Two-dimensional soft-input source decoding for robust transmission of compressed images

J. Kliewer and N. Goertz

The transmission of compressed images over highly corrupted channels is addressed. The implicit residual redundancy inherent in the quantised subband images and the bit-reliability information at the channel output are utilised for error protection. As a novelty a known estimation technique is extended to the two-dimensional case, where both horizontal and vertical correlations are exploited. Especially for very noisy channels the quality of the reconstructed image is greatly increased compared to one-dimensional approaches.

Introduction: Recent joint source-channel decoding for the transmission of compressed images includes techniques where the residual redundancy after source encoding is used to improve the error resilience [1–3]. In these schemes the output signal of the decoder is estimated depending on bit-reliability information at the channel output and on the source statistics. As a novelty, this estimation technique is extended to spatial image correlations in two dimensions. We use a simple wavelet-based image source coder, which deliberately leaves some redundancy in the source-encoded bit stream; this *a priori* knowledge is exploited at the receiver. The scheme can easily be combined with conventional forward error protection using soft-output channel decoding [4].

Transmission system: The block diagram of the underlying transmission system is depicted in Fig. 1. Herein, the two-dimensional subband images are scanned after analysis filtering to obtain the one-dimensional source signal vectors $\boldsymbol{U}^{(\ell)} = [U_0^{(\ell)}, U_1^{(\ell)}, \ldots, U_k^{(\ell)}, \ldots]$, with $\ell = 0, 1, \ldots, K-1$ denoting the subband number and k the sample index. For simplicity the subband index ℓ will be omitted below.

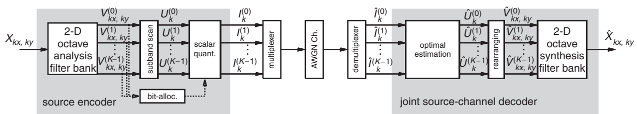

Fig. 1 *Experimental image transmission system*

The quantisation of the subband signal vectors leads to indices I_k which can be represented by N bits. We model the residual correlation between the indices I_k by a first-order stationary Markov process, which is described by the index-transition probabilities $P(I_k = \lambda | I_{k-1} = \mu)$, $\lambda, \mu = 0, 1, \ldots, 2^N - 1$. In what follows we will write $I_k^{(\lambda)}$ for the hypothesis $I_k = \lambda$. The bits $i_{l,k}$ of the source-encoded indices $I_k = \{i_{l,k}, l = 1, \ldots, N\}$ are transmitted over a binary-input additive white Gaussian noise (AWGN) channel, where after a mapping to bipolar bits we add white Gaussian noise samples $n_{l,k}$ that have zero mean and a variance of $\sigma_n^2 = N_0/2E_s$. E_s denotes the energy used to transmit each bit and $N_0/2$ represents the power spectral density of the channel noise.

Soft-input source decoding: The *a posteriori* probabilities (APPs) $P(\boldsymbol{I}_k^{(\lambda)} | \hat{I}_0^k) = P(I_k = \lambda | \hat{I}_0, \ldots, \hat{I}_{k-1}, \hat{I}_k)$, $\lambda = 0, \ldots, 2^N - 1$, describe the probability that the index $I_k = \lambda$ has been transmitted, given all received soft-bit vectors $\hat{\boldsymbol{I}}_0^k = [\hat{I}_0, \ldots, \hat{I}_{k-1}, \hat{I}_k]$ up to the time k. With the transition probabilities $P(\hat{I}_k^{(\lambda)} | I_{k-1}^{(\mu)})$ of the Markov model and the conditional probability density function (pdf) $p(\hat{I}_k | I_k^{(\lambda)})$ that describes the channel, the APPs can be calculated via the recursion [3, 5]

$$P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^k) = c_k \cdot p(\hat{I}_k | I_k^{(\lambda)}) \sum_{\mu=0}^{2^N-1} P(I_k^{(\lambda)} | I_{k-1}^{(\mu)}) P(I_{k-1}^{(\mu)} | \hat{\boldsymbol{I}}_0^{k-1}) \qquad (1)$$

The factor $c_k \in 2\,IR$ ensures that the $P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^k)$ are true probabilities (which sum up to one), and the initialisation for $k = 0$ can be carried out with the unconditional source-index probabilities $P(I_k^{(\lambda)})$. The APPs obtained by (1) can be used for an estimation of the decoder output. In the sequel we apply a mean-square (MS) estimator which is well suited for waveform-like signals, since it directly corresponds to an SNR-maximisation [3]:

$$\hat{U}_k = \sum_{\lambda=0}^{2^N-1} U_q(\lambda) \cdot P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^k) \qquad (2)$$

The MS estimation $\hat{U}_k$ equals the conditional expectation of the quantisation levels $U_q(\lambda)$.

Extension to two dimensions: Since subband images have spatial correlations in the horizontal and vertical direction, this additional knowledge can be exploited by two-dimensional (2-D) APPs. This is depicted in Fig. 2, where some received indices of a subband image are displayed, and $\hat{\boldsymbol{I}}_0^M = [\hat{I}_0, \hat{I}_1, \ldots, \hat{I}_M]$ denotes a received row or column vector of length $M+1$. The *a priori* knowledge is restricted to all indices in the boxes drawn with bold lines.

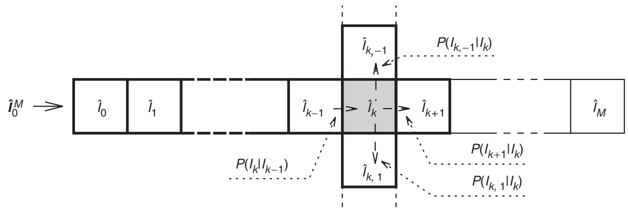

Fig. 2 *Received subband values and corresponding Markov models*

By use of the Bayes rule the 2-D APPs can be written according to

$$P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^{k+1}, \hat{I}_{k,-1}, \hat{I}_{k,1}) = c_k \cdot p(\hat{I}_{k+1}, \hat{I}_{k,-1}, \hat{I}_{k,1} | I_k^{(\lambda)}) \cdot P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^k) \qquad (3)$$

Both the memoryless property of the channel and the Markov property of the source are utilised in (3). The joint pdf $p(\hat{I}_{k+1}, \hat{I}_{k,-1}, \hat{I}_{k,1} | I_k^{(\lambda)})$ from (3) can now be further decomposed as

$$\begin{aligned} p(\hat{I}_{k+1}, \hat{I}_{k,-1}, \hat{I}_{k,1} | I_k^{(\lambda)}) = & \sum_{\mu=0}^{2^N-1} \sum_{\nu=0}^{2^N-1} \sum_{\kappa=0}^{2^N-1} p(\hat{I}_{k+1} | I_{k+1}^{(\mu)}) \\ & \times p(\hat{I}_{k,1} | I_{k,1}^{(\nu)}) p(\hat{I}_{k,-1} | I_{k,-1}^{(\kappa)}) \\ & \times \underbrace{P(I_{k+1}^{(\mu)}, I_{k,1}^{(\nu)}, I_{k,-1}^{(\kappa)} | I_k^{(\lambda)})}_{=: P(\mu,\nu,\kappa|\lambda)} \end{aligned} \qquad (4)$$

The term $P(\mu, \nu, \kappa | \lambda)$ describes the joint transition probabilities from the actual index I_k (the one under consideration) to all possible combinations of the neighbouring indices I_{k+1} and $I_{k,\pm 1}$. However, to store this joint probability a four-dimensional matrix is needed, leading to high memory requirements especially for larger N. A simplification is possible when the diagonal correlations in the subband image are neglected, which leads to $P(\mu, \nu, \kappa | \lambda) \simeq P(I_{k+1}^{(\mu)} | I_k^{(\lambda)})\ P(I_{k,1}^{(\nu)} | I_k^{(\lambda)})\ P(I_{k,-1}^{(\kappa)} | I_k^{(\lambda)})$. By inserting this approximation into (4) and using (3) we obtain the final expression for the APPs as

$$\begin{aligned} P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^{k+1}, \hat{I}_{k,-1}, \hat{I}_{k,1}) \simeq\ & c_k \cdot P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^k) \cdot \sum_{\mu=0}^{2^N-1} p(\hat{I}_{k+1} | I_{k+1}^{(\mu)}) P(I_{k+1}^{(\mu)} | I_k^{(\lambda)}) \\ & \times \sum_{\nu=0}^{2^N-1} p(\hat{I}_{k,1} | I_{k,1}^{(\nu)}) P(I_{k,1}^{(\nu)} | I_k^{(\lambda)}) \\ & \times \sum_{\kappa=0}^{2^N-1} p(\hat{I}_{k,-1} | I_{k,-1}^{(\kappa)}) P(I_{k,-1}^{(\kappa)} | I_k^{(\lambda)}) \end{aligned} \qquad (5)$$

In (5), the reliability values $P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^k)$ from (1) are weighted with three sum-terms. Each of these contains the corresponding transition probability and the channel term for the additionally considered indices I_{k+1}, $I_{k,1}$ and $I_{k,-1}$, respectively. An MS estimation at the decoder can again be carried out by simply replacing the probabilities in (2) with the values $P(I_k^{(\lambda)} | \hat{\boldsymbol{I}}_0^{k+1}, \hat{I}_{k,-1}, \hat{I}_{k,1})$ from (5).

Robust transmission of compressed images: The proposed estimation approach is now applied to the image transmission system in Fig. 1. For the subband decomposition we utilise a wavelet-based octave filter bank with L levels. Prior to quantisation the two-dimensional subband images are scanned in a meander-type fashion to obtain an one-dimensional subband vector $\boldsymbol{U}^{(\ell)}$. Since the bit-allocation information is highly sensitive to channel errors, we assume that this information is protected by a sufficiently strong channel code. At the receiver side, the estimation of the reconstructed subband coefficients $\hat{U}_k^{(\ell)}$ is carried out independently for every subband image as described above.

This experimental image transmission system is applied to the 'Goldhill' test image of pixel dimension 512×512 for a three level decomposition and a source coding rate of 0.36 bits per pixel (bpp) including all side information. We compare estimation techniques with the APPs given by (1) (denoted as '1-D') and with the APPs given by (5) (denoted as '2-D'). Furthermore, different techniques for obtaining the index transition probabilities at the decoder are utilised:

- From the original subband images ('Orig.'), which is the practically infeasible best case.
- From a training set ('Tr.') of images (i.e. faces, landscapes, 'Goldhill' image not included).
- From two-dimensional first-order auto-regressive modelling ('AR(1)') of the original input image, where the horizontal and vertical AR-coefficients are assumed to be known at the receiver.

In Fig. 3 the average peak-SNR (PSNR) values of the reconstructed 'Goldhill' image against the channel SNR E_s/N_0 are displayed. The '2-D, Orig.' technique outperforms all other methods especially for low channel SNRs. The best realisable decoding scheme can be obtained from the training set using the novel '2-D, Tr.' approach, which is superior compared to the one-dimensional version ('1-D, Tr.'). For lower channel SNRs it even leads to better results than the 1-D optimal case '1-D, Orig.'. Utilising the AR(1) model yields poor reconstruction PSNR compared to the other approaches.

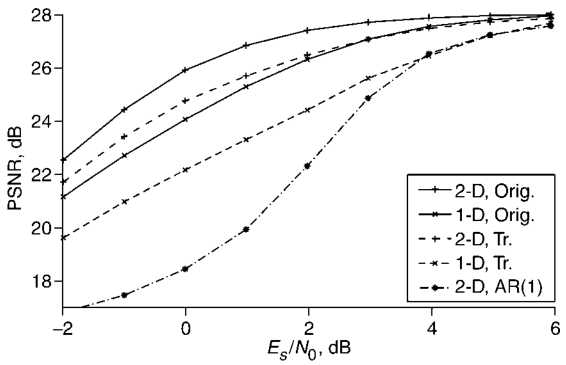

Fig. 3 *Performance for 'Goldhill' image*
$R = 0.36$ bpp, $L = 3$

Conclusion: The residual redundancy inherent in the quantised subband images can be utilised for error protection. As a novelty we have included the two-dimensional spatial correlations of the subband images into decoding which leads to a much better reconstruction compared to one-dimensional approaches. The subband image statistics can be best approximated from a large image training set.

26 October 2004
Electronics Letters online no: 20057590
doi: 10.1049/el:20057590

J. Kliewer (*School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom*)

N. Goertz (*School of Engineering and Electronics, The University of Edinburgh, Mayfield Rd., Edinburgh EH9 3JL, Scotland, United Kingdom*)

E-mail: norbert.goertz@ed.ac.uk

J. Kliewer: On leave from University of Kiel Institute for Circuits and Systems Theory, 24143 Kiel, Germany

References

1 Sayood, K., and Borkenhagen, J.C.: 'Use of residual redundancy in the design of joint source/channel coders', *IEEE Trans. Commun.*, 1991, **39**, (6), pp. 838–846
2 Hagenauer, J.: 'Source-controlled channel decoding', *IEEE Trans. Commun.*, 1995, **43**, (9), pp. 2449–2457
3 Fingscheidt, T., and Vary, P.: 'Robust speech decoding: a universal approach to bit error concealment'. Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, Munich, Germany, April 1997, pp. 1667–1670
4 Hagenauer, J., Offer, E., and Papke, L.: 'Iterative decoding of binary block and convolutional codes', *IEEE Trans. Inf. Theory*, 1996, **42**, (2), pp. 429–445
5 Phamdo, N., and Farvardin, N.: 'Optimal detection of discrete Markov sources over discrete memoryless channels—applications to combined source-channel coding', *IEEE Trans. Inf. Theory*, 1994, **40**, (1), pp. 186–193