available, the hard bit plane by bit plane decoding (fig. 3(a)) is applied. This is the classical decoding procedure which is used in state-of-the-art Slepian-Wolf decoders (based on convolutional turbo codes). The first bit plane of the quantization symbol is decoded. Afterwards the next plane is decoded, whereas the already decoded planes are used to improve the decoding performance.

For the hard plane by plane decoding the proposed soft reconstruction algorithm cannot increase the reconstruction quality compared to the minimum MSE reconstruction algorithm (hard MSE). This is shown in figure 8. For high data rates ($rate > 60\,\mathrm{kbps}$) the decoding for the last plane was successful and there is no significant difference between the reconstruction quality of the minimum MSE and proposed soft reconstruction algorithm. Furthermore, there is still a gap in reconstruction quality between the hard (side information clipping) and the proposed soft reconstruction. The reason for this gap is the minimum MSE approach which is used by the soft reconstruction algorithm. Therefore the soft input values do not increase the quality here.

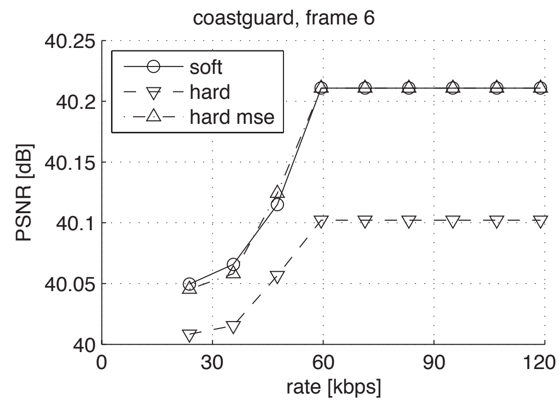


Figure 8. Last bit plane soft decoding, PSNR vs. rate for the LSB, `coastguard`, QCIF, 30 fps, $M = 2$

B. Semi Feedback Channel Usage

For semi feedback channel usage we use the last bit plane soft decoding scenario. The decoding is done plane by plane. In contrast to case A the sum rate is fixed. Hence, the rate for the last plane is not sufficient for successful decoding. For this reason the last bit plane is decoded partially (non error free decoding).

The simulation results (fig. 8 and 9) for this case show that an increased quality (PSNR) of up to $0,05\,\mathrm{dB}$ in comparision to the minimum MSE (hard mse) reconstruction is gained. Furthermore, the quality is increased by $0,15\,\mathrm{dB}$ in comparison to the known hard reconstruction algorithm. But there are also cases, where the quality of the soft reconstruction is worse than the minimum MSE reconstruction quality, as shown in figure 8. Therefore an adaptive selection of reconstruction algorithm is needed. But by now there is now algorithm which can decide which method is the best.

C. Non Feedback Channel Usage

Our inter bit plane decoding algorithm (fig. 3(b), [5]) and the proposed soft reconstruction is used for a fixed rate setup without feedback channel. This is also a very promising case to enable feedback channel free systems. The encoder sends a fixed rate for each plane and the decoder does inter bit plane decoding. It decodes all planes in parallel with information exchange and not in sequential order (plane-by-plane). Subsequently our soft reconstruction algorithm can do reconstruction also in case of non successful decoding for all or some bit planes. It is not mandatory for the encoder to send enough data for successful decoding of all planes. The

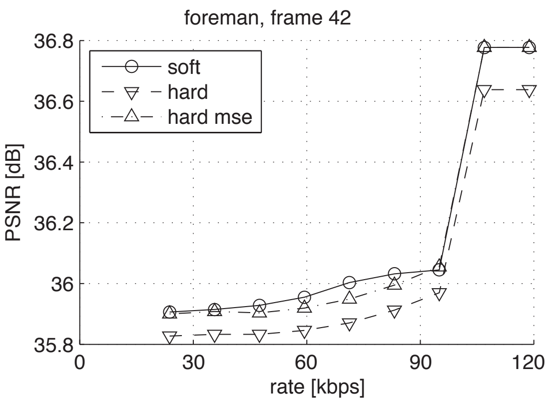


Figure 9. Last bit plane soft decoding, PSNR vs. rate for the LSB, `foreman`, QCIF, 30 fps, $M = 2$

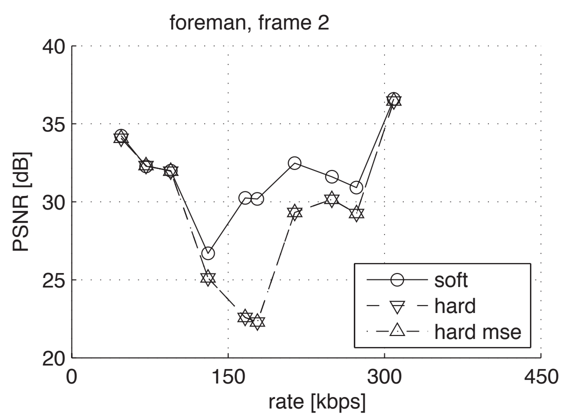


Figure 10. Inter plane soft, `foreman`

decoder can transfer data between the bit planes and can do reconstruction in case of non successful decoding.

A significantly increased reconstruction quality is achieved for the inter bit plane decoding case. Figure 10 shows the reconstruction quality over sum bit rate for one frame of the foreman sequence. Coding was done for two bit planes (M=2). The rate for the least significant bit plane (LSB) was twice the rate of the most significant bit plane (MSB). Thus a rate ratio of $rateRatio = R_{MSB}/R_{sum} = 1/3$ was used.. The reconstruction quality for `soft` reconstruction and minimum MSE (`hard mse`) reconstruction is slightly better than the reconstruction quality of the `hard` reconstruction algorithm in case of a successful decoding (highest rate). But if less bits are transmitted and thus the decoding was not successful, the soft reconstruction algorithm achieves a PSNR gain of up to $8,5\,\mathrm{dB}$.

VI. Conclusion

In this paper an algorithm for minimum MSE reconstruction was proposed that uses the probability for each quantization symbol $\Pr(\hat{q})$ instead of only a hard decision for one symbol. This reconstruction algorithm was tested for 3 different scenarios: full, semi and non feedback channel usage.

It is shown that the quality (PSNR) is increased by up to $8,5\,\mathrm{dB}$ for the non feedback channel decoding case. This offer a setup with a fixed data rate for each bit plane (no feedback channel). Furthermore, the reconstruction quality (PSNR) for the semi feedback channel usage scenario (fixed sum rate) is increased up to $0,15\,\mathrm{dB}$.

Further work on this topic will include complexity reduction of the proposed reconstruction algorithm and performance enhancements.

Acknowledgement

I would like to thank my colleagues for some helpful discussions and editorial remarks. Furthermore thanks to the reviewers for there detailed remarks.

References

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *Information Theory, IEEE Transactions on*, vol. 19, no. 4, pp. 471–480, Jul 1973.

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

[3] A. Aaron and B. Girod, "WYNER-ZIV VIDEO CODING WITH LOW ENCODER COMPLEXITY," in *Proc. Picture Coding Symposium (PCS)*, SF/USA, December 15-17 2004.

[4] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Ouaret, "THE DISCOVER CODEC: ARCHITECTURE, TECHNIQUES AND EVALUATION," in *Picture Coding Symposium (PCS)*, Lisboa, November 2007.

[5] R. Hänsel and E. Müller, "Simplifying the Rate Control Scheme for Distributed Video Coding by Flexible Slepian-Wolf Decoding," in *Proc.Pacific Rim Symposium on Image and Video Technology (PSIVT)*, 2009.

[6] C. Brites and F. Pereira, "Encoder Rate Control for Transform Domain Wyner-Ziv Video Coding," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2007.

[7] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv Coding of Motion Video," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2002.

[8] D. Kubasov, J. Nayak, and C. Guillemot, "Optimal Reconstruction in Wyner-Ziv Video Coding with Multiple Side Information," in *Proc. IEEE 9th Workshop on Multimedia Signal Processing MMSP 2007*, J. Nayak, Ed., 2007, pp. 183–186.