

VIDEO COMMUNICATION OVER BROADCAST CHANNELS

Susie J. Wee

Hewlett-Packard Laboratories
Palo Alto, California USA
swee@hpl.hp.com

Michael O. Polley

Texas Instruments Incorporated
Dallas, Texas USA
polley@hc.ti.com

ABSTRACT

This paper investigates the problem of communicating video over a broadcast channel. The broadcast channel is expressed in terms of the channel capacity that exists between the transmitter and each receiver in the broadcast area – Shannon's separation theorem does not apply for video communication over this class of channels. Digital (discrete time, discrete amplitude) and hybrid (discrete time, discrete/continuous amplitude) transmission and video coding methods are discussed. Joint source and channel coding principles are employed to effectively couple these methods to form efficient systems for communicating video over broadcast channels. A framework is presented for characterizing and bounding the performance of these systems; the results lead to interesting directions for future work.

1. INTRODUCTION

Typical video communication systems are based on purely digital source and channel coding methods. These systems are usually optimized for point-to-point communication environments, in which one has a priori knowledge of the channel SNR or available bit rate. In this work, we address the problem of video communication over broadcast channels, in which a single transmitter delivers a coded video signal to a number of receivers operating under a wide range of channel SNRs. These channels are common in applications such as indoor wireless communication and terrestrial television broadcasting.

This paper is concerned with issues that arise when communicating video over broadcast channels. It describes the behavior of digital and hybrid transmission in a broadcast environment. The relevance of joint source/channel coding is discussed and basic descriptions of digital and hybrid video coding are given. The performances of the resulting digital and hybrid systems are evaluated and the relative merits of each system are discussed.

2. CONVENTIONAL APPROACHES

A number of bitstream-scalable, digital video coders have been developed for progressive transmission of images and video [1, 2]. In these coders, images and video are encoded into prioritized data streams that can be decoded up to

A portion of this work was performed in collaboration with Prof. W.F. Schreiber while the authors were with the Research Laboratory of Electronics at the Massachusetts Institute of Technology, Cambridge, Massachusetts USA.

any point in the data stream to reconstruct video with a quality that depends on the amount of decoded data. However, these papers do not address the broader problem of video communication. A number of multiresolution advanced television systems have been designed for terrestrial video broadcasting [3, 4, 5]. This paper focuses on transmitting single-resolution images and video over the broadcast channel using a joint source/channel coding approach; this focus enables a better comparison between the various approaches.

3. BROADCAST CHANNELS

In the broadcast channel, a single transmitter delivers a signal to a number of receivers operating under a wide range of channel signal-to-noise ratios (SNRs). The challenge in communicating video over broadcast channels is to jointly optimize the video quality decoded by the various receivers. Intuitively, an effective solution would enable the video quality decoded by each receiver to depend on the channel SNR – a higher SNR should result in higher video quality. This property is often referred to as graceful degradation because it describes the gradual decrease in video quality that should occur with a gradual decrease in channel quality.

The channel capacity dictates the maximum data rate that a transmitter can deliver to a receiver across an additive white Gaussian noise channel with a given SNR, and is given by $C = \frac{1}{2} \log_2(1 + \text{SNR})$ in bits per 1-D symbol. Shannon's noisy channel coding theorem states that for point-to-point communication, data can be transmitted reliably at rates up to the channel capacity. In the broadcast channel, receivers with higher SNRs have the capacity to receive higher data rates than those with lower SNRs [6]. The problem that remains is to find practical methods that best utilize the available capacity in order to provide the highest possible video quality to all the receivers.

In this work, we consider the transmission of *discrete-time* signals over broadcast channels. These signals can have *discrete-amplitude* and/or *continuous-amplitude* components, which are hereafter referred to as *digital* and *analog* components, respectively. We focus our investigation on the performance of purely digital transmission methods and hybrid digital/analog transmission methods over broadcast channels.

Single-rate digital transmission involves sending a discrete-time, discrete-amplitude signal over the channel. The maximum data rate R is limited by the channel capacity of

the receivers operating at a threshold SNR. Thus, receivers with a channel SNR that exceeds the threshold can decode the data with rate R , while receivers with lower SNRs can not decode any data. Figure 1 illustrates the decoded data rate of an ideal digital system as a function of channel SNR.

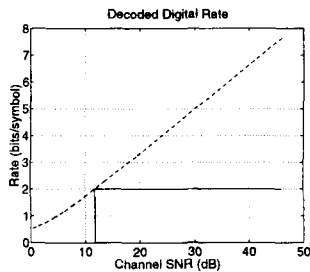


Figure 1: *Digital transmission.* The digital signal is characterized by the decoded digital data rate as a function of the channel SNR. The dashed line shows the channel capacity as a function of the channel SNR.

The concept of hybrid transmission was proposed by Schreiber in [7]. The novelty lies in its improved use of the channel capacity available in the broadcast channel. The hybrid data stream is a *discrete-time* signal composed of a digital *discrete-amplitude* component and an analog *continuous-amplitude* component.

The hybrid signal is formed by first creating a digital signal, and then superimposing an analog signal. A simple example of a one-dimensional signal space for a hybrid data stream is shown in Figure 2. The signal can be decoded in the following steps. First, the digital information is determined based on the polarity of the received signal. In this step, the analog information is seen as noise. Once the digital component is determined, it is extracted from the hybrid signal and the remaining amplitude represents the received analog value, which is degraded by the channel noise.

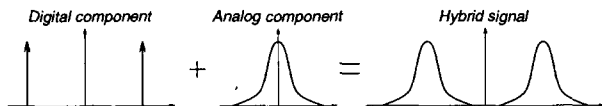


Figure 2: *A simple one-dimensional hybrid constellation.* The hybrid signal is composed of a digital discrete-amplitude component and an analog continuous-amplitude component.

The signal power of the transmitted hybrid signal, P_H , depends on the powers of its digital and analog components, P_d and P_a . If the components are independent then $P_H = P_d + P_a$. Proper embedding of the digital and analog components yields the following capacities and corresponding SNRs without loss of efficiency [8]:

$$C_d = \frac{1}{2} \log_2 (1 + \text{SNR}_d) \quad \text{SNR}_d = \frac{P_d}{P_a + N} \quad (1)$$

$$C_a = \frac{1}{2} \log_2 (1 + \text{SNR}_a) \quad \text{SNR}_a = \frac{P_a}{N}. \quad (2)$$

The concept of "capacity" for the analog component is used to indicate the maximum SNR (or fidelity) with which the

analog component can be received. In other words, the capacity of the analog component is defined as the highest rate at which a digital signal could be transmitted across a channel with SNR_a , which is determined by P_a , P_d , and the channel SNR.

The received hybrid signal is characterized by the data rate of its decoded discrete-amplitude component and the SNR of its decoded continuous-amplitude component. Figure 3 shows an example of this data rate and SNR as a function of channel SNR. This transmission scheme is efficient in that the analog SNR results in an analog capacity that equals the distance between the digital capacity and the channel capacity for each receiver in the broadcast area. The discrete-amplitude data rate and continuous-amplitude SNR can be traded off by varying the power distribution between the two components.

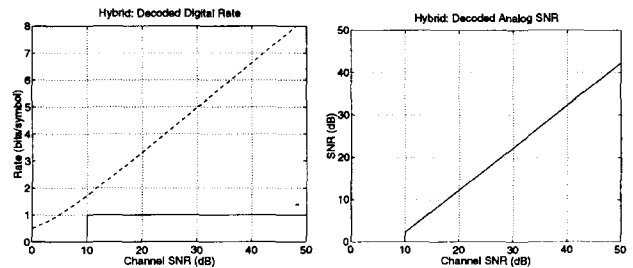


Figure 3: *Hybrid transmission.* The hybrid signal is characterized by the decoded data rate of its discrete-amplitude component and the received SNR of its continuous-amplitude component as a function of the channel SNR. The dashed line shows the channel capacity as a function of the channel SNR.

Multirate digital transmission can be viewed as a special case of hybrid transmission in which the analog component is restricted to a discrete number of amplitudes. This type of transmission does not achieve true graceful degradation, but achieves a property often referred to as stepwise graceful degradation. For the sake of simplicity, we will not discuss this topic any further.

4. VIDEO COMMUNICATION

Shannon's source-channel coding theorem is often interpreted as the separation theorem, which states that separate source and channel coding is as effective as joint coding, i.e. the source and channel coders can be designed separately without loss of optimality. This assumes that the source coder removes all the redundancy from the data and the channel coder reinserts the redundancy necessary for reliable transmission across the channel. The rate of the source-coded data equals the source entropy in the case of lossless coding, and it equals $R(D)$, the minimum rate for a given distortion D , for lossy coding. Lossless coding is seldom applicable in video communication environments because of the large data rates inherent to video and the narrow bandwidths of typical channels. Meanwhile, the existence of a lossy separation theorem for broadcast channels is unknown at this time [9]. Even if such a theorem exists, a number of practical issues would likely make it inappropriate for video communication, e.g. long block-length codes

are impractical and source redundancy can not be completely eliminated. Thus, joint design of the source and channel coders can be beneficial for video communication over broadcast channels.

Common video coding techniques include transform coding and motion-compensated transform coding. In transform coding, each video frame is coded independently with transform coding techniques. Only the transform coefficients need to be coded into the transmitted data stream. In motion-compensated transform coding, predictive methods are used to form a prediction of the current frame based on previously coded frames, and the prediction error, known as the residual, is coded with conventional transform coding techniques. The prediction is completely specified by a set of motion vectors which must be coded into the transmitted data stream along with the residual transform coefficients.

The motion vectors are critical when reconstructing video from a motion-compensated transform coder. For this reason we assume that if motion compensation is used, the motion vectors are accurately encoded into the digital portion of the data stream. The problem that remains is to code the residual transform coefficients into the transmitted data stream. Since transform coding is essential for both transform coding and motion-compensated transform coding, the remainder of this paper considers the problem of coding transform coefficients into a data stream that will be transmitted over the broadcast channel.

A digital transform coder codes an input video signal into a digital data stream as shown in Figure 4. The input signal is transformed with an orthogonal transformation. The low-frequency (DC) coefficients are quantized and coded into a digital data stream. The high-frequency (AC) coefficients are also quantized. After quantization, many of the low-amplitude coefficients are set to zero, and only the locations and amplitudes of the nonzero coefficients need to be coded into a digital data stream. The data rate required to code the digital data is given by

$$R = R_{DC} + R_{location} + R_{amplitude} \quad (3)$$

where R_{DC} , $R_{location}$, and $R_{amplitude}$ are the digital data rates needed to code the DC, AC location, and AC amplitude information.

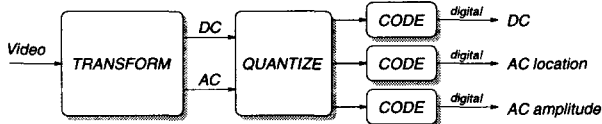


Figure 4: *Digital transform coding.* In conventional digital transform coding, the coefficients are quantized, and the location and amplitude information of the nonzero quantized coefficients are encoded into a digital data stream.

A hybrid transform coder was designed for hybrid transmission [10, 11]. This coder encodes the transform coefficients into a hybrid data stream, which is composed of a digital and an analog component as described in Section 3. The hybrid video coder is shown in Figure 5. As in the conventional digital system, the input signal is transformed with an orthogonal transform and the DC transform coefficients are quantized and coded into a digital data stream. However, in the hybrid system the AC coefficients are not

quantized. Rather, a number of these coefficients are *selected* for transmission and coded into a hybrid data stream. The amplitudes of the selected coefficients are transmitted in full precision in the analog component of the hybrid signal, and their locations are coded into the digital component of the hybrid signal.

The hybrid signal can be characterized by the decoded digital data rate and the received analog SNR which are given by

$$R = R_{DC} + R_{location} \quad SNR = SNR_{amplitude} \quad (4)$$

where R_{DC} and $R_{location}$ are the digital data rates needed to code the DC coefficients and the AC location data and $SNR_{amplitude}$ is the received SNR of AC amplitude data.

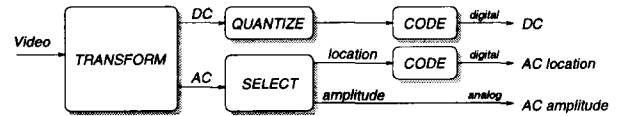


Figure 5: *Hybrid transform coding.* In hybrid transform coding, transform coefficients are selected for coding. The DC coefficients are coded separately into a digital data stream. The locations of the selected AC coefficients are encoded into a digital data stream, and the amplitudes are transmitted in full precision.

A generalization of this hybrid video coder can lead to improved performance at the expense of added complexity. In the method described above, the locations and amplitudes of a set number of selected coefficients are encoded into the digital and analog components of the hybrid data stream. A more general representation would allow the data to be *ordered* or *packed* more efficiently – the digital component can be used to robustly represent the ordering information and the analog component can be used to represent the values of the ordered data. In this scenario, the digital component describes the ordering information while the analog component contains the ordered data.

5. JOINT SOURCE/CHANNEL CODING

Achieving high performance in a video communication system requires an effective matching of the source and channel coders. Similarly, the transmission methods described in Section 3 must be carefully coupled to the video coding methods described in Section 4.¹ In this section, we describe the digital and hybrid video communication systems. We characterize and bound the system performance as a function of the channel SNR. This allows us to compare the relative merits of the two systems for video communication over broadcast channels.

In the digital video communication system, combining the video coding and transmission methods is quite straightforward. The broadcast area is specified by a threshold SNR, and all receivers with a channel SNR exceeding the threshold must be able to receive the video signal. Thus, the maximum data rate is limited by the capacity of the receivers at the threshold SNR. Since the threshold SNR and data rate are specified in advance, the task of the digital video coder is to efficiently encode the video into a digital

¹This is not meant to imply that source and channel coding are precisely analogous to video coding and transmission.

data stream with the appropriate data rate. All the receivers in the broadcast area will decode the same video, and the quality of this video depends on the capacity of the receivers at threshold.

The hybrid video communication system delivers different video qualities to receivers depending on their individual channel SNRs – receivers with higher channel SNRs can decode higher-quality video. In this system, a percentage of the transform coefficients are *selected* for coding. The locations of the selected coefficients are coded into a digital data stream (discrete time, discrete amplitude), and their amplitudes are coded into an analog data stream (discrete time, continuous amplitude). The hybrid signal power can be traded off between the digital and analog components to achieve the desired distribution between the digital data rate and analog SNR. Digital transmission is used to robustly communicate sensitive video data such as motion vectors and coefficient location information while analog transmission is used to transmit less sensitive data types such as coefficient amplitude information.

Notice that in the digital system, the selection process implicitly occurs during quantization. After quantization, the coefficients with nonzero amplitudes are the selected coefficients. Selection by quantization ensures that the largest-amplitude weighted coefficients are coded into the data stream. This optimizes the reconstructed video for the given number of coefficients in a weighted least-squares sense, but it does not necessarily optimize the performance for the total data rate. In hybrid video coding, the selection process is more general in that any set of coefficients can be selected for transmission as long as it can be coded into a hybrid data stream that meets the requirements of the system. When designing a selection scheme, consideration must be given to a number of criteria, including the visual quality of the reconstructed video, the data rate needed to represent the location data, and the noise performance of the amplitude data.

The sources of error can be categorized into two classes: the error from the unselected coefficients and the error from the selected coefficients, keeping in mind that the set of selected coefficients is not necessarily the same in the two systems. In both systems, the error from the unselected coefficients equals the energy in those coefficients. The error from the selected coefficients, however, has different sources in the two systems. In the digital system, the error is due to the discretization of the amplitudes during the quantization process, and is the same for all the receivers in the broadcast area. In the hybrid system, the selected coefficient amplitudes are not discretized; rather, they are transmitted in full precision. However, channel noise degrades these values during transmission. Thus, the selected coefficients are received with a fidelity that depends on the channel SNR.

The video communication system can be characterized for broadcast channels by plotting the decoded video quality as a function of channel SNR. For the digital system, the quality is constant for all receivers with channel SNRs that exceed the threshold. This video quality is limited by the unselected coefficients and the discretized amplitudes of the selected coefficients. For the hybrid system, the quality starts at some baseline value for the threshold receivers and increases with the channel SNR. The improvement contin-

ues until the selected coefficients are received with their full precision, at which point the quality saturates and stays constant for higher channel SNRs. Thus, the video quality is limited by the unselected coefficients and the channel noise on the selected coefficients. Notice that in this simple system, the maximum quality is limited by the coefficients that were not selected for transmission.

The performance of the hybrid video communication system can be improved by using a more sophisticated hybrid video coder in conjunction with the hybrid transmission scheme. By more efficiently packing the data into the hybrid data stream, both the rate of improvement and the point of saturation can be increased. The authors are currently investigating these possibilities. This problem is similar but not identical to the problem addressed by bit-stream scalable image and video coders [1, 2]. The difference lies in the distinction between the analog and digital components of the hybrid data stream. In essence, the hybrid video transmission problem can be viewed as a two-channel problem in which the digital component can be used to robustly deliver important data over a reliable channel while the analog component can be used to transmit less sensitive data over a noisy channel.

6. ACKNOWLEDGEMENTS

The authors would like to thank Prof. W.F. Schreiber for motivating and supervising much of this work and Prof. M.D. Trott for a discussion regarding the efficient packing of hybrid data

7. REFERENCES

- [1] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445–3462, December 1993.
- [2] D. Taubman and A. Zakhor, "Multi-rate 3-D subband coding of video," *IEEE Transactions on Image Processing*, vol. 3, pp. 572–588, September 1994.
- [3] W. Schreiber, "Considerations in the design of HDTV systems for terrestrial broadcasting," *SMPTE Journal*, pp. 668–677, September 1991.
- [4] W. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions," *Proceedings of the IEEE*, vol. 83, pp. 958–981, June 1995.
- [5] K. Ramchandran, A. Ortega, K. Uz, and M. Vetterli, "Multiresolution broadcast for digital HDTV using joint source/channel coding," *IEEE Journal on Selected Areas in Communications*, vol. 11, pp. 6–23, January 1993.
- [6] T. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 2–14, January 1972.
- [7] W. F. Schreiber, E. Adelson, A. Lippman, Gong Rongshu, P. Monta, A. Popat, H. Sallic, P. Shen, A. Tom, K. Zangi, and A. Netravali, "A compatible high-definition television system using the noise-margin method of hiding enhancement information," *SMPTE Journal*, vol. 98, Dec. 1989.
- [8] T. Cover and J. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, Inc., 1991.
- [9] M. D. Trott, "Unequal error protection codes: Theory and practice," in *IEEE Information Theory Workshop*, (Haifa, Israel), June 1996.
- [10] S. Wee, *Scalable Video Coding*. PhD thesis, Massachusetts Institute of Technology, 1996.
- [11] M. Polley, *Spectrum-Efficient Modulation and Channel Coding for Terrestrial Broadcasting*. PhD thesis, Massachusetts Institute of Technology, 1996.