Error-Resilient Transmission of Compressed Images over Very Noisy Channels Using Soft-Input Source Decoding

Jörg Kliewer and Norbert Görtz
University of Kiel
Institute for Circuits and Systems Theory
Kaiserstr. 2, 24143 Kiel, Germany
Phone: +49 (431) 77572-406, Fax: +49 (431) 77572-403
E-Mail: {joerg.kliewer, norbert.goertz}@ieee.org

Abstract

In this contribution we utilize an optimal estimation approach for error-resilient transmission of compressed images over AWGN channels. In contrast to other methods, we mainly rely on the implicit residual redundancy inherent in the subband images and the bit-reliability information at the channel output for error protection. It is shown that by including a-priori knowledge about the spatial correlations in the subband images, we obtain a highly increased quality of the reconstructed image. This holds especially for very noisy channels and a mean-square estimation of the subband source coefficients.

1. Introduction

Joint source-channel coding for error-resilient image transmission has recently gained strong interest in the literature, since the classical separation of source and channel coding seems not to be justified especially for finite-length source signals. A certain subset of these methods is denoted by joint source-channel *de*coding, where the residual redundancy of the source signal is exploited in order to improve the results from the channel decoding stage [3–5, 9].

In the following we present an error-concealment approach for image transmission over AWGN-channels, which is based on the source-decoding scheme proposed in [3]. This method optimally estimates the output signal at the decoder depending on bit-reliability information at the channel output and on the source statistics. In contrast to many other recent approaches in the literature (e.g. in [7, 10]) we do not use any channel codes for error protection, except for transmitting the error-sensitive values of the bit-allocation. Instead, by only using simple scalar quantizers for the subband images, we deliberately leave some residual redundancy in the source-encoded bitstream, which then is exploited for error-concealment at the receiver.

2. Soft-Input Source Decoding by Optimal Estimation

2.1. Transmission system

Let us consider the simple model in Fig. 1. The source signal vector $\boldsymbol{U} = [U_0, U_1, \ldots, U_k, U_{k+1}, \ldots]$ with k denoting the sample index, is generated by an one-dimensional scan of a two-dimensional subband image (see Section 3). Due to complexity and delay constraints the

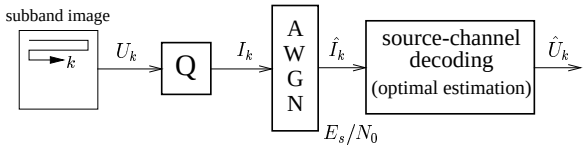

Figure 1. Simple model of the transmission system

source correlation can never be completely removed by subband decomposition. Therefore, we can generally assume that the quantized indices I_k, which can be represented with a limited number of N bits, are more or less correlated. In the following, the correlation of the indices is modeled by a first-order stationary Markov-process, where the Markov-model is described by the index-transition probabilities $P(I_k = \lambda \,|\, I_{k-1} = \mu)$, $\lambda, \mu = 0, 1, \ldots, 2^N - 1$, which are assumed to be known in advance or to be approximated at the decoder. In order to allow a more compact notation we will also use the expression $I_k^{(\lambda)}$ instead of $I_k = \lambda$ in the following.

The index I_k is then transmitted over an AWGN-channel, where for a single bit this scenario is depicted in Fig. 2.

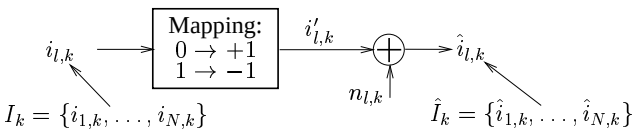

Figure 2. Transmission model for one bit $i_{l,k}$

After a mapping to bipolar bits we add a white Gaussian noise sample $n_{l,k}$. The noise samples have zero mean and a variance of $\sigma_n^2 = \frac{N_0}{2E_s}$, where E_s denotes the energy used to transmit each bit and N_0 represents the one-sided power spectral density of the channel noise. For the conditional p.d.f. $p(\hat{i}_{l,k}|i_{l,k})$ we then obtain

$$p\big(\hat{i}_{l,k} \,|\, i_{l,k}\big) = \frac{e^{-\frac{1}{2\sigma_n^2}(\hat{i}_{l,k}-1)^2}}{\sqrt{2\pi}\sigma_n} \; e^{-\frac{1}{2\sigma_n^2}\cdot 4\cdot \hat{i}_{l,k}\cdot i_{l,k}} \qquad (1)$$

with $i_{l,k} \in \{0,1\}$ and $\hat{i}_{l,k} \in \mathbb{R}$. Using this, the conditional p.d.f. for the overall received soft-bit vector $\hat{I}_k$, when a certain $I_k \in \{0,1\}^N$ is given, can be written as

$$p\big(\hat{I}_k \,|\, I_k\big) = \prod_{l=1}^{N} p\big(\hat{i}_{l,k} \,|\, i_{l,k}\big). \qquad (2)$$

After that, this channel-based knowledge and the residual source correlation is exploited in the source-channel decoding block in Fig. 1 in order to maximize the signal-to-noise ratio (SNR) at the decoder output.

2.2. A-posteriori probabilities

The a-posteriori probabilities (APPs)

$$P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k) = P(I_k = \lambda \,|\, \hat{I}_0, \ldots, \hat{I}_{k-1}, \hat{I}_k)$$

at sample index k represent the probability that the index $I_k = \lambda$ has been transmitted, given all already received soft-bit vectors $\hat{\boldsymbol{I}}_0^k = [\hat{I}_0, \ldots, \hat{I}_{k-1}, \hat{I}_k]$ at the decoder. The value of $P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k)$ can be regarded as reliability information for the hypothesis $I_k^{(\lambda)}$, and will be used for the optimal estimation of the reconstructed subband coefficients at the decoder output (see Section 2.3). Given the index transition probabilities of the Markov-model and the conditional p.d.f.s in (1) and (2), resp., the a-posteriori probability can be recursively calculated as [3, 6]

$$P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k) = c_k \cdot \overbrace{p\big(\hat{I}_k \,|\, I_k = \lambda\big)}^{\text{channel term}} \cdot \cdot \sum_{\mu=0}^{2^N-1} \underbrace{P\big(I_k = \lambda \,|\, I_{k-1} = \mu\big)}_{\text{trans. probabilities, Markov source}} \; P(I_{k-1}^{(\mu)} \,|\, \hat{\boldsymbol{I}}_0^{k-1}). \qquad (3)$$

for $\lambda = 0, 1, \ldots, 2^N - 1$. The factor c_k normalizes the sum in (3) such that $P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k)$ is a true probability, and the initialization of (3) can be carried out with the unconditional source index probabilities $P(I_k = \lambda)$.

When we now assume that we have already knowledge of L_k future samples, where $L_k = M - k$ and M denoting the overall length of the source-vector $\boldsymbol{U}$, this additional information can be exploited for increasing the reliability of the APPs. In the general case these probabilities are given as $P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^M)$, which can in principle be calculated with the BCJR-algorithm [2]. Unfortunately, application of this algorithm leads to high computational costs especially for large N and M. Therefore, in the following we only regard the special case for $L_k = 1$. The corresponding APPs $P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^{k+1})$ can be expressed by known quantities, which are the APPs from (3), the channel p.d.f. in (2), and the transition probabilities $P(I_k^{(\lambda)} \,|\, I_{k-1}^{(\mu)})$ and $P(I_{k+1}^{(\lambda)} \,|\, I_k^{(\mu)})$ for the corresponding Markov models, respectively.

Using the Bayes rule, the Markov property of the source, and the memoryless property of the AWGN-channel we first write the APPs as

$$\begin{aligned} P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^{k+1}) &= P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k, \hat{I}_{k+1}) \\ &= c'_k \cdot p(\hat{I}_{k+1} \,|\, I_k^{(\lambda)}) \cdot P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k), \end{aligned} \qquad (4)$$

where the factor c'_k again ensures that the APPs on the left-hand side are true probabilities. The term $p(\hat{I}_{k+1} \,|\, I_k^{(\lambda)})$ in (4) can be further decomposed for a memoryless channel and by using the Bayes rule as

$$p(\hat{I}_{k+1} \,|\, I_k^{(\lambda)}) = \sum_{\mu=0}^{2^N-1} p(\hat{I}_{k+1} \,|\, I_{k+1}^{(\mu)}) \, P(I_{k+1}^{(\mu)} \,|\, I_k^{(\lambda)}). \qquad (5)$$

Inserting (5) into (4) leads to the desired result for the APPs:

$$\begin{aligned} P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^{k+1}) = c'_k \cdot P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k) \cdot \\ \cdot \sum_{\mu=0}^{2^N-1} p(\hat{I}_{k+1} \,|\, I_{k+1}^{(\mu)}) \, P(I_{k+1}^{(\mu)} \,|\, I_k^{(\lambda)}). \end{aligned} \qquad (6)$$

In order to utilize the additional knowledge of one future received sample at time-instant $k+1$ the APPs from (3) are weighted with a sum-term containing the index transition probabilities and the channel term for the index $\hat{I}_{k+1}$.

2.3. Optimal estimation

The APPs from equation (3) and (6), resp., can now be used to optimally estimate the vector of subband coefficients $\boldsymbol{U}$ in such a way that the resulting vector $\hat{\boldsymbol{U}} = [\hat{U}_0, \hat{U}_1, \ldots, \hat{U}_k, \hat{U}_{k+1}, \ldots]$ after the estimation minimizes the value of a suitable "overall distortion". For the sake of clarity we will use $P_{AP}(\lambda, k)$ as a general expression for the APPs $P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^k)$ in (3) and $P(I_k^{(\lambda)} \,|\, \hat{\boldsymbol{I}}_0^{k+1})$ in (6), resp., in the following descriptions of the estimators.

(a) Maximum a-posteriori estimation. The maximum a-posteriori (MAP) estimator minimizes the decoding error probability in the reconstructed values, where the MAP-estimated indices $\hat{U}_k^{(\mathrm{MAP})}$ are obtained as

$$\hat{U}_k^{(\mathrm{MAP})} = U_q(I_k = \lambda_{\mathrm{map}}), \quad P_{AP}(\lambda_{\mathrm{map}}, k) \geq P_{AP}(\lambda, k). \tag{7}$$

Here, $U_q(\lambda)$ denotes the entry of the quantization table belonging to the index λ. Note that classical maximum likelihood (ML) decoding, which in the absence of channel codes is only a simple hard decision decoding, represents a special case of the MAP-estimator. This case is characterized by independent and equally distributed source indices with $P(I_k = \lambda \mid I_{k-1} = \mu) = P(I_k = \lambda) = 1/2^N$ for all μ, λ.

(b) Mean-squares estimation. The mean-square estimator (MS) directly corresponds to the demand for maximal SNR. Thus, it may be well suited when dealing with "waveform-like" signals, as for example images. The mean-square-estimated values $\hat{U}_k^{(\mathrm{MS})}$ are derived according to

$$\hat{U}_k^{(\mathrm{MS})} = \sum_{\lambda=0}^{2^N-1} U_q(\lambda) \cdot P_{AP}(\lambda, k), \tag{8}$$

where each $U_q(\lambda)$ is weighted with the APP corresponding to the index λ prior to summation. Note that this estimator has a "graceful degradation" property. This can be best seen from the worst case example for a severely distorted channel, where all $P_{AP}(\lambda, k)$ are almost equally distributed. In this case the estimated value $\hat{U}_k^{(\mathrm{MS})}$ takes on the mean of $U_q(\lambda)$ for all λ.

3. Application to Error-Resilient Transmission of Compressed Images

The experimental image codec used in our approach is based on a separable two-dimensional wavelet-based octave filter bank with L levels, where the well known 9-7 filters [1] are used as subband filters. For every subband image the transmission and estimation of the subband source coefficients U_k in Fig. 1 is carried out independently.

In order to obtain the vector-elements U_k from the two-dimensional subband image we employ the scanning methods depicted in Fig. 3, where for the sake of clarity only the first level of the decomposition is shown. As we can see, the scanning is carried out in a meander-type style. Thus, the spatial correlations inherent in the subband images are projected onto correlations between (several) adjacent vector-elements U_k for all k in the resulting one-dimensional vector $\boldsymbol{U}$. Note that this does not necessarily hold for the vectors created by simply concatenating all rows or columns of the individual subband images. Furthermore, the scanning orientation in Fig. 3 is chosen such that we try to maximize the correlation between the vector-elements. In the HL- and LH-subband this corresponds with the orientation

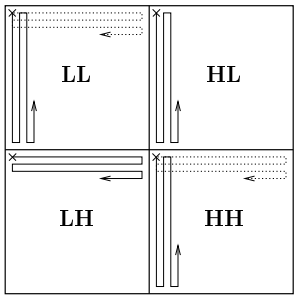

Figure 3. Scanning methods for the individual subband images

of the lowpass filtering, in the LL- and HH-subband, however, the scanning orientation may be arbitrarily chosen.

After the subband decomposition and scanning, the resulting subband vectors are then quantized with simple scalar quantizers, where the bit-allocation is performed in a rate-distortion optimal sense according to the method presented in [8]. Since the bit-allocation values are highly sensitive to channel errors, we assume that these values are transmitted without errors to the decoder. The bit-rate information only contributes about 1% to the overall bit rate, so that we can easily protect the corresponding bits with a sufficiently strong channel code at the expense of only a minor bit rate increase.

4. Results

We apply this experimental image transmission system to the "Goldhill" test image of pixel dimension 512×512 using a $L = 5$-level wavelet filter bank and a source coding rate (including all side information) of $R = 0.35$ bit per pixel (bpp), where the following different estimation / decoding techniques are compared:

- Classical maximum-likelihood (hard decision) decoding ("ML").
- MS estimation without using a-priori knowledge about the source statistics ("MS, napr.").
- MS / MAP estimation with $P_{AP}(\lambda, k) = P(I_k^{(\lambda)} \mid \hat{\boldsymbol{I}}_0^k)$ ("MS, P" / "MAP, P").
- MS/MAP estimation with $P_{AP}(\lambda, k) = P(I_k^{(\lambda)} \mid \hat{\boldsymbol{I}}_0^{k+1})$ ("MS, PF1" / "MAP, PF1").

In the last two cases the transition probabilities are directly derived from the original image, which corresponds to the optimal case. However, this is infeasible in real transmission scenarios, so that approximations of the transition probabilities have to be obtained at the decoder, which will be discussed below. The simulation results are depicted in Fig. 4, where the average peak SNR (PSNR) values over

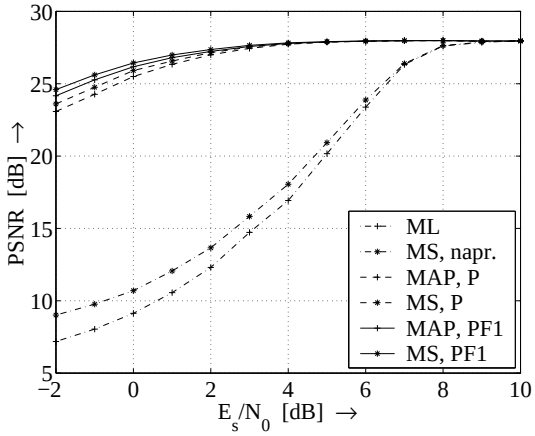

Figure 4. Performance for the "Goldhill" image using the original transition probabilities ($R = 0.35$ bpp, $L = 5$)

100 independent trials versus the ratio E_s/N_0 are displayed. From Fig. 4 we can see that exploiting the residual redundancy in the optimal estimation leads to a highly increased quality of the reconstructed image especially for a very noisy channel. Here, an additional gain of up to 1 dB is obtained when we utilize the extended a-priori knowledge in the "MS / MAP, PF1"-technique compared to the "MS / MAP, P"-approach. Furthermore, when using the same type of a-priori information, the MS estimation performs slightly better than the MAP technique.

In order to approximate the transition probabilities at the decoder they are computed from a training set of 129 images (i.e. faces, landscapes, satellite images), where the "Goldhill" and "Lena" images are not included in the set. This approximation method (denoted as "Tr.") is now compared with the optimal case from above using the original transition probabilities ("Orig."). Fig. 5 shows the results for the "Goldhill" and "Lena" images, resp., using a $L = 3$ level filter bank and a source coding rate of $R = 0.36$ bpp. We here use only three levels for the subband decomposition (compared to $L = 5$ in Fig. 4), since by using a training set with only 129 subband images the transition probabilities for the low-frequency subband images (and especially for the LL-subband) may not be representative for larger L. It can be seen from Fig. 4 that the best feasible estimation approach can be obtained by using the "MS, PF1, Tr."-technique. However, this method is inferior to all approaches using the original transition probabilities. For example, for the "Goldhill" image (Fig. 5(a)) in the worst case for $E_s/N_0 = 0$ dB we are about 2 dB away from the best optimal result, for the "Lena" image in Fig. 5(b) the deviation from the optimal case is even worse.

Finally, Fig. 6 compares the performance of the image transmission system for the "Goldhill" image with $L = 3$

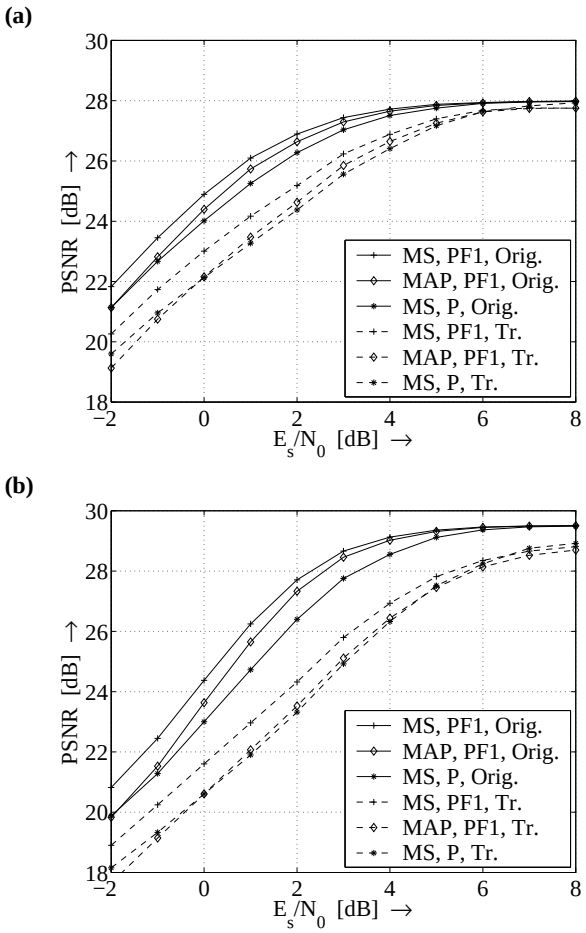

Figure 5. Performance using different methods for representing the transition probabilities at the decoder ($R = 0.36$ bpp, $L = 3$): (a) "Goldhill", (b) "Lena" image.

and $R = 0.36$ bpp for an exact match of the channel-SNR at the decoder (which has been assumed in all earlier simulations) and under channel mismatch conditions ("mism."). In the latter case the decoder is designed for a channel-SNR of 1 dB. We can see from Fig. 6 that the overall transmission system is very robust against a channel mismatch. Even a mismatch of ± 3 dB only leads to a PSNR-loss of about 1 dB.

An example of the good reconstruction quality is given in Fig. 7 for a highly corrupted channel ($E_s/N_0 = 0$ dB) using the "MS, PF1" estimator, where the transition probabilities are derived from the original image in Fig. 7(a), and from the training set in Fig. 7(b).

5. Acknowledgement

The authors would like to thank Mr. Artur Watkowski and Mr. Ragnar Thobaben for their help in implementing the simulation software.

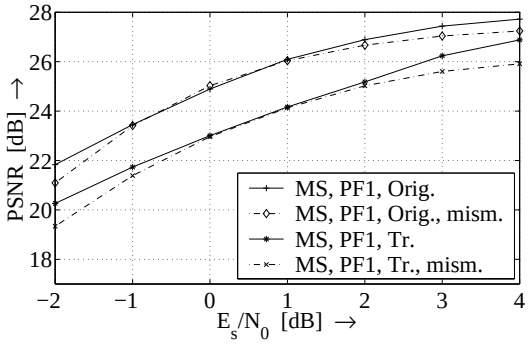

Figure 6. Performance under channel mismatch conditions ("mism."), and comparison to the matched case ($R = 0.36$ bpp, $L=3$). The decoder is designed for a fixed channel-SNR of $E_s/N_0 = 1$ dB.

6. Conclusion

We have shown for image transmission over AWGN-channels that especially for highly corrupted channels it seems advantageous to leave some residual redundancy at the output of the source coder. This implicit redundancy can be exploited for error-concealment at the decoder with only a negligible amount of additional channel coding using an optimal estimation approach based on bit-reliability information at the channel output and on the subband image statistics. The proposed transmission system is fairly robust against a mismatch of the channel-SNR at the decoder. Furthermore, the subband image statistics can be approximated at the decoder from a large image training set without leading to a large quality loss in the reconstructed image.

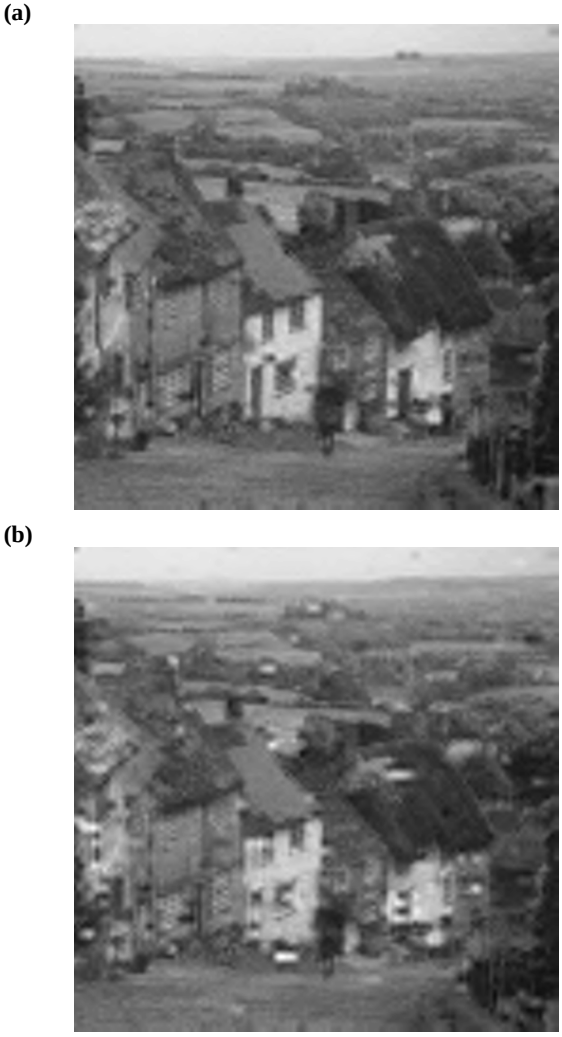

Figure 7. Reconstructed Goldhill image (estimator: "MS, PF1", $E_s/N_0 = 0$ dB (bit error probability 7.9%), $R = 0.36$ bpp): Transition probabilities (a) derived from the original image (PSNR $= 25.27$ dB), (b) derived from the training set (PSNR $= 23.29$ dB).

References

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Trans. on Image Processing*, 1(2):205–220, Feb. 1992.

[2] L. R. Bahl, J. Cooke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. on Inform. Theory*, pages 284–287, Mar. 1974.

[3] T. Fingscheid and P. Vary. Robust speech decoding: A universal approach to bit error concealment. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 1667–1670, Munich, Germany, Apr. 1997.

[4] J. Hagenauer. Source-controlled channel decoding. *IEEE Trans. on Comm.*, 43(9):2449–2457, Sept. 1995.

[5] T. Hindelang, J. Hagenauer, and S. Heinen. Source-controlled channel decoding: Estimation of correlated parameters. In *Proc. 3. ITG-Conf. "Source and Channel Coding"*, pages 251–258, Munich, Germany, Jan. 2000.

[6] N. Phamdo and N. Farvardin. Optimal detection of discrete Markov sources over discrete memoryless channels – applications to combined source-channel coding. *IEEE Trans. on Information Theory*, 40(1):186–193, Jan. 1994.

[7] P. Raffy, C. Pépin, and R. M. Gray. Robust stack-run image coding for noisy channels. In *Proc. 33rd Asilomar Conference on Signals, Systems and Computers*, pages 352–356, Pacific Grove, USA, Oct. 1999.

[8] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Trans. on Image Processing*, 2(2):160–175, Apr. 1993.

[9] K. Sayood and J. C. Borkenhagen. Use of residual redundancy in the design of joint source/channel coders. *IEEE Trans. on Comm.*, 39(6):838–846, June 1991.

[10] P. G. Sherwood and K. Zeger. Progressive image coding for noisy channels. *IEEE Signal Processing Letters*, 4(7):189–191, July 1997.