Robust Transmission of Image Sequences over Channels with Memory

(Robuste Übertragung von Bildsequenzen über Kanäle mit Fehlerbündelstrukturen)

Tilo Strutz and Andreas Ahrens

Abstract—**This contribution proposes a new framework of combined source-channel coding. It describes a wavelet-based coding scheme, which is especially designed for error-resilient transmissions and covers both compression of still images and the processing of video using predictive techniques. A well-tuned interface between source- and channel-coding stages enables a stable transmission synchronization using an unequal error protection (UEP) including multi-resolution features. The channel coding algorithms are based on Reed-Solomon (RS) codes and were developed for transmission over channels with memory.**

Keywords—**combined source-channel coding, video compression, channel with memory**

I. Introduction

MODERN communication technologies have a growing importance due to the increased desire for communication and information interchange. The transmission of video and still images especially plays a significant role. Unfortunately, transmission of raw image data requires enormous bandwidth and transmissions over real channels induce transmission errors. The first disadvantage can be combatted using efficient source coding methods. The elimination of redundancy and irrelevancy decreases the correlation and statistical dependencies between samples, but it also yields a high dependency of bits in the resulting bitstream and an increased vulnerability to transmission errors. The protection of data against errors is the task of channel coding. Shannon [12] proved that in the ideal source and channel coding could be solved independently and then combined for optimal performance. However, one may not forget that this independence holds only for restricted models of source and channel, and at the price of infinite delay and infinite computing power. In practical conditions, joint coding can yield significantly better performance. The idea behind is to spend intentionally more redundancy for source coding enabling a significant reduction of the expense of channel coding. The demands on channel coding can be further reduced, if the receiving side is error-tolerant. In this case, the transmission does not have to be 100% error-free.

T. Strutz and A. Ahrens are with the Institute for Communications and Information Electronics, University of Rostock, Germany.
E-mail: tilo.strutz@ntie.e-technik.uni-rostock.de
WWW: http://www-nt.e-technik.uni-rostock.de/~ts/

This contribution proposes a new source-channel coding framework. It does not concern with joint but combined coding, since there is a distinct demarcation between the source and channel-coding stages. Nevertheless, both stages are well coordinated.

The reduction of the overall-bitrate depends greatly on the efficiency of the source codec. For that reason Cosman et al [5] and Sherwood and Zeger [13] have chosen the SPIHT-codec from Said and Pearlman [11]. In contrast to the coding scheme proposed in this paper, both approaches aim to transmit of single images only. There is no mechanism to maintain a synchronization of image sequences.

The new source codec follows a multi-resolution concept facilitating a decoding at variable spatial resolutions. Besides, this property can be used for a source-encoder controlled UEP. Dependent upon the importance of the actual part of image information, the source encoder sends a weighting information to the channel encoder, which chooses a suitable grade of error protection. This strategy can also be used to adapt the protection strength to the channel state. The coding of single images is extended to the processing of image sequences based on intra- and inter-coding methods.

The simulations of the channel-coding techniques are based on a digital model of the HF channel, which takes into account the dependences between errors. Contrary to approaches assuming memoryless channels like additive white Gaussian noise (AWGN) channels [6, 14], slowly changing channels [3], or channel coding methods which maps channels with memory onto memoryless channels using fully bit-interleaving, our approach uses solely a symbol-interleaving (5-bit symbols) with a depth of 24 to break up the structures of bursty errors.

To avoid source-decoding failures, it is desirably to have available information about the reliability of single symbols. As a special feature of the new coding scheme, an interchange of so-called soft-out information is implemented. Under circumstances, the channel decoder is not able to make a successful correction of all errors. In these cases, the channel decoder delivers additional information about the location of these errors, that can be used for a significant improvement of image quality. The benefits of soft-out information were already shown in

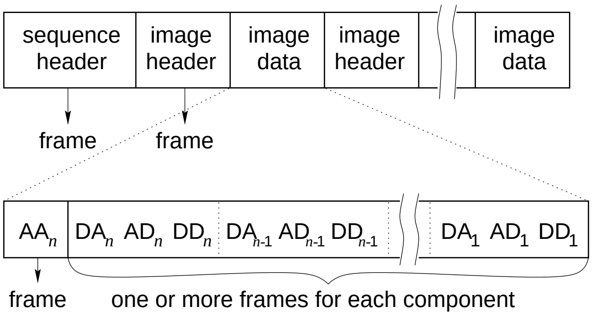

Fig. 1. Structure of source-coding output

[7] and [10].

The following section describes the cooperation of source and channel coding. After that, the source coding algorithm is explained. Section IV discusses several ways to a suitable UEP. Section V contains the description of the channel coding scheme. A presentation of the coding results finishes this contribution.

II. System Description

The aim of joint or combined source-channel coding is to minimize image distortions, while satisfying a constraint on the bitrate. Since channel coding increases the amount of data, a source-coding unit with high compression performance is necessary and one has to find a trade-off between both stages.

A wavelet-based coding scheme was developed, which provides an image transmission at high compression rates with spectral progression. **Fig. 1** depicts the structure of the source-coding output. The video bitstream starts with a sequence header containing global parameters like image format, type of wavelet transform, and subsampling format of color components. It is succeeded by an alternation of image headers and image data blocks. The image header determines a quantization parameter and the kind of image. It can be distinguished between intra (I), inter (P), and skipped images. The latter is used if the bitrate is to low to maintain a minimum image quality. Both I- and P-images are processed in the same way.

The image data, i.e. the wavelet coefficients (**Fig. 2**) to be encoded are ordered dependent on their importance. First, a coarse image information (low-pass band AA_n) is transmitted, which is followed by successive refinements (subbands containing higher frequencies).

The wavelet information is segmented subbandwise into data units, whose location in the wavelet domain is predetermined. Since image signals are non-stationary, the compression rates vary from unit to unit yielding different amounts of data. These units will be called in the following contribution as frames. They serve as interface to the channel coder. Each frame gets a header containing additional information. First, the channel encoder has to know how many number of bits a frame contains, and secondly, the source encoder informs about the importance of the frame data and includes a weighting factor. Dependent upon this information the channel encoder chooses a suitable error protection.

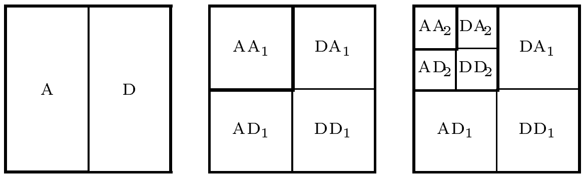

Fig. 2. Dyadic wavelet transformation

Furthermore, the frame-wise processing reduces the error sensitivity of the source decoder. The occurrence of errors in a frame does not imperil the decoding of succeeding frames. The developed channel-decoding algorithm is able to recognize, whether the error correction was successful or not and tells the source decoder which frame is error-free and which has noncorrectable errors, respectively. Besides, the channel decoder marks the location of these errors. This additional information (soft-out) can be used for an improvement of image quality.

A further task of the channel coder is transmission synchronization. Independent on the channel conditions, it has to guarantee a stable connection of transmitter and receiver. That means to the proposed scheme that the number of frames handed over to the channel encoder at the transmitting side has to be identically to the number of frames at the receiving side.

The source codec is able to process color images in YCbCr format and different subsampling modes. For simplification, the following investigations are limited to the processing of the luminance component Y.

III. Source Coding

A. Wavelet Transformation and Quantization

The source encoding is based on dyadic wavelet transformation. Each image is decomposed successively in subbands (**Fig. 2**). The number of decompositions is dependent on the image size. Assuming $size = \min(w, h)$ is the minimum of image width w and image height h, the number of decomposition n is calculated by

$$n = \left\lfloor \log_2 \left(\frac{size}{9} \right) \right\rfloor .$$

For the investigations in this paper, 9/7- [2] and 5/3- [4] wavelet filters are applied. Better objective performance in terms of peak-signal-to-noise ratio (PSNR) can be achieved with the 9/7 wavelet [15]. However, the 5/3 filters enable a faster implementation using the lifting scheme [16] and provide

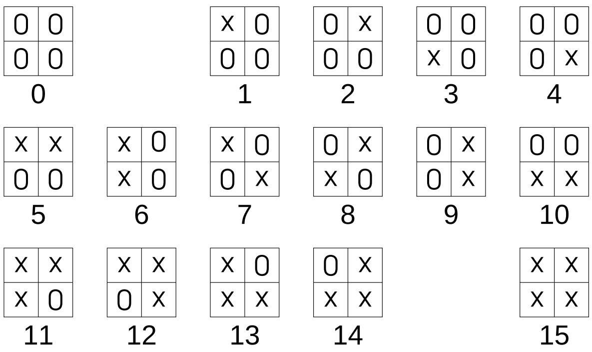

Fig. 3. Projection of 4 values onto new symbols

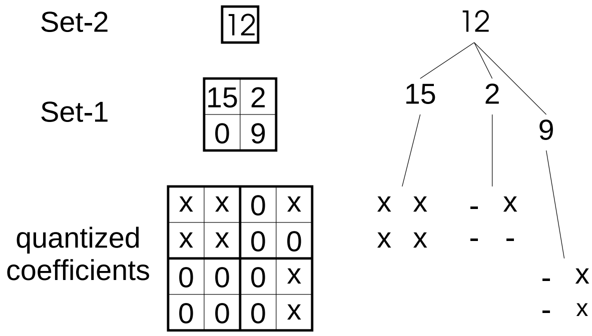

Fig. 4. Example for hierarchical precoding

the same or even better subjective image quality in most cases. The impulse responses of the analysis low-pass filters are scaled to provide a gain factor of 1.0.

The transformation is followed by scalar quantization. The basic width Δ of quantization intervals is determined by an external parameter. For every subband j containing three components DA_j, AD_j, and DD_j, the quantization intervals are given by

$$\Delta_j = 2^{n-j} \cdot \Delta \ .$$

Each transform coefficient d_i is divided by the actual Δ_j. Rounding towards zero yields an integer quantization index q_i. At the decoder stage, the coefficient $[d_i]_q$ is reconstructed from q_i

$$q_i = \mathrm{sgn}(d_i) \cdot \left\lfloor \frac{|d_i|}{\Delta_j} \right\rfloor \qquad [d_i]_q = q_i \cdot \Delta_j + \mathrm{sgn}(q_i) \cdot \frac{\Delta_j}{2}$$

with

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ +1 & \text{for } x > 0 \end{cases}$$

Following this quantization strategy, a near constant subjective image quality can be guaranteed with constant Δ.

It has to be mentioned that $\lfloor |x| \rfloor$ is a rounding towards zero and induces a deadzone forcing the output of values q_i equal to zero.

B. Coding Scheme

Dependent on Δ, the quantization results into more or less quantization indices equal to zero. This can be utilized by an efficient quadtree-based precoding algorithm. Adjacent values are projected onto a set of 16 different data symbols (see **Fig. 3**). 'x' marks values unequal to zero. A repetition of this mapping yields a hierarchical tree. **Fig. 4** shows an example with two steps. On the one hand, the projection can reduce the number of symbols (in the example from 16 to 11), on the other hand it separates the symbols into sets of different probability distributions. These are excellent conditions for entropy coding. The efficiency of this kind of quadtree coding is determined by the clustering of zeros. Remembering the graduation of Δ_j, it is advantageous to adapt the depth of the hierarchical trees to the subband level. As a good approximation the following connection could be found

subband j	1	2	3	4	5	...	n
projections	4	3	2	1	1	...	1

The coding tree is worked out using the 'deep first'-method. Starting with the set symbol at the top of a tree, each sub-tree is coded completely before the next sub-tree is processed. In the example the order would be: 12, 15, xxxx, 2, x, 9, xx.

The use of prefix codes is preferable to arithmetic coding in terms of faster implementation and higher robustness to bit errors. For this reason the symbols of each set and the remaining quantization indices are coded with fixed Huffman codes.

One has to distinguish between three symbol alphabets. The set at the top of a hierarchical tree contains 16 symbols, whereas the other set alphabets contain merely 15 different symbols. The 0-symbol does not belong to these alphabets due to the prediction from projection level to projection level. In each case four Huffman-code tables based on different distributions are at disposal. During the precoding process of a frame, the symbols of each set are counted. The algorithm chooses now that Huffman-code table that produces the lowest number of code bits, and informs the decoder about its choice. This adaptation to the image content is simple and very fast.

In contrast to the set alphabets, the alphabet for the quantization indices is open. The number of different values depends on several factors, like type of used wavelet and quantization, and is therefore unknown. To keep the Huffman-code tables small, the assumption is made that values from 1 to 15 are most probably. Including an extra-symbol (ESC), the basic code table contains 16 entries. Analog to the encoding of set symbols, the algorithm chooses one of four code tables. If a quantization index greater than 15 is found, the ESC symbol indicates an exception handling. The values are subdivided into six categories with decreasing probability (16–31, 32–63, 64–127, 128–255, 256–511, and 512–1023). The Huffman code belonging to the current category is transmitted followed by a fixed-length code determining the final value.

It is clear that if a transmission error occurs, the decoder would read a wrong Huffman code and lose the synchronization. That is why it is important to reset the encoding at well-determined points enabling a re-synchronization. For that reason the precoding and entropy coding processes described above are applied frame per frame. In the error case, the decoding of the current frame has to be stopped to avoid a reconstruction of false coefficient values, however, this does not influence the decoding of the next frame.

If the error checking of the channel decoder should fail, the source decoding stage is able to detect unmarked bit errors. The coding method for quantization indices leaves some redundancy in the data. Theoretically, a quantization index can show an absolute value from 1 to 1023. However, the domain of definition depends on Δ_j. Strictly speaking it can be found that $|q_i| < \lfloor 1024/\Delta_j \rfloor$. The source decoder checks each decoded value using that threshold. In the case of a wrong value the decoding of the current frame is stopped to prevent further image distortions.

C. Processing of the Approximation Signal

The encoding of the approximation signal AA_n follows an other strategy. This subband contains the most important image information, a false decoding yields to distinct distortions. For that reason fixed-length codes are used for the transmission of these quantization indices.

The difference to the quantization of the other subbands is that no deadzone is used.

$$q_i = \mathrm{sgn}(a_i) \cdot \left\lfloor \frac{|a_i|}{\Delta} + 0.5 \right\rfloor$$

To minimize the needed code length, the domain of definition of the quantized transform coefficients q_i is examined by

$$domain = max - min$$

with

$$max = \max_i(q_i) \qquad \text{and} \qquad min = \min_i(q_i).$$

The code length is now calculated by

$$l_q = \lfloor \log_2 (domain/\Delta) + 1 \rfloor .$$

All quantization indices $q'_i = q_i - min$ are transmitted using l_q bits and combined to one frame. As described above, the channel decoder informs the source decoder with a limited accuracy about the location of errors. Exploiting this information, the source decoder is able to reconstruct all received coefficients correctly and to interpolate the missing ones.

The reconstructed coefficients can be calculated by

$$[a_i]_q = (q'_i + min) \cdot \Delta$$

Both *domain* and *min* have to be transmitted as side information in the image header.

D. Motion Compensation and Bitrate Control

Video data contain in general a high temporal redundancy. If one removes this unnecessary data, the performance of the compression scheme can be improved significantly. For that purpose, an ordinary block matching with half-pixel accuracy is applied. To speed up the motion estimation, a multistep search with decreasing step sizes is used. For the investigations in this paper a block size of 16x16 and a search area of ± 12.5 in both directions were chosen.

Every transmitted image is also reconstructed at the encoder stage and regarded as reference image. The motion estimation calculates motion vectors, which describe the displacement between blocks of the next image to be coded and reference blocks. These vectors are necessary for the motion compensation and have to be transmitted. The x- and y-components of each vector are coded differentially using a simple prediction from the left neighbor to the current block. All block vectors belonging to the most left block column are coded without prediction. Analog to the wavelet coefficients, the coding is provided using Huffman codes. The motion information is inserted into the image-header frame.

If errors should occur during the transmission of motion vectors the decoding is still continued to collect as much information as possible. The missing vectors are set to $(0, 0)$ and the corresponding blocks perform an incorrect motion. In cases of little motion, this is not disturbing.

On noisy channels, predictive coding is rather dangerous, due to the propagation of errors. Therefore a periodical interrupt of the prediction by intracoded images is necessary to update the image content.

There are two modes for coding control: quality control and bitrate control. Quality control means, that every reconstructed image should have the same subjective quality (in error-free transmissions). This is simply performed by a constant strength of quantization (Δ = const.). The second mode tries to keep a constant bitrate. At the beginning of the sequence encoding, the mean number N_B of bytes per image is calculated. During the encoding process the encoder output fills a first-in-first-out (FIFO) buffer. Before the coding of the next image, N_B bytes are sent from the buffer to the channel. The control mechanism has to guarantee that neither an overflow nor an underflow of the buffer occurs. This is done by varying the quantization parameter. Should the coding of the current image consume to much bytes, then Δ is increased to attain a higher compression rate for the next image and vice versa.

If the capacity of the buffer is exceeded despite of that adaptation, the following image has to be skipped.

IV. Unequal Error Protection (UEP)

UEP means an adaptation of the protection strength to the importance of data parts. Information which is more important should be more protected than less important information. The channel coding stage does not know anything about image data. It processes merely anonymous data frames received from the source encoder. That is why the source encoder has to address the importance using a weighting factor.

As described above, the encoded data are segmented into frames. The source coder assigns weightings from 0 to 5 to each of them. It is obvious that the headers (one sequence and all image headers) are most important. They get a high weight that should guarantee an almost error-free transmission.

The spectral UEP is two-fold. In terms of objective quality (PSNR), all subbands have nearly the same importance, since the spreading of low-frequency coefficients errors is compensated by the higher incidence of high-frequency coefficients. However, image distortion induced by errors in low-frequency subbands are more visible than errors in subbands with higher frequencies generally. Thus, low-frequency subbands should be better protected. The approximation signal AA_n makes an exception. It is less vulnerable, since fix-length codes (see Sect. III-C) enable a continuous decoding without loss of synchronization.

In predictive video coding one has furthermore to decide between intra- and inter-coded images. The information of inter-coded images is less important for two reasons. First, less following images depends on that information, and secondly, the most changes of the image content from one image to the next image is realized by the motion compensation. The prediction error is just necessary for the updating of non-compensated image parts. If that information is absent, the subjective quality remains good in most cases, however, the objective quality can decrease dramatically.

V. Channel Coding

The channel encoder processes all frames produced by the source encoder. For synchronization purposes, the frames are segmented into data packets of the same length and a header packet is added (**Fig. 5**). This header contains a frame number (modulo 16), information about the type of error protection used for the data packets, and a start-of-image flag. If the recognition of a header packet fails on the receiver side, the channel decoder detects this by checking the successive frame numbers and includes a dummy frame to keep the frame synchronization steady. Should this mechanism fail, then the source decoder will wait for the next frame with an activated start-of-image flag.

The channel coding is based on a concatenation of inner coding for error detection and correction and outer coding for error detection. **Fig. 6** depicts a single-level concatenated coding system in which an outer code and an inner code are combined. If both code strategies are chosen properly, the concatenated coding system achieves high performance with moderate decoding complexity. The investigations are based on a hard-decision inner-code decoding and a hard-decision outer-code decoding with soft-output.

The error structures on channels with memory require coding strategies, which have a high robustness against bursty errors. Therefore, the inner decoding stage uses codes employing over $\mathrm{GF}(2^q)$ exclusively [8]. The inner code is a $(2^q-1, 2^q-2\cdot t_k-1, t_k)_q$ Reed-Solomon (RS) code and is able to correct any pattern of t_k or fewer symbol errors. **Tab. I** contains an overview of weighting factors (delivered by the source encoder), corresponding RS codes, and the resulting code rates R_c. The application of codes with different error correction capacities enables a suitable UEP.

The performance of the RS codes is shown in **Fig. 7** and **Fig. 8**. The channel simulations are

TABLE I

RS codes for unequal error protection

weighting	RS code	R_c
0	$(31,29,1)_5$	0.935
1	$(31,27,2)_5$	0.871
2	$(31,23,4)_5$	0.742
3	$(31,19,6)_5$	0.613
4	$(31,15,8)_5$	0.484
5	$(31,11,10)_5$	0.355
header packet	$(31,11,10)_5$	0.355

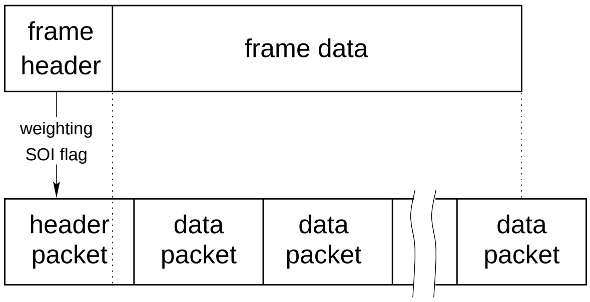

Fig. 5. Separation of frame data into packets

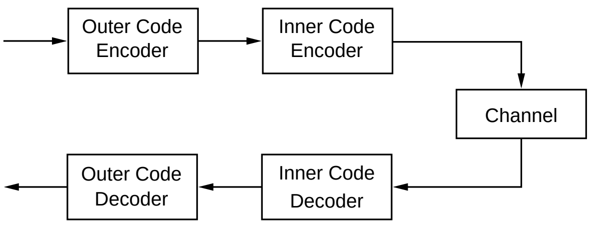

Fig. 6. Single-level concatenated channel coding system

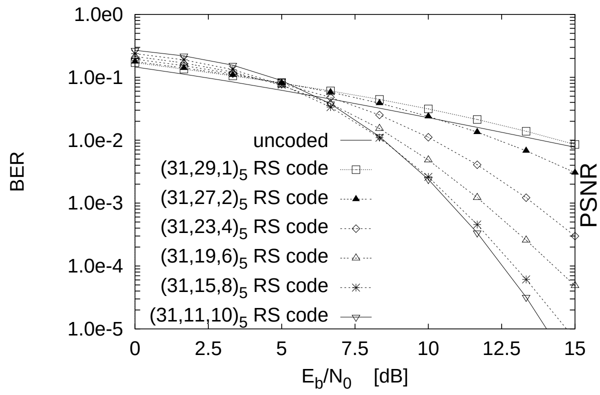

Fig. 7. Performance of the inner coding stage on channels with an error concentration value of $(1-\alpha)=0.2$ and a symbol interleaving with a depth of 24

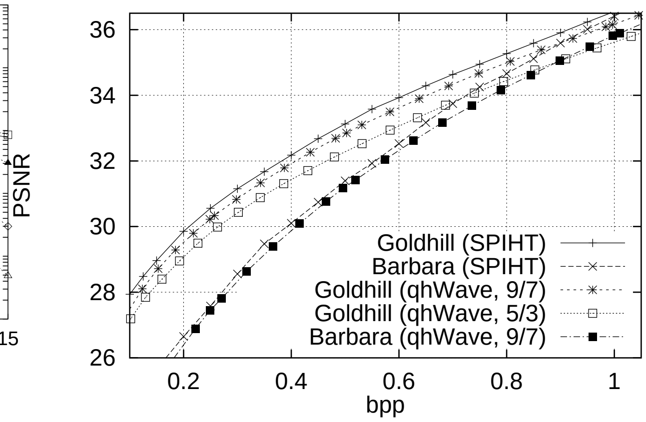

Fig. 9. Performance of source coding

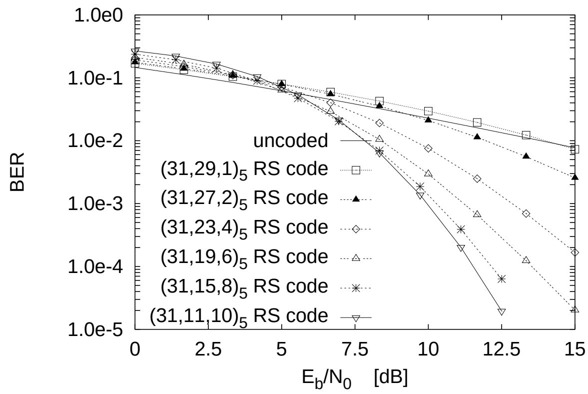

Fig. 8. Performance of the inner coding stage on channels with an error concentration value of $(1-\alpha)=0.3$ and a symbol interleaving with a depth of 24

based on the channel model proposed in [1]. This model depends on two parameters: the channel-signal-to-noise-ratio (CSNR) and a value $0 \leq (1-\alpha) \leq 0.5$ describing the error concentration. $(1-\alpha)=0$ is related to uniform-distributed errors (channel without memory). The connection of bit-error rate (BER) and CSNR $= R_c \cdot E_b/N_0$ is given by

$$\mathrm{BER} = \frac{1}{2} \cdot \left(1 - \sqrt{1 - \frac{1}{1 + R_c \cdot E_b/N_0}}\right) \qquad [9]$$

A typical approach to error correction for bursty channels is to introduce a bit interleaver which spreads out adjacent bits by the interleaver depth. The goal is to produce a memoryless channel and then to use conventional error correction strategies (e.g. convolutional coding). For channels with high error concentration, fully bit interleaving is impractical, since the realization of the required interleaver depth would be to expensive.

Cyclic redundancy codes (CRC) of 8 bits realize the outer coding for each data packet. An important property of the outer code is that the CRC provides a high probability indication of the decoding success or failure of a packet. Due to the importance of header packets, they are protected with an applied CRC-12 and a fixed RS code with high error-correction capacity. To improve the security of header-packet detection a special 8-bit pattern (1111 1111) is included. The probability is very low that this bit sequence occurs in the source encoder output. Thus, a tree-level verification is performed to decide, whether a packet is a header packet or not.

The sensibility of the source decoder to transmission errors must be taken into account in the system design. After receiving data, the channel decoder tries to correct the identified bit-errors and resamples the encoded image information into frames. The frame headers contain now information about the correctness of data packets. A packet-wise error information is advantageous, because it is not necessary to omit the whole frame. The information up to the occurrence of the first corrupted data packet of subbands and all error-free packets of the approximation signal can be used for a better image reconstruction.

The overall size for each packet is 155 bits and is determined by the RS codes. Deducting the synchronization and error-correcting information, 27 bits remain in the header packet for transmission of data bits. The number of data bits in the data packets depends on the used RS code.

The generation of the soft-out information includes two steps. First, an error information is determined by the analysis of the syndrome $S_i(x)$ and is passed to the outer decoding stage. In the case of $S_i(x) \neq 0$, errors were recognized and a Chien search decides, whether it is an uncorrectable or a correctable error pattern. Secondly, the outer decoding stage delivers statements about the faultlessness of the received packet.

VI. Results

A. Source Coding

The efficiency of the source codec is depicted in **Fig. 9** for two test images ('Barbara' and 'Goldhill',

TABLE II
CODING PERFORMANCE FOR VIDEO CODING (200 IMAGES)

sequence	bits per pixel (bpp)	mean PSNR
foreman	0.25	28.70
foreman	0.50	32.40
carphone	0.25	31.85
carphone	0.50	36.16
salesman	0.25	33.69
salesman	0.50	41.55

TABLE III
CHANNEL BITRATES FOR COMBINED SOURCE-CHANNEL CODING AND EQUAL ERROR PROTECTION (FOREMAN SEQUENCE, 200 IMAGES)

weighting	bpp on channel	code rate
-	0.500	1.000
0	0.681	0.735
1	0.722	0.692
2	0.834	0.600
3	0.995	0.502
4	1.254	0.340
5	1.730	0.289

each with 512x512 pixels). The PSNR is calculated by

$$\text{PSNR [dB]} = 10 \cdot \log_{10} \frac{255^2}{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}.$$

x_i denotes the greyvalues of the original image, whereas $\hat{x}_i$ denotes the reconstructed values. Though the new coding algorithm (called as 'qh-Wave') is very simple, nearly the same performance can be attained compared to the SPIHT codec of Said and Pearlman [11].

Tab. II shows some results for the coding of different image sequences (Y component) in QCIF format (176x144 pixels) with an image rate of 25 Hz.

B. Source-Channel Coding

First, the overhead and the additional redundancy for error protection and detection are investigated. The test sequence 'foreman' (176 x 144 Pixel, luminance component, 200 images) was encoded with bitrate control at 0.5 bpp. The minimum distance between adjacent intra images was 5 (one I-image and four P-images per group) for this and the following simulations if not stated differently. After that, a channel encoding with an equal protection for each data packet was performed. The header packets were encoded using the strongest RS code (weighting 5) as described above. **Tab. III** contains the resulting code rates and the bit-per-pixel values. It has to mentioned that the code

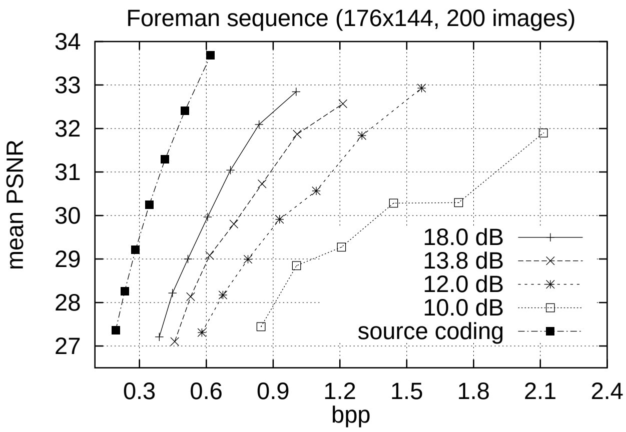

Fig. 10. Performance of combined source-channel coding under different channel conditions and adapted protection strength

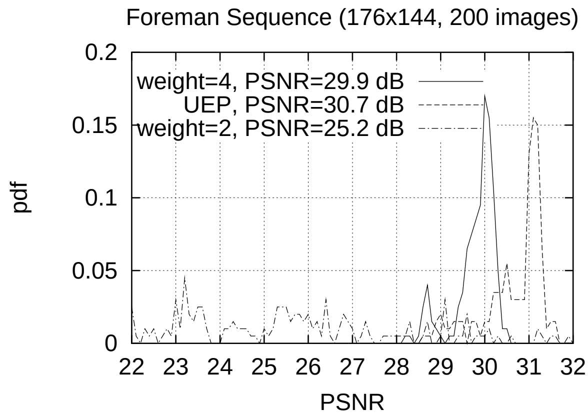

Fig. 11. Distribution of quality of reconstructed images at 0.7 bpp, CSNR= 13.8 dB, $(\alpha - 1) = 0.2$, and different protection strength

rates also include all mechanisms for synchronization of image sequences.

Fig. 10 depicts the performance of the combined source-channel coding under different channel conditions in terms of channel-signal-to-noise-ratio (CSNR) with an error concentration of $(1-\alpha) = 0.2$. The most left curve gives an impression, which image quality can be attained with pure source coding without transmission errors. In combined coding, the protection strengths were chosen properly to attain highest PSNR values. In practice, the channel conditions could be examined using a feedback channel. The resulting PSNR curves would be smoother, if one would increase the sample size. However, 200 images should be enough to get a sufficient impression.

Fig. 11 gives an example for the distribution of objective image quality (mean PSNR), if the protection strength is varied. A high protection for all frames (weight=4) results to a distinct peak at 30.0 dB. Just a few images are corrupted. If one performs an unequal protection, the weightings can

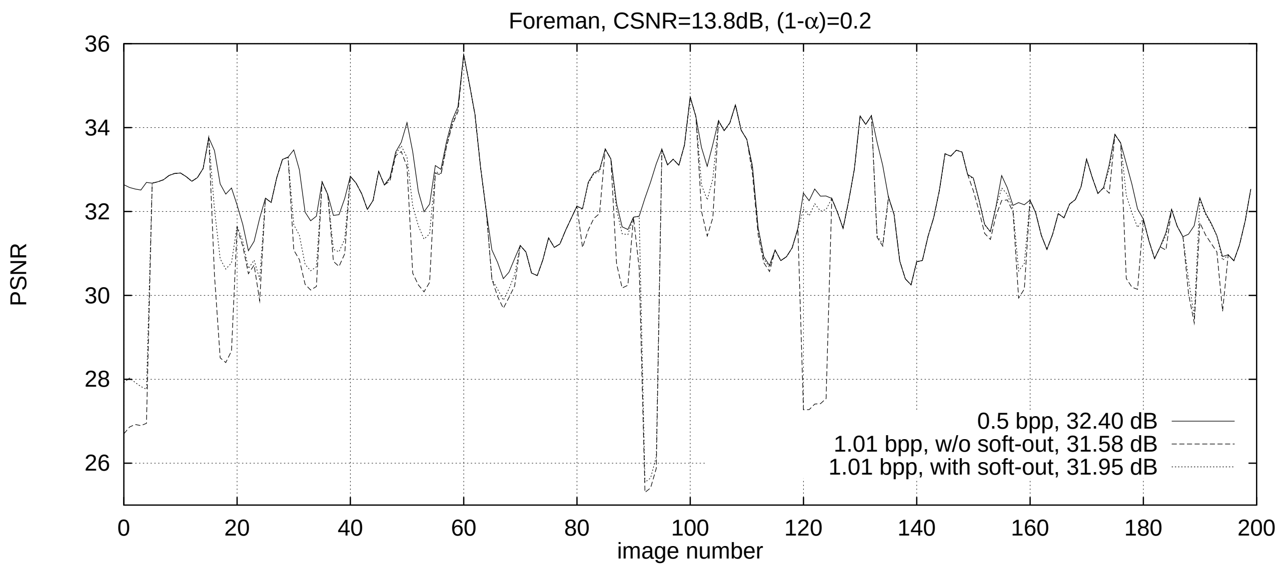

Fig. 12. Example of combined source-channel coding with and without utilization of soft-out information

be reduced for some frames. Less bits are necessary for channel coding and can be used now to improve the source coding. The peak moves to a higher image quality. However, the distribution function becomes more flat. If the protection is to low, many images are distorted by channel noise, the image quality falls significantly.

Dependent on the demands on the transmission, the bitrate can be decreased, if not the objective but just the subjective quality is important. For that, the UEP could be varied like discussed in Section IV.

The progress of single PSNR values can be seen in **Fig. 12**. The source coding takes 0.5 bpp. Using the same UEP as in Figure 10 at CSNR=13.8 dB, the bitrate increases to 1.01 bpp on channel. The corresponding curves show the changes in PSNR caused by channel noise. If the soft-out mechanism is disabled, the decoding omit each erroneous frame. Using soft-out information, the decoding stops only if the first faulty data packet of a frame is detected. Hence, the image quality can be improved. The influence of I-images on P-images is clearly to be seen for images 0 – 4 and 120 – 124. If the intra-coded image is corrupted, the distortions propagates in the following P-images. If the decoding of a P-image fails, the collapse of objective quality is shorter (images 92–94).

The codec performance in dependence on the error concentration $(1-\alpha)$ is investigated in **Fig. 13**. The UEP was optimized for $(1-\alpha)=0.2$. If the error concentration increases, the protection performs better. Otherwise, if the errors are more equipartitioned, the performance get worse.

C. Error Propagation in Predictive Video Coding

In predictive video coding, one has to find a trade-off between higher compression rates by

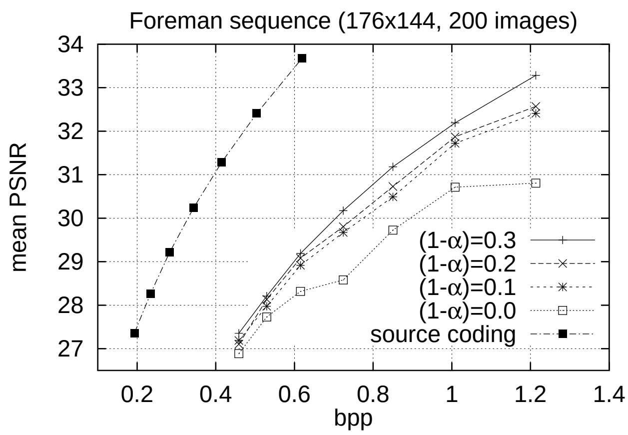

Fig. 13. Performance of combined source-channel coding depending on error concentration at CSNR=13.8 dB

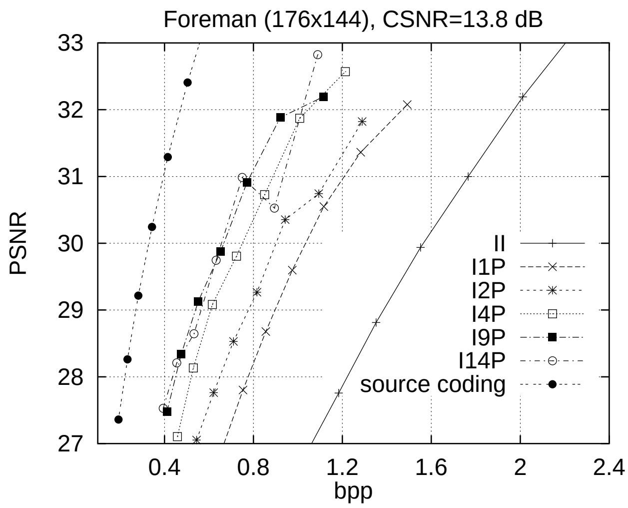

Fig. 14. Trade-off between distances between adjacent I-images and error propagation (CSNR=13.8 dB)

means of longer distances between adjacent I-images and the resulting decrease in terms of PSNR caused by error propagation. **Fig. 14** shows some examples. The notion 'IxP' means that one I-images is followed by x P-images in each group. First, one can see that the introduction of only one P-image between I-images (I1P) can already decrease the bitrate significantly in comparison to pure intra coding (II). If the group is longer than about 5, a stronger protection is required to avoid image distortions that could propagate. However, this would be in contradiction to the idea of error-tolerant transmission. Should the protection fail despite stronger codes, the errors can propagate through the whole image group and decrease the mean PSNR significantly (curve 'I14P' at 0.9 bpp).

VII. Summary

With this contribution a new source-channel-coding framework has been presented. The channel-coding stage was designed specifically for transmissions over channels with memory and shows a high robustness against bursty errors. The simulations are based on a new channel model which was introduced, too. For source coding, a new, simple coding algorithm based on wavelet transformation, uniform quantization, quadtree-like precoding, and Huffman coding was developed.

The simulations have shown that this new method of combined source-channel coding guarantees a stable synchronization for image-sequence transmission at relative low bitrates and high image qualities even at high error rates on channel. The transmission performance was improved by predictive coding up to an intra-image distance of about 5, and by the utilization of soft-out information.

Appendix

I. CRC Codes

The outer codes use a 12-bit CRC defined by the polynomial $x^{12} + x^{11} + x^3 + x^2 + x^1 + 1$ or an 8-bit CRC defined by $x^8 + x^2 + x^1 + 1$, respectively.

References

[1] Ahrens, A.: *A digital Model of the HF-Channel and its Application.* [in German], Frequenz, Fachverlag Schiele & Schön GmbH, Berlin-Kreuzberg, Vol.53, No.1–2, 1999, pp.7–11.

[2] M. Antonini; M. Barlaud; P. Mathieu; I. Daubechies: *Image Coding Using Wavelet Transform.* IEEE Transactions on Image Processing, Vol.1, No.2, April 1992, pp.205-220

[3] G. Cheung; A. Zakhor: *Joint Source/Channel Coding of Scalable Video over Noisy Channels.* Proceedings of ICIP'96, International Conference on Image Processing, Lausanne, Switzerland, September 16–19, 1996, CDROM

[4] A. Cohen; I. Daubechies; J.-C. Feauveau: *Biorthogonal Bases of Compactly Supported Wavelets.* Communications on Pure and Applied Mathematics, Vol.45, pp.485-560, 1992

[5] P.C. Cosman; J.K. Rogers; P.-G. Sherwood; K. Zeger: *Combined Forward Error Control and Packetized Zerotree Wavelet Encoding for Transmission of Images Over Varying Channels.* IEEE Transactions on Image Processing, revised Dec.1998

[6] A. Fuldseth; T.A. Ramstadt: *Channel-Optimized Subband Video Coding for Channels With a Power Constraint.* Proceedings of ICIP'97, International Conference on Image Processing, Santa Barbara, CA, October 26-29, 1997, Vol.3, pp.428–431

[7] Joachim Hagenauer: *Soft-In/Soft-Out.* Third International Workshop on Digital Signal Processing Techniques applied to Space Communications, Noorwijk, pp. 7.1-7.15, 1992.

[8] R.W. Johnson, M.B. Jorgensen and K.W. Moreland: *Error Correction Coding for single Tone HF-Transmission.* Seventh International Conference on HF Radio Systems and Techniques, Nottingham, pp. 80-84, 1997.

[9] Proakis, J. G.: *Digital Communications.* McGraw-Hill, New York, Third Edition, 1995

[10] B. Rislow, T. Maseng and O. Trandem: *Soft Information in Concatenated Codes.* IEEE Transaction on Communications, pp. 284-286, 1996.

[11] A. Said; W.A. Pearlman: *A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees.* IEEE Transactions on Circuits and Systems for Video Technology, Vol.6, June 1996, pp.243-250

[12] C.E. Shannon: *Collected Papers*, edited by N.J.A. Sloane and A. Wyner, IEEE Press, NY, 1993

[13] Sherwood, P.-G.; Zeger, K..: *Error Protection for Progressive Image Transmission over Memoryless and Fading Channels.* Proceedings of ICIP'98, International Conference on Image Processing, Chicago, Illinois, October 1998

[14] M. Srinivasan; R. Chellappa; Ph. Burlina: *Adaptive Source-Channel Subband Video Coding for Wireless Channels.* Proceedings of IEEE 1st Workshop on Multimedia Signal Processing, June 23–25, 1997, Princeton, New Jersey, USA, pp.407–412

[15] Strutz, T.: *Untersuchungen zur skalierbaren Kompression von Bildsequenzen bei niedrigen Bitraten unter Verwendung der Wavelet-Transformation.* Dissertation, Institut für Nachrichtentechnik und Informationselektronik, Universität Rostock, 1997, published by Shaker Verlag GmbH Aachen, ISBN 3-8265-3600-2, 1998

[16] W. Sweldens: *The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Construction.* Proceedings of SPIE - The International Society for Optical Engineering, Vol.2569, San Diego, CA, USA, July 1995, pp.68-79