CROSS-LAYER OPTIMIZATION WITH POWER CONTROL IN DS-CDMA VISUAL SENSOR NETWORKS

Elizabeth Serena Pynadath, Lisimachos P. Kondi

332 Bonner Hall, Dept. of Electrical Engineering
State University of New York at Buffalo, Buffalo, NY 14260
Email: { pynadath,lkondi }@eng.buffalo.edu

ABSTRACT

In this paper, we propose an approach for cross-layer optimization with power control for a Direct Sequence-Code Division Multiple Access (DS-CDMA) visual sensor network where nodes monitor scenes with varying levels of motion. Our technique simultaneously assigns a source coding rate, a channel coding rate, and a power level to all nodes in the network based on two criteria that maximize the quality of video of the entire network as a whole, subject to a constraint on the total chip rate. One method minimizes the average end-to-end distortion amongst all nodes, while the other method minimizes the maximum distortion of the network. To reduce the computational complexity of the solution, Universal Rate-Distortion Characteristics (URDCs) are obtained experimentally to relate bit errors probabilities to the distortion of corrupted video. The URDCs are used in conjunction with channel characteristic plots obtained by using Rate-Compatible Punctured Convolutional (RCPC) codes for channel coding.

Index Terms— Visual system, Networks, Power control, Video Coding, Code division multiaccess, Convolutional codes, Multimedia communication

1. INTRODUCTION

There has been a rapidly increasing demand for real-time video transmission over wireless networks. Sensor networks previously focused on networks that transmit scalar information such as temperature, pressure, acoustic data, etc. Visual sensor networks are much more challenging due to the high bit rates and delay constraints required for video transmission. These networks are comprised of typically low-weight distributed sensor nodes that can communicate with a centralized control unit. In this work, we will use the terms, "nodes" and "users", interchangeably. The centralized control unit performs channel and source decoding to obtain the received video from each node. The control unit transmits information to the nodes in order to request changes in transmission parameters, such as source coding rate, channel coding rate, and transmission power. Applications of visual sensor networks include surveillance, automatic tracking and signaling of intruders within a physical area, command and control of unmanned vehicles, and environmental monitoring.

DS-CDMA channels allow several users to use the same frequency band at the same time. Users' transmissions are distinguished through the use of different spreading codes. These channels differ from traditional wired networks in the types of errors they introduce, as well as, the severity of these errors. Problems like multipath fading and cochannel interference can cause high bit error rates that can result in a devastating degradation of the quality of the transmitted video. Even if the spreading codes used are orthogonal to each other, transmissions of one node cause interference to the other nodes, due to possible asynchronous transmissions and multipath fading. Since all nodes transmit on the same frequency, interference within a channel plays a significant role in determining the system's capacity and its quality-of-service (QoS). The transmit power for each node must be minimized to limit the interference experienced by other nodes in the system. This is also important for sensor networks, since the nodes are typically battery-operated and have limited energy. But, at the same time, a node's power should be high enough to maintain its own quality.

Although Shannon's principle of separability states that it is possible to design source and channel coding separately without loss of optimality, the principle assumes that the source and channel codes are of arbitrarily long lengths. Since this assumption does not hold in practical situations due to limitations on computational power and processing delays, it is useful to consider source and channel coding jointly [1]. In [2], the source coding rate, channel coding rate, and processing gain are optimized for the user-of-interest, with the assumption that the powers of all the users are given. The authors consider only the user-of-interest instead of solving the multi-user problem. Assuming the powers of all users are given is not a realistic assumption, and they do not utilize power control to address the interference-limited nature of CDMA systems.

Using power control as an indirect way of controlling a user's associated error probability in order to achieve a certain quality-of-service (QoS) was first proposed by [3], but they did not jointly optimize the source coding rate and the power level to achieve the desired QoS. Using power control alone is not enough to ensure adequate quality if video was transmitted. In [4], a joint source coding-power control approach was presented that allocates a source coding rate and the energy-per-bit to the multiple-access interference (MAI) density to each user to maximize the per-cell capacity and the end-to-end QoS for individual users. The authors kept a fixed channel coding rate instead of assigning an optimal channel coding rate to each user, as well [4]. However, for transmission over error-prone channels, it is imperative that an acceptable choice for the channel coding rate is made to offer an adequate level of protection for the data transmitted.

In this paper, we develop a novel cross-layer optimization technique that operates across the physical, data link, and application layers of the system. At the physical layer, the transmission power will be determined. At the data link layer, the channel coding rates will be selected. And at the application layer, the source coding rate for video compression will be determined. Our algorithm simultaneously allocates a source coding rate, a channel coding rate, and a power level to all nodes in a DS-CDMA visual sensor network. We assume the nodes in the network are equipped with a video camera deployed to survey a large area. To create a more realistic scenario with varying levels of motion, some nodes will be imaging a stationary field while other nodes will be imaging scenes with a high level of motion. Video sequences with less motion can be source encoded at a lower bit rate while still yielding good picture quality. The centralized control unit should be able to request that the video at specific nodes be transmitted at a lower bit rate, if it is deemed as being capable of still producing an adequate picture quality. These nodes that compress their video at a lower bit rate are left with more bits for channel coding and can afford to transmit at a lower power, so that they will cause less interference to the other node. For this reason, DS-CDMA is an appropriate choice for use in our visual sensor network set-up. Unlike previous methods, we recognize that different nodes of a sensor netowrk can have different source coding rate requirements due to different scene activity and propose to jointly optimize all nodes using two criteria. Our first method will result in the minimal average end-to-end distortion over all the nodes in the network while our second method will minimize the maximum distortion amongst all nodes. The goal will be to maximize the received video quality of all nodes in the network while maintaining that each node transmits at an appropriate level of power. The optimization algorithm proposed uses URDCs along with channel characteristic plots to reduce the computational complexity.

The rest of the paper is organized as follows. In section 2, we describe the channel coding parameters. In section 3, the cross-layer optimization algorithm is explained. In section 4, experimental results are presented, and in section 5, conclusions are drawn.

2. CHANNEL CODING

In this work, we use RCPC codes for channel coding. With convolutional coding, the source data is convolved with a convolutional matrix **G**, which specifies which delayed inputs to add to the current input. This process is equivalent to passing the input data through a linear finite-state register where the tap connections are defined by **G**. Unlike linear block codes that have a number of channel code symbols for a corresponding block of source symbols, convolutional coding generates one codeword for the entire source data. Convolution is the process of modulo-2 addition of the current source bit with previously delayed source bits. The rate of the convolutional code is defined as *k/n* where *k* is the number of input bits and *n* is the number of output bits.

Commonly, decoding convolutional codes is done with the Viterbi algorithm, which is a maximum-likelihood sequence estimation procedure [5]. There are two types of Viterbi decoding: soft and hard decoding. In soft decoding, decision statistics of the channel output are passed to the decoder. Usually, the distortion metric used is the Euclidean distance. In hard decoding, the decision of the received bit is made before the received data is input into the Viterbi decoder. The distortion metric commonly used for hard decoding is the Hamming distance [6].

Punctured convolutional codes were mainly developed to simplify Viterbi decoding for rate k/n with two branches arriving at each node instead of 2^k branches. Puncturing is the process of deleting bits from the output sequence in a predefined manner so that fewer bits are transmitted than in the original code. The idea of puncturing was extended to include the concept of rate compatibility. Rate compatibility requires that a higher-rate code be a subset of a lower-rate code, or that lower-protection codes be embedded into higher-protection codes. This is accomplished by puncturing a "mother" code of rate $1/n$ to achieve higher rates. One major benefit of these RCPC codes with the same mother code is that they all can be decoded by the same Viterbi decoder [7].

Using RCPC codes allows us to utilize Viterbi's upper bounds on the bit error probability, P_b, given by

$$P_b \leq \frac{1}{P} \sum_{d=d_{free}}^{\infty} c_d P_d \tag{1}$$

where P is the period of the code, d_{free} is the free distance of the code, c_d is the information error weight, and P_d is the probability that the wrong path at distance *d* is selected [7]. An AWGN channel with binary phase-shift keying (BPSK) modulation has a P_d given by

$$P_d = Q\left(\sqrt{\frac{2dR_cE_b}{N_0}}\right) \tag{2}$$

where R_c is the channel coding rate and E_b/N_0 is the energy-per-bit normalized to the single-sided noise spectral density measured in Watts/Hertz.

Ignoring thermal noise and background noise due to spurious interference allows us to assume that N_0 is entirely due to interference from other users in the systems. It is a reasonable assumption that the probability distribution of the interfering users is a zero-mean Gaussian random variable [8]. Since user *i* has an associated power level in Watts, $S_i = E_iR_i$, the energy-per-bit to MAI ratio becomes

$$\frac{E_i}{N_0} = \frac{\frac{S_i}{R_i}}{\sum_{j\neq i}^{N} \frac{S_j}{W_T}}; i = 1, 2, 3, ..., N \tag{3}$$

where E_i is the energy-per-bit, $N_0/2$ is the two-sided noise power spectral density due to MAI in Watts/Hertz, S_i is the power of the user-of-interest in Watts, R_i is the transmitted bit rate in bits per second, S_j is the power of the interfering user in Watts, and W_T is the total bandwidth in Hertz [4]. R_i is taken to be the total bit rate used for source and channel coding. Assuming N users, R_i can be expressed as

$$R_i = \frac{R_{s,i}}{R_{c,i}}; i = 1, 2, 3, ..., N \tag{4}$$

where $R_{s,i}$ is the source coding rate for user i and $R_{c,i}$ is the channel coding rate for user i. Since $R_{s,i}$ has units of bits per second and $R_{c,i}$ is a dimensionless number, R_i will be measured in bits per second. [9].

3. OPTIMAL RESOURCE ALLOCATION

There are two methods we will utilize to optimally allocate the network resources to each user in the system. The constraint for both algorithms is that the chip rate be the same for all CDMA users. The first technique we will employ can be formally stated as follows: Given an overall chip rate, R_{budget}, optimally allocate a source coding rate, R_s, a channel coding rate, R_c, and a power level, *S*, to all users such that either the overall distortion D_{ave} over all users is minimized.

$$\min D_{ave} \text{ subject to } R_{chip} = R_{budget} \tag{5}$$

where R_{chip} is the chip rate for each user and D_{ave} is the resulting expected distortion averaged over all users in the system which is due to both source coding errors and channel errors. Assuming *N* users, D_{ave} is expressed by

$$D_{ave} = \frac{1}{N} \sum_{i=1}^{N} D_{s+c,i} \tag{6}$$

where $D_{s+c,i}$ is the expected distortion for user *i*. The distortion due to source coding is a result of the quantization process and is deterministic. However, the distortion due to channel errors is stochastic. Thus, the total distortion for each user is also stochastic, and we use its expected value. The second method we will use to allocate resources to the users on the network minimizes the maximum distortion.

$$\min\{\max D_{s+c,i}\} \text{ subject to } R_{chip} = R_{budget} \tag{7}$$

where R_{chip} is the chip rate for each user. Our constraint is that the chip rate be the same for all CDMA users. The problem is a discrete optimization problem, that is, $R_{s,i}$, $R_{c,i}$, and S_i can only take values from discrete sets $\mathbf{R_s}$, $\mathbf{R_c}$, and $\mathbf{S}$, respectively, i.e., $R_{s,i} \in \mathbf{R_s}$, $R_{c,i} \in \mathbf{R_c}$, $S_i \in \mathbf{S}$ [9].

Since it would be prohibitively complex to experimentally obtain the expected distortion for each user for all possible combinations of source coding rate, channel coding rate, and power level, we instead have chosen to relax the optimality of the algorithm and utilize URDCs. These characteristics show the expected distortion as a function of the bit error probabilities, P_b, after channel coding. P_b is calculated using equations (1)-(4) for the set of channel coding rates and power levels. It acts as a reference for the performance of channel coding over the specified channel with the given parameters. An illustration plot showing the performance of channel coding as a function of given channel parameters is shown in Fig. 1 for a set of channel coding rates R_{cn},$n = 1, ..., N$ We use these plots, called channel characteristic plots, that show the channel bit error probability as a function of the channel parameters in conjunction with URDCs.

As in [9] and [10], we assume the following model for the URDC for each user *i*

$$D_{s+c,i} = a \left[\log_{10}\left(\frac{1}{P_b}\right)\right]^b \tag{8}$$

where *a* and *b* are such that the square of the approximation error is minimized. Thus, instead of calculating the URDCs based on experimental results for every possible P_b, we instead experimentally calculate the expected distortion for a few bit P_b's. We then use the model, given in (8), to approximate the distortion for other bit error rates and power levels. The distortion for a particular user, $D_{s+c,i}$, given a particular source coding rate, $R_{s,i}$, is a function of the bit error rate. Therefore, URDCs will give a family of $D_{s+c,i}$ versus $1/P_b$ curves given a set of source coding rates for each type of user. An illustration plot is shown in Fig. 2 [10] [11].

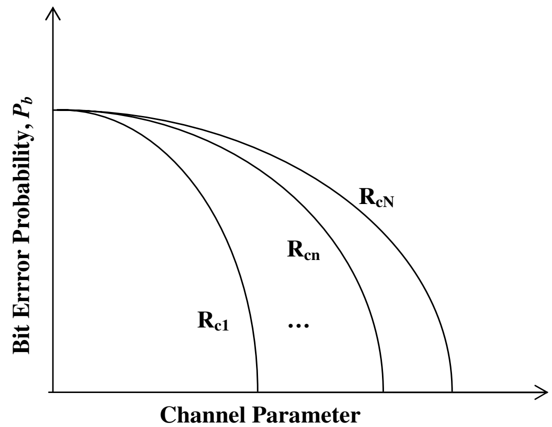

Fig. 1. Channel Characteristic Plots

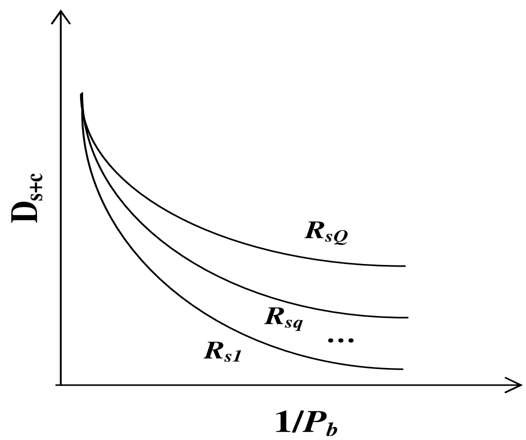

Fig. 2. Universal Rate-Distortion Characteristics

4. EXPERIMENTAL RESULTS

We performed the optimization procedure discussed in Section 3 using the proposed model for URDCs. The data points used to obtain the parameters *a* and *b* were obtained by corrupting the video bitstream with errors based on a calculated P_b, decoding the corrupted video bitstream with the MPEG-4 codec, calculating the distortion, repeating this experiment 100 times and then taking the average distortion. We assume that there are two possible motion levels viewed by the sensor nodes, low motion and high motion. The *Akiyo* sequence is used to represent the low-motion node, and the *Foreman* sequence is used to represent the high-motion node. It is necessary to have two sets of URDC curves, one for each level of motion. The characteristics were obtained for both video sequences at a frame rate of 10 f/s.

We used BPSK modulation, and RCPC codes with mother code rate 1/4 from [7] for channel coding. We assume the same processing gain for all users, so our constraint that the chip rate be the same for all CDMA users translates into a constraint on the transmitted bit rate given in (4). The constraint on the target bit rate was set at 384000 bits per second. The set of admissible source coding rates are $R_{s,i} \in \{128kbps, 192kbps, 256kbps\}$ and the corresponding set of channel coding rates to be $R_{c,i} \in \{1/3, 1/2, 2/3\}$. The power levels in Watts were chosen from $S_i \in \{5, 10, 15\}$. The total bandwidth, W_t, was set to 98MHz.

In Tables 1 through 5, we show what resources should be assigned to the two types of users for various combinations of users with *Akiyo's* parameters represented by R_{s1}, R_{c1}, and S_1 and

Foreman's parameters represented by R_{s2}, R_{c2}, and S_2. "MAD" corresponds to the the method of Minimizing the Average end-to-end Distortion over all users, and "MMD" corresponds to the technique of Minimizing the Maximum Distortion.

We see that in all cases, *Foreman* users are assigned a source coding rate of 256kbps. This is because the drop in the end-to-end distortion when increasing the source coding rate for a high-motion video is much more significant than the effect of employing stronger channel coding. However, the distortions for the *Akiyo* sequence remain relatively low even when the source coding rate is decreased, so it can afford to transmit at a lower source coding rate in some cases. Since low-motion video sequences are more robust to errors, *Akiyo* users are assigned a lower power than *Foreman* even when they both have the same channel coding rate.

	R_{s1}	R_{c1}	S_1	R_{s2}	R_{c2}	S_2	D_{s+c}
MAD	256k	2/3	10	256k	2/3	15	6.276
MMD	256k	2/3	5	256k	2/3	15	6.425

Table 1. 45 *Akiyo* users and 45 *Foreman* users

	R_{s1}	R_{c1}	S_1	R_{s2}	R_{c2}	S_2	D_{s+c}
MAD	256k	2/3	10	256k	2/3	15	7.600
MMD	192k	1/2	5	256k	2/3	15	7.712

Table 2. 20 *Akiyo* users and 70 *Foreman* users

	R_{s1}	R_{c1}	S_1	R_{s2}	R_{c2}	S_2	D_{s+c}
MAD	256k	2/3	10	256k	2/3	15	4.978
MMD	256k	2/3	5	256k	2/3	15	5.026

Table 3. 70 *Akiyo* users and 20 *Foreman* users

5. CONCLUSIONS

In this paper, we have presented two methods for assigning a source coding rate, R_s, a channel coding rate, R_c, and a power level, S, to each user transmitting a video sequence in a wireless visual sensor network. Our system assumes that all noise experienced by a user is due entirely to the interference due to other users in the system. To create a realistic DS-CDMA visual sensor network, different levels of motion were assumed to be imaged by the nodes. By utilizing the parametric model for the URDCs, we found each user's expected distortion for only a small number of source coding rates, channel coding rates, and power levels and used the model to estimate the distortion for other rates and power levels. This reduced the computational complexity of the solution significantly. We presented the combinations of $\{R_s, R_c, S\}$ for each user that result in the minimal average end-to-end distortion over all users in the system and the combinations that minimize the maximum distortion.

6. REFERENCES

[1] K. Ramchandran and M. Vetterli, *Multiresolution Joint Source-Channel coding for Wireless Channels,* chapter in Wireless Communications: A Signal Processing Perspective, V. Poor and G. Wornell ed. Prentice-Hall, 1998.

[2] L. P. Kondi, D. Srinivasan, D. A. Pados, and S. N. Batalama, "Layered Video Transmission over Multirate DS-CDMA Wireless Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1629–1637, December 2005.

[3] L. C. Yun and D. G. Messerschmitt, "Variable quality of service in CDMA systems by statistical power control," in *Proc. ICC*, vol. 2, Seattle, WA, 1995, pp. 713–719.

[4] Y. S. Chan and J. W. Modestino, "A Joint Source Coding-Power Control Approach for Video Transmission Over CDMA Networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1516–1525, December 2003.

[5] G. D. Forney, "The Viterbi Algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[6] J. B. Cain, G. C. Clark, and J. M. Geist, "Punctured convolutional codes or rate (n-1)/n and simplified maximum likelihood decoding," *IEEE Transactions on Information Theory*, vol. IT-25, pp. 97–100, 1979.

[7] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Transactions on Communications*, vol. 36, no. 4, pp. 389–400, April 1988.

[8] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, J. L. A. Weaver, and C. E. W. III, "On the Capacity of a Cellular CDMA System," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 2, pp. 303–312, May 1991.

[9] L. P. Kondi, F. Ishtiaq, and A. K. Katsaggelos, "Joint Source-Channel Coding for Motion-Compensated DCT-Based SNR Scalable Video," *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 1043–1052, September 2002.

[10] M. Bystrom and T. Stockhammer, "Modeling of operational distortion-rate characteristics for joint source-channel coding of video," in *Proc. of the IEEE International Conference on Image Processing*, Kobe, Japan, 1999, pp. 359–362.

[11] L. P. Kondi and A. K. Katsaggelos, "Joint source-channel coding for scalable video using models of rate-distortion functions," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, Utah, 2001, pp. 1377–1380.

	R_{s1}	R_{c1}	S_1	R_{s2}	R_{c2}	S_2	D_{s+c}
MAD	256k	2/3	10	256k	2/3	15	8.404
MMD	128k	1/3	5	256k	2/3	15	8.441

Table 4. 5 *Akiyo* users and 85 *Foreman* users

	R_{s1}	R_{c1}	S_1	R_{s2}	R_{c2}	S_2	D_{s+c}
MAD	256k	2/3	5	256k	2/3	10	4.212
MMD	256k	2/3	5	256k	2/3	15	4.221

Table 5. 85 *Akiyo* users and 5 *Foreman* users