IMAGE TRANSMISSION WITH ADAPTIVE POWER AND RATE ALLOCATION OVER FLAT FADING CHANNELS USING JOINT SOURCE CHANNEL CODING

Greg H. Håkonsen, Tor A. Ramstad and Anders Gjendemsjø
Department of Electronics and Telecommunications, Norwegian University of Science and Technology
O.S. Bragstads plass 2B, 7491 Trondheim, Norway.

Keywords: Joint source channel coding, nonlinear mappings, fading channel, image coder.

Abstract: A joint source channel coder (JSCC) for image transmission over flat fading channels is presented. By letting the transmitter have information about the channel, and by letting the code-rate vary slightly around a target code-rate, it is shown how a robust image coder is obtained by using time discrete amplitude continuous symbols generated through the use of nonlinear dimension changing mappings. Due to their robustness these mappings are well suited for the changing conditions on a fading channel.

1 INTRODUCTION

For transmission over wireless channels it is important to have a robust system as the channel conditions vary as a function of time. Traditional *tandem systems* use a channel code that is designed for a worst case *channel signal-to-noise ratio* (CSNR) where it can guarantee a *bit error rate* (BER) below a certain level. A source coder is then matched to the bit rate for which the channel code is designed. There are two main problems with this system. One is that this system breaks down very fast if the true CSNR falls below the design level. On the other hand, if the true CSNR is much higher than the design level, this system suffers from what is called the *leveling-off effect*. As the CSNR rises, the BER decreases, but the performance of the system remains constant after a certain threshold. This is due to the lossy part of the source coder: the design of the quantizer sets a certain distortion level which yields the target rate of the channel code.

To increase the average bit rate over a channel, one strategy is to divide the CSNR range into a set of regions and use different constellations for the different regions, see e.g. (Holm et al., 2003). This does not, however, combat the problem that most channel codes need to group very long bit sequences into blocks to perform well. So switching fast between these constellations becomes a problem.

Shannon's *separation theorem* (Shannon, 1948), says that a source coder and channel coder can be separately designed and still obtain an optimal communication system. To achieve this, the two coders have to have infinite complexity and infinite delay. When taking complexity and delay into account, *joint source channel coding* (JSCC) is a promising solution. The idea in JSCC is to optimize the source and channel codes together, and in that manner make the source information better adapted to the channel. Work on JSCC can be split into three main categories, one is fully digital systems, where the different parts of the source is given unequal error protection (UEP) on the channel, see e.g. (Tanabe and Farvardin, 1992). A second is hybrid digital-analog (HDA) systems, where some digital information is sent, but an analog component is sent as refinement, see e.g. (Mittal and Phamdo, 2002). In this paper the focus is on continuous amplitude systems, where the main information is sent without any kind of channel code. We use nonlinear Shannon mappings to adapt the transmission of an image over a flat fading channel without the use of any quantization or channel codes. This image communication system is partly based on the system presented for an *additive white Gaussian noise* (AWGN) channel in (Coward and Ramstad, 2000).

2 SYSTEM OVERVIEW

The system considered in this paper is a joint source channel coder for image transmission over flat fading

H. Hakonsen G., A. Ramstad T. and Gjendemsjø A. (2006).
IMAGE TRANSMISSION WITH ADAPTIVE POWER AND RATE ALLOCATION OVER FLAT FADING CHANNELS USING JOINT SOURCE CHANNEL CODING.
In *Proceedings of the International Conference on Wireless Information Networks and Systems*, pages 158-163

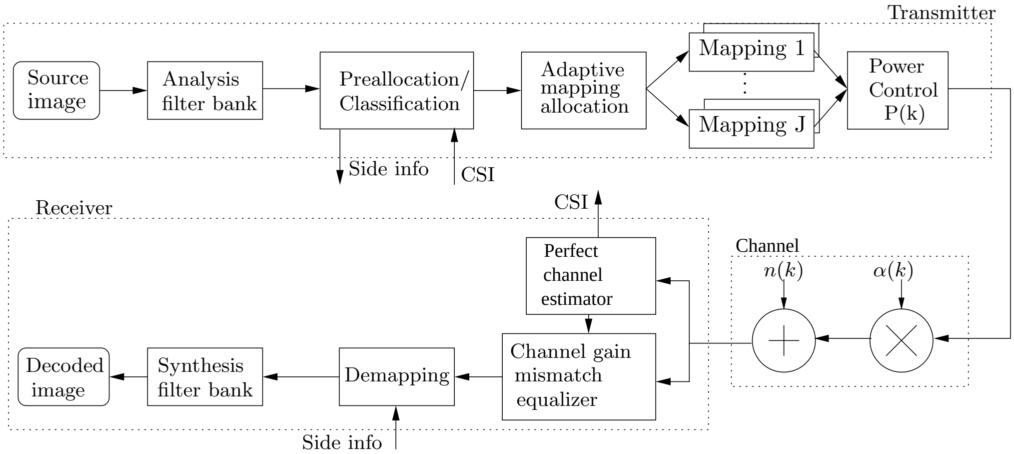

Figure 1: Proposed image transmission system.

channels. The structure of the system is as seen in Figure 1.

We assume that the return channel from the receiver to the transmitter has no delay, so that the transmitter has the same information as the receiver about the channel. To be able to decode the main information, the receiver needs some side information. The side information sent from the transmitter is assumed to be error free. At least at high rates, the size of the side information will have little impact on the total bandwidth. Any further investigation of the size of the side information is outside the scope of this paper. Transmission and coding are based on the use of nonlinear dimension changing mappings. The properties for these mappings will not be analyzed here, but can be found in e.g., (Coward and Ramstad, 2000; Fuldseth and Ramstad, 1997). The channel samples are transmitted as amplitude continuous, time discrete PAM symbols. The mappings work by taking g *source samples* and represent these samples by b *channel samples*, where g and b are integers such that $\hat{r} = b/g$. The resulting ratio $\hat{r}$, will from now on be denoted as the rate of the mapping. A mapping of $\hat{r} = 2$ represents each source sample as 2 channel samples, thus adding redundancy to protect that sample.

A mapping of $\hat{r} = 1/2$ is shown in Figure 2. Compression is achieved by representing two source samples by one channel sample. A $2D$ vector composed of two source samples is represented by (*). This point is then mapped to the closest point on the spiral, represented by (o). This point will then be transmitted as a continuous amplitude PAM symbol. Points on the dotted line can be represented by negative channel symbols, while points on the solid line can be represented by positive channel symbols. Channel noise will move this point so that the received point will have a different value, represented by ($\Diamond$). The total distortion in the reconstructed sample will be due to the approximation and channel noise.

Figure 2 can also be used to explain a mapping of $\hat{r} = 2$. Letting the spiral represent the source space, a source sample, represented by ($\Diamond$), can take any value along the spiral. By representing this sample as a $2D$ vector, the value of each coordinate can be transmitted as a continuous amplitude PAM symbol. After noise is added, the resulting $2D$ vector has been moved in the noise-space, represented by (*). The receiver knows that the original point has to be on the spiral, and maps it into the closest point along the spiral, represented by (o).

Designing good mappings for large dimension changes is increasingly more difficult as the dimensionality goes up, due to the high number of parameters that need to be optimized, similar to vector quantization (Gersho and M. Gray, 1992). Due to this, the mappings we use in this paper have rates

$$\hat{r}_j \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{2}{3}, 1, 2\}, j = 0, \ldots, 5. \quad (1)$$

Each mapping can be optimally designed for a set of given CSNRs, $\{\gamma_m\}_{m=1}^{M}$. The obtained signal representations are robust, so the system improves and degrades the performance gracefully around the design CSNR.

2.1 Source Description

The source in this paper is an image which is decorrelated by using a *filter bank* designed by Balasingham in (Balasingham, 1998). This filter bank consists of an 8×8 band uniform filter bank, and three different 2×2 band filter banks, used in a tree structure so each filter is applied to the low-low band of the previous stage. The filter bank is *maximally decimated* which

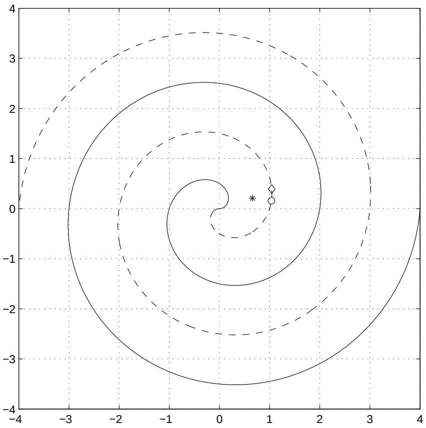

Figure 2: Example of mapping with $r_j = \frac{1}{2}$. $2D$ input vector marked by (*) is mapped to the closest point (o) in the channel space (the spiral). The channel noise will move the point along the spiral (◊).

keeps the number of pixels equal before and after filtering. Images have non-zero means, which implies a non-zero mean in the resulting low-low band, so to reduce the power of the image, the mean is subtracted from this band and sent as side information.

To cope with the local statistical differences within each band, the subband filtered image is split into N blocks. By using small blocks, the local statistics are captured better, but then the number of blocks will be larger, resulting in larger side information. To be able to decode the transmitted signal the receiver needs knowledge about the variances of all the source blocks. This information is sent as side information. Taking practical considerations into account, we have found 8×8 subband pixels to be a suitable block size. The variance $\sigma^2_{X_n}$ of each block is estimated using the RMS value.

We assume that the transform coefficients within a block can be modeled by a Gaussian pdf (Lervik and Ramstad, 1996). For a given distortion μ, the block variance can then be used to find the rate (bits/source sample), for each sub-source(block) from the following expression (Berger, 1971),

$$R_n = \frac{1}{2}\log_2\left(\frac{\sigma^2_{X_n}}{\sigma^2_{D_n}}\right) \text{ bits/source sample,} \quad (2)$$

where $\sigma^2_{D_n} = \min(\mu, \sigma^2_{X_n})$.

Assuming that all the N blocks are independent, the total average rate is given by

$$R = \frac{1}{2N}\sum_{n=0}^{N-1}\log_2\left(\frac{\sigma^2_{X_n}}{\sigma^2_{D_n}}\right) \text{ bits/source sample.} \quad (3)$$

The overall *signal-to-noise ratio* (SNR) is given by

$$\text{SNR} = \frac{\sum_{n=0}^{N-1}\sigma^2_{X_n}}{\sum_{n=0}^{N-1}\sigma^2_{D_n}}. \quad (4)$$

By using (4), it is possible to find the value of μ that yields a wanted SNR. This can be the scenario if the receiver needs a certain quality in the received image. Another scenario is when there is a time requirement, i.e. the transmission has be be finished within a certain time. This will put a rate constraint on (3), and the corresponding μ for this rate needs to be found. Large values of μ imply that some of the source blocks are discarded according to rate distortion theory (Berger, 1971).

2.2 Channel Adaptation

We consider a frequency-flat fading channel where it is assumed that the received signal $y(k)$ can be written as

$$y(k) = \alpha(k)v(k) + n(k), \quad (5)$$

where $\alpha(k)$ is a stationary channel gain, $v(k)$ is the sent signal and $n(k)$ is AWGN.

Let $\bar{P}$ denote the average transmit power and $N_0/2$ be the noise density on the channel of bandwidth B. The instantaneous CSNR, $\gamma(k)$, is defined as

$$\gamma(k) = \frac{\bar{P}|\alpha(k)|^2}{N_0 B}, \quad (6)$$

and let the expected value of the CSNR, $\bar{\gamma}$, be

$$\bar{\gamma} = \frac{\bar{P}}{N_0 B}. \quad (7)$$

To avoid an infinite amount of *channel state information* (CSI), the CSNR range is divided into $M+1$ regions, similar to e.g. (Holm et al., 2003). In this paper we choose to let a set of parameters $\{\gamma_m\}_{m=1}^{M}$ represent the different CSNR regions where we transmit, and let the region with worst CSNR represent outage. These regions will be denoted as channel states. The representation points γ_m, are placed inside each channel region similar to representation points in scalar quantization. The reason why this system is optimized for a representation point within a region instead of the edges which is common in *adaptive coded modulation* (ACM), is that the mappings used in this system degrades and improves gracefully in an area around the design CSNR instead of breaking down as is the case of conventional digital codes. So a mismatch between the CSNR for which a mapping is coded, and the actual CSNR, does not have so big impact as it would in traditional systems.

For a channel state γ_m, the instantaneous channel capacity is then given by, for $m = 1, \ldots, M$,

$$C_m = \frac{1}{2}\log_2(1+\gamma_m) \text{ bits/channel sample.} \quad (8)$$

2.3 Matching Source to Channel

Since the source is represented by source blocks, and the channel is divided into discrete channel states, it is interesting to look at the instantaneous rate, given by

$$r_{n,m} = \frac{\log_2\left(\frac{\sigma^2_{X_n}}{\sigma^2_{D_n}}\right)}{\log_2\left(1+\gamma_m\right)} \text{channel/source samples.} \quad (9)$$

For a given channel state γ_m, (9) gives the rate-change needed to be able to transmit a source block with variance $\sigma^2_{X_n}$ and distortion μ. Using a continuous set of dimensional changing mappings is, however, highly impractical, so in practice $r_{n,m}$ has to be approximated to the closest match in (1). And since the mappings are not ideal, the actual performance of each mapping has to be considered.

To set up a reasonable strategy for source-channel adaptation, the different source blocks are *preallocated* to the different channel states so that the block with largest variance is matched with the best channel state. The preallocation is done using the long term statistics of the channel. For a single transmission the channel does not behave exactly according to the statistics. So if, during transmission, there are no more blocks preallocated to a given state, blocks from a better channel state are transmitted, but then a new mapping has to be used calculated from (9). In this way the system can adapt to the varying conditions of the channel.

2.4 Finding the Representation Points

Since the source blocks are preallocated to different channel states, it is natural to include them in the calculation of channel representation points γ_m. Finding the best representation points γ_m will also include finding the best thresholds $\hat{\gamma}_m$ that defines the different channel regions $R_m = \{\gamma : \hat{\gamma}_m < \gamma \leq \hat{\gamma}_{m+1}\}$.

Since the rate-change needed is given by the ratio between the source rate and instantaneous channel capacity, (9), it is natural to include this into the optimization problem. By looking at the total average rate-change needed in a given channel state m

$$r_{\text{avg}_m}(\gamma_m) = \frac{\frac{1}{|I_m|}\sum_{n\in I_m}\log_2(\frac{\sigma^2_{X_n}}{\sigma^2_{D_n}})}{\log_2(1+\gamma_m)}, \quad (10)$$

where I_m is the set of source blocks allocated to the m'th channel state, we can find the mean squared error (MSE) denoted, ϵ, for all the M transmission channel states and all source blocks by

$$\epsilon = \sum_{m=1}^{M}\int_{R_m}\left(r_{\text{avg}_m}(\gamma) - r_{\text{avg}_m}(\gamma_m)\right)^2 p_\gamma(\gamma)d\gamma. \quad (11)$$

Minimization of ϵ with respect to R_m and γ_m has to be done through an iteration process, since the number of blocks in each channel state varies.

2.5 Optimization of Power Distribution

The mapping rate, $\hat{r}_n$, for each block, does not match the needed rate $r_{n,m}$ exactly. This leads to a variation in the resulting distortion that is not necessarily optimal. It is possible to minimize the distortion after the source blocks have been preallocated to a channel state and mapping rate by adjusting the power allocation. With a power constraint, and when taking the real distortions of each mapping into account, a Lagrangian can be set up

$$\mathcal{L} = \sum_{n=0}^{N-1}\sigma^2_{X_n}D_n(P_n) + \lambda\frac{\sum_{n=0}^{N-1}P_n\hat{r}_n}{\Psi}, \quad (12)$$

where, for practical mappings, $D_n(P_n)$ is the tabulated distortion of the mapping the n'th block is given, with power P_n, and Ψ is the total rate given by

$$\Psi = \sum_{n=0}^{N-1}\hat{r}_n(1+\frac{p_\text{o}}{p_\text{t}}), \quad (13)$$

where p_o and p_t are the probabilities of being in outage or transmit mode respectively.

To find the optimal P_n, (12) is differentiated and set to zero.

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}P_n} = \sigma^2_{X_n}D'_n(P_n) + \frac{\lambda\hat{r}_n}{\Psi} = 0 \quad (14)$$

$$\Updownarrow$$

$$-D'_n(P_n) = \frac{\lambda\hat{r}_n}{\sigma^2_{X_n}\Psi}. \quad (15)$$

The optimal λ, has to be found by checking the total power used by inserting the P^*_n from the optimal $D'_n(P_n)$ into

$$\bar{P} = \frac{\sum_{n=0}^{N-1}P^*_n\hat{r}_n}{\Psi}. \quad (16)$$

3 REFERENCE SYSTEM

To check the performance of the system we have chosen to compare with a traditional tandem source and channel coding system. As a source coder we have chosen to use JPEG2000 (Taubman and Marcellin, 2001)[1]. For transmission, the reference system employs an ACM scheme consisting of M transmission

[1]Version 5 of executables is downloaded from www.kakadusoftware.com

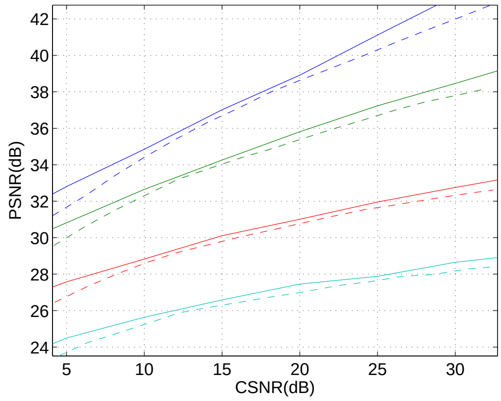

Figure 3: Performance of proposed system (dashed) compared to reference system (solid). "Goldhill" image. r_{avg} in channel samples/pixel, from below: $0.03, 0.1, 0.3, 0.5$.

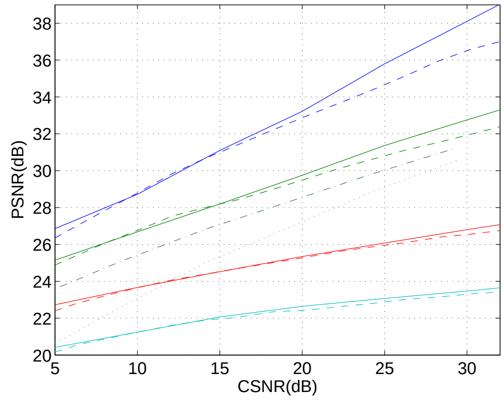

Figure 4: Performance of proposed system (dashed) compared to reference system (solid). "Bridge" image. r_{avg} in channel sample/pixel, from below: $0.03, 0.1, 0.3, 0.5$. Proposed system for $r_{\text{avg}} = 0.3$ with channel information every 1000'th (dash-dotted), and 20000'th channel symbol (dotted).

rates (bit/channel symbol), each assumed to achieve AWGN capacity for a given CSNR. To further maximize the *average spectral efficiency* (ASE) continuous power adaptation to is used within each CSNR region (Gjendemsjø et al., 2005).

The examples of the proposed coder are given for a certain CSNR, ($\bar{\gamma}$), and an overall target compression-rate, r_{avg} (channel samples/pixel). To be able to compare with the reference system, a bitrate in bits/channel sample, R_c, is found for a given $\bar{\gamma}$ for the channel code, and the resulting source bitrate, R_s, in bits/pixel is found and given as a parameter to the reference image coder through

$$R_s = r_{\text{avg}} R_c. \tag{17}$$

4 RESULTS

The following results are for images transmitted over a Rayleigh fading channel. The Doppler frequency is set to $f_d = 100$ Hz, and the carrier frequency, f_c, is set to 2 GHz, resulting in a mobile velocity of 15 m/s. The channel is simulated according to the Jakes' correlation model (Jakes, 1974). The number of channel states with transmission, M, is set to four for the proposed system and the reference system, thus allowing one outage state for both cases.

The quality of the received image is measured by *peak signal-to-noise ratio* (PSNR), defined as the ratio between the squared maximum pixel value, and the mean squared error (MSE) on a pixel basis for the whole image.

From Figure 3 and Figure 4 we see that the proposed system performs slightly poorer compared to the reference system. When comparing the two systems, it is however important to remember that the reference system uses near capacity achieving channel transmission. In practice the reference system would require infinite complexity and infinite delay. Looking at robustness we can see from Figure 5, that

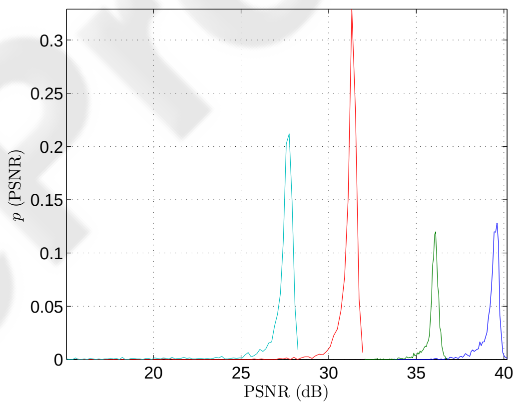

Figure 5: Estimated distributions of PSNR for target average rate r_{avg}, from left: $0.03, 0.1, 0.3, 0.5$. For $\bar{\gamma} = 22$ dB. "Goldhill" image.

even without any error protecting or correcting code, the robustness in the proposed system is quite high. Where as the reference channel transmission system will break down without perfect channel knowledge, the proposed system degrades gradually. In Figure 4, we have included the case where the transmitter does not have perfect channel knowledge. Two cases are included, one where the transmitter only knows the channel state for every 1000'th channel symbol, and one for every 20000'th symbol. For $r_{\text{avg}} = 0.3$ the total number of channel symbols is about 78300 for a 512×512 image. We will however not analyze this any further in this paper. It should noted that the reference system uses an optimal threshold for outage, while the proposed system uses a fixed outage threshold of $\hat{\gamma}_1 = 2$ for simplicity.

It should be emphasized that the performance for the proposed system is an average. Due to the limited number of symbols needed to transmit an image, the channel will not be fully ergodic. The result of this can be seen in the distribution of the average rate r_{avg} in Figure 6. The assumed rate is based on the preallocation of the blocks, but since the probabilities of each channel state will vary, the number of channel symbols will also vary. Since r_{avg} varies, the actual CSNR will vary slightly as well. So when reading the results in Figure 3 and Figure 4, one should keep in mind that the CSNR and PSNR is plotted as an average. An estimated distributions of the PSNR for different r_{avg} is given in Figure 5.

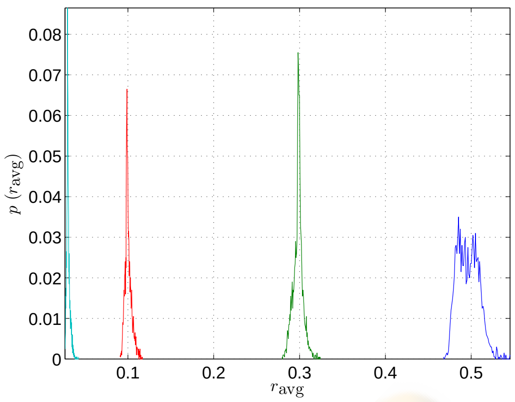

Figure 6: Estimated distributions of average transmission rates r_{avg} for target average rate from left: $0.03, 0.1, 0.3, 0.5$. For $\bar{\gamma} = 22$ dB. “Goldhill” image.

In Figure 3 and Figure 4, it can be seen that the performance of the proposed coder is not parallel with the performance of the reference coder. For high CSNR values this is due to the number of mapping rates available is to low in the rate-range of 0 to 2. For low CSNR values, there is a need for mappings with rates higher than 2.

5 CONCLUSION

We have shown how a joint source channel image coder system can achieve robust performance for transmission over a Rayleigh fading channel, when allowing the average rate to vary slightly around a target rate. This is done by choosing nonlinear mappings best suited for the current channel condition, the importance of the transmitted image block, and by allocating the power to minimize the distortion. The proposed system has been shown to be comparable to a reference tandem system using capacity-approaching codes.

REFERENCES

Balasingham, I. (1998). *On Optimal Perfect Reconstruction Filter Banks for Image Compression*. PhD thesis, Norwegian University of Science and Technology, Norway.

Berger, T. (1971). *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice Hall, Englewood Cliffs, NJ.

Coward, H. and Ramstad, T. (2000). Robust image communication using bandwidth reducing and expanding mappings. volume 2, pages 1384–1388, Pacific Grove, CA, USA. Conf. Record 34th Asilomar Conf. Signals, Syst., Comput.

Fuldseth, A. and Ramstad, T. A. (1997). Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, volume 4, pages 3093–3096.

Gersho, A. and M. Gray, R. (1992). *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, MA, USA.

Gjendemsjø, A., Øien, G. E., and Holm, H. (2005). Optimal power control for discrete-rate link adaptation schemes with capacity-approaching coding. In *Proc. IEEE Global Telecommunications Conference*, pages 3498–3502, St. Louis, MO.

Holm, H., Øien, G. E., Alouini, M.-S., Gesbert, D., and Hole, K. J. (2003). Optimal design of adaptive coded modulation schemes for maximum spectral efficiency. pages 403–407. Proc. Signal Processing Advances in Wireless Communications (SPAWC).

Jakes, W. C. (1974). *Microwave Mobile Communication*. Wiley, New York.

Lervik, J. M. and Ramstad, T. A. (1996). Optimality of multiple entropy coder systems for nonstationary sources modeled by a mixture distribution. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, volume 4, pages 875–878, Atlanta, GA, USA. IEEE.

Mittal, U. and Phamdo, N. (2002). Hybrid digital-analog(hda) joint source-channel codes for broadcasting and robust communication. *IEEE Trans. Inform. Theory*, 48(5):1082–1102.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423.

Tanabe, N. and Farvardin, N. (1992). Subband image coding using entropy-coded quantization over noisy channels. *IEEE J. Select. Areas Commun.*, 10(5):926–943.

Taubman, D. and Marcellin, M. (2001). *JPEG2000 Image compression fundamentals, standards and practice*. Kluwer Academic Publishers.