A NOVEL JSCC SCHEME FOR SCALABLE VIDEO TRANSMISSION OVER MIMO SYSTEMS

Xinggong Zhang, Chao Zhou, Zongming Guo †

Institute of Computer Science & Technology
Peking University, Beijing, 100871, P.R. China
Email: {zhangxg, zhouchaoyf, guozongming}@pku.edu.cn

ABSTRACT

MIMO recently emerges as one of promising techniques for wireless video streaming. It is still a challenge to provide un-equal error protections by joint source-channel coding (JSCC) over multiple diverse MIMO sub-channels. In this paper, a joint source-channel coding and antenna mapping scheme for scalable video transmission over MIMO systems is proposed. Bandwidth are elaborately allocated between video source and channel protections by layer extracting and FEC coding. For the extracted layers, we determine *i*) which antenna will they be transmitted over and *ii*) how much redundancy bits will be added for error protections. We formulate this scheme into a non-linear integer optimization problem, whose complexity is very high. Instead, a low-complexity branch-and-bound algorithm is presented. Source layers are partitioned into subsets of layers, and the selected layer are mapped to antennas using Min-max scheduling algorithm. By branching and pruning, the computation complexity are reduced significantly. We carry out extensive numerical experiments under various network conditions. The results demonstrate our algorithm's efficiency and the overall transmission quality is improved significantly.

1. INTRODUCTION

Recent years have seen a rapid growth of video communication over wireless networks. Multi-input multi-output (MIMO) [1] systems, which have been investigated to simultaneously transmit multiple bits streams to achieve high data rate, have emerged as one of the most prominent techniques. By spatial multiplexing, MIMO channel can be decomposed into series of independent sub-channels [2], which generally have various signal-to-noise ratio (SNR) strength and bandwidth. It has attracted many research interests [3, 4] on how to exploit channel diversity of MIMO to improve quality-of-service (QoS) for wireless video transmission.

*This work was supported by National Natural Science Foundation of China under contract No. 61071082, National Basic Research Program of China under contract No. 2009CB320907 and Doctoral Fund of Ministry of Education of China under contract No. 20090001120027.

†Corresponding Author

There are mainly two challenges to provide un-equal error protections (UEP) for video streaming over MIMO systems:

- How to do joint source-channel coding (JSCC) over multiple MIMO sub-channels with various SNR? Since each sub-channel has different capacities and bit error rates (BER), it is a challenge to allocate source rate and FEC redundancies over multiple sub-channels.
- How to map video layers to MIMO antennas (sub-channels)? For diverse sub-channels and prioritized video layers, a proper layer-antenna mapping can provide better resource provisions.

In many pioneering works, channel diversity of MIMO have been exploited for wireless video streaming. Song and Chen *et al.* [4] has proposed a SNR-based antenna mapping scheme to protect scalable H.264/SVC video layers unequally. They assign good sub-channels (with higher SNR) to more important layers. But they did not consider channel protection with JSCC. On the other hand, there are also many literatures [3, 5] about UEP for scalable video streaming. They consider a video with temporal-quality scalability and employ a genetic algorithm (GA) [5] to allocate FEC redundancies among layers. But they ignore the benefits of MIMO channel diversities.

In this paper, we attempt to combine the advances of JSCC and channel diversity together, and to provide unequal error protections for scalable video transmission over MIMO wireless networks. The source rate is adapted by selectively extracting a suitable subset of video layers, i.e., dropping some video layers in the source. Then, for the selected layers, UEP is provided by adding different FEC redundancies and being mapped to appropriate antennas. By finding the optimal layer extracting, FEC coding and antenna mapping scheme, the end-to-end distortion is minimized. We formulate this scheme as an integer optimization problem, whose computation complexity is very higher. Instead, a low-complexity branch-and-bound algorithm is proposed. The original problem is partitioned into some simple sub-problems by layer extracting. While for each sub-problem, a min-max scheduling scheme [6] is employed to obtain the optimal redundancy allocation and antenna mapping. With the bounding and pruning techniques, the computation complexity is greatly reduced. At last, the experimental results show that our proposed algorithm is efficient, and the quality of reconstructed video is improved significantly compared with the existing schemes.

The paper is organized as follows: Section 2 formulates the problem. In Section 3, we present a branch-and-bound solution. Finally, we illustrate the simulation results in Section 4 and conclude the paper in Section 5 respectively.

2. PROBLEM FORMULATION

Consider a $N \times N$ MIMO wireless system that both the transmitter and receiver are equipped with N antennas. According to the singular value decomposition (SVD) technique [2], this MIMO channel can be decomposed into N sub-channels. Assume that the channel condition matrix is known perfectly at the receiver end using ML or MMSE estimation methods [1], thus the bit error ratio (BER) of sub-channels can be easily estimated using pilot symbols. We denote $\mathbf{p} = [p_1, p_2, ..., p_N]$ as the BER vector, where p_i denotes the i-th sub-channel's BER. We further assume that the bandwidth of sub-channels is known perfectly at the transmitter, which is denoted as the vector $\mathbf{R} = [R_1, R_2, ..., R_N]$.

Meanwhile, current H.264/SVC video coding standard provides temporal, spatial, and SNR quality scalabilities. In this paper, we only consider the scalability in temporal and SNR quality dimensions. Assume that a video sequence is encode into L layers, the additive quality and rate of each layer is denoted with w_l and r_l respectively.

We now define a layer-antenna mapping matrix $\mathbf{X}$ as

$$\mathbf{X}_{L\times N} = \begin{pmatrix} x_{1,1} \cdots x_{1,N} \\ \vdots \quad \ddots \quad \vdots \\ x_{L,1} \cdots x_{L,N} \end{pmatrix} \quad (1)$$

where $x_{l,n} \in \{0,1\}$. Let $x_{l,n} = 1$ if the l-th video layer is transmitted over the n-th sub-channel, and $x_{l,n} = 0$ otherwise. Noting that for any video layer, it would be either transmitted over only one certain sub-channel, or dropped at the transmitter. On the other hand, multiple video layers can be transmitted over the same sub-channel. For this multi-to-one mapping problem, we have the following condition:

$$\sum_{n=1}^{N} x_{l,n} = \{0,1\} \;\; 1 \le \forall l \le L \quad (2)$$

Specially, if $\sum_{n=1}^{N} x_{l,n} = 0$, it denotes that the l-th video layer is dropped at the transmitter, and the source rate is reduced for allocating more bandwidth to channel protections.

Then we define $\hat{\mathbf{p}}$ as the BER vector after layer-antenna mapping

$$[\hat{\mathbf{p}}]^{\mathrm{T}} = \mathbf{X}\mathbf{p}^{\mathrm{T}} \quad (3)$$

and $[\cdot]^{\mathrm{T}}$ denote the transpose operation. $\hat{p}_l \in \hat{\mathbf{p}}$ denotes the BER of the antenna which the l-th video layer are transmitted over.

We denote the vector of channel coding (FEC) rate as $\mathbf{C} = \{c_l\}$, where c_l is the FEC rate of the l-th video layer. Assume the block length is fixed at M, FEC error probability for the l-th video layer, P_l^{FEC}, is written as

$$P_l^{FEC}(\mathbf{X},\mathbf{C}) = \mathbf{1} - \sum_{\mathbf{k=0}}^{\mathbf{M(1-c_l)}} \binom{M}{k} (\hat{\mathbf{p}}_\mathbf{l})^{\mathbf{k}} \, (\mathbf{1} - \hat{\mathbf{p}}_\mathbf{l})^{(\mathbf{M-k})} \quad (4)$$

Now, considering the dependences among video layers, the overall end-to-end transmission distortion can be written as follows by applying the distortion model in [7]

$$D(\mathbf{X},\mathbf{C}) = \sum_{l\le L} w_l (1 - \prod_{i\le l} \left(1 - P_i^{FEC}(\mathbf{X},\mathbf{C})\right)) \quad (5)$$

At last, we formulate our JSCC scheme as

$$\begin{aligned} &\{\mathbf{X}^*, \mathbf{C}^*\} = \arg\min_{\{\mathbf{X},\mathbf{C}\}} D(\mathbf{X},\mathbf{C}) \\ &\text{s.t.} \\ &\sum_{l\le L} \frac{r_l}{c_l} x_{l,n} \le R_n, \; 1 \le \forall n \le N \\ &Equ.(5) \\ &Equ.(4) \\ &Equ.(2) \\ &x_{l,n} \in \{0,1\}, \; 1 \le \forall l \le L, 1 \le \forall n \le N \end{aligned} \quad (6)$$

The objective of this optimization problem is to minimize the total distortion by video source adaptation, FEC rate allocation and antenna mapping. The first constraint corresponds to the bandwidth constraint of each sub-channel, and the rest are the constraints of antenna mapping . This is a non-linear integer optimization problem, it is well known that this problem is NP-hard [8].

3. SOLUTION ALGORITHM

The above optimization problem can be solved using a exhaustive search method, which searches over the whole solution space with size $M^N(N+1)^L$, where M is the total number of FEC rates. In order to reduce the computational cost and make it suitable for practical implementation, we propose a heuristic low-complexity branch-and-bound algorithm.

3.1. Branching and pruning

Branch-and-bound is an iterative method for solving optimization problems, especially for discrete and combinatorial problems. A important component in the branch-and-bound procedure, called *branching*, is to partition a problem into subproblems. Here, we also employ it to partition the original video source layers into mutually exclusive sub-sets of layers. Each branch, which is represented by a node, is a set of extracted layers for transmission. For one node, the bound of distortion is obtained by antenna mapping and FEC coding.

We consider a scalable video source with quality-temporal scalability such as in Fig. 1. The temporal scalability results in a divergence from existing rate-distortion models, since the quality metric by PSNR is not enough to present the quality contribution of frame rates changes. Thus, we adopt a subject quality model and rate model proposed by Yao [7] to weight the quality contributions w_l of video layers.

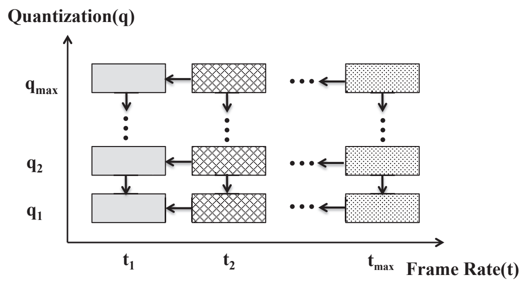


Fig. 1. Branch-tree for quality-temporal scalable video.

Considering the dependence among video layers as in Fig. 1, the *branching* is implemented to extract the sub-sets of layers for transmission. Starting from the layer with the maximal frame rate index t_{max} and quantization-step index q_{max}, we iterate every combination of quantization steps-size q and frame rate t. For a node with index (t, q), it includes all (i, j)-th video layer where $1 \leq \forall i \leq t, 1 \leq \forall j \leq q$. And it has two child nodes, $(t-1, q)$ and $(t, q-1)$. Then we obtain a *branch-and-bound tree*, and each node denotes a sub-set of extracted layers.

To reduce the search space, *pruning* is implemented. For a node with index (t, q), we use a heuristic method as in Section 3.2 to find a feasible solution. If founded, the solution is considered as the upper-bound of the node, and it is no need to be branched any more. Otherwise, this node is branched to two child nodes $(t - 1, q)$ and $(t, q - 1)$. The procedures are repeated recursively until there is no more nodes to be branched. The node with the least upper-bound is the optimal solution.

3.2. Upper-bounding

For a node in the branch-and-bound tree, the subset of extract layers for transmission is determined. We use a heuristic algorithm to map the layers to appropriate sub-channels, and obtain the end-to-end transmission distortion by Equ. (5).

Assume that FEC coding can achieve the theory upper-bound of Shannon limits, we are able to treat the n-th lossy sub-channel as a lossless channel with the effective bandwidth:

$$\mathcal{R}_n = R_n * (1 - p_n) \tag{7}$$

By this approximation, the antenna mapping problem is converted to a **knapsack loading problems**: there are total $\mathcal{L}$ layers and N channels, each layer has different quality weight w_l and rates r_l. The problem is how to put the layers into the channels so that the sum of weight for the layers in the channels is maximized.

$$\{\mathbf{X}^*\} = \arg\min_{\{\mathbf{X}\}} \sum_{l \leq \mathcal{L}} w_l \sum_{n \leq N} x_{l,n}$$

s.t.

$$\begin{aligned} & \sum_{l \leq \mathcal{L}} r_l x_{l,n} \leq \mathcal{R}_n, \ \forall 1 \leq n \leq N \\ & \sum_{n=1}^{N} x_{l,n} \leq 1, \ \forall 1 \leq l \leq \mathcal{L} \\ & x_{l,n} \in \{0, 1\}, \ \forall 1 \leq l \leq \mathcal{L}, 1 \leq n \leq N \end{aligned} \tag{8}$$

This is an integer combinational optimization problem, whose search space is $(N+1)^{\mathcal{L}}$. To reduce the computation complexity further more, a classical min-max scheduling scheme[6] is employed. We first sort all layers in a descendent order of rates, then select the channel with maximal free bandwidth, and put the layers into channels one by one. If all video layers can be put into sub-channels under the bandwidth constraint as in Equ. (8), it is a feasible solution. Otherwise it is not a feasible solution and will be branched.

From the above analysis we can see that the deleted nodes are either unfeasible solutions or non-optimal solutions. Then in the rest feasible solutions, we can find a sub-optimal solution by Equ. (6).

4. EXPERIMENTAL RESULTS

In this section, we present illustrative simulation results to shows both the effectiveness and efficiency of our proposed scheme. The video sequences, *akiyo*, *city*, *crew*, and *football*, are selected for testing for their different characteristic motion and spatial details. All test sequences are encoded using JSVM-9.12 [9] to generate embedded scalable video streams with 25 layers (5 quantization steps and 5 frame rates), and transmitted over a 4×4 MIMO system.

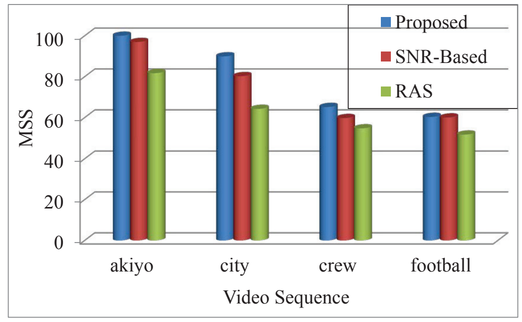
First, we demonstrate the *computation complexity* of our proposed branch-and-bound algorithm and the exhaustive search method (ESM). Under the same network conditions (channels' bandwidth is $400kbps$ and BER is 20%), ESM and the proposed are both implemented to find the optimal solution of Equ. (6). The number of iterations are presented in Table 4. ESM searches all solution space, but with awesome complexity. While by branching and pruning techniques, our proposed algorithm find the solutions with much lower complexity.

We also compare the three JSCC schemes in video transmission over MIMO: random antenna selection (RAS), SNR-based mapping, and our proposed algorithm. RAS maps video layers to antennas randomly, while SNR-based method assigns good channels to important layers. And both of them

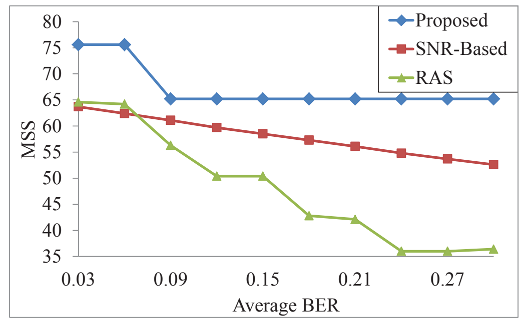
Table 1. Number of iterations

	Exhaustive Search	Branch-and-bound
akiyo	4^{25}	100
city	4^{25}	132
crew	4^{25}	160
football	4^{25}	160

do JSCC respectively by the resources provided by antenna mapping. The mean subject quality score (MSS) of reconstructed video is illustrated in Fig. 2. First, we compare the three schemes fort various video sequences under the same network conditions. The total bandwidth of MIMO is set to $400kbps$ and average BER is 20%. As showed in Fig. 2 (b), the proposed algorithm always outperforms other methods for all sequences. It is also noted that the sequence *Akiyo* achieves the best quality for all schemes. Since the source needs the least bit rates under the same distortion, it is true that *akiyo* has better protections compared with others. We also compared the three schemes under various MIMO channel conditions for *akiyo* . As showed in Fig. 2 (b), the quality of our proposed algorithm keeps stable, while for the other two methods, the quality is dropped with increased BER. The proposed method is able to resists channel errors until the BER is lager than its tolerances. Then it would find a new optimal antenna mapping and do JSCC under the new channel conditions. But for RAS and SNR-based method, the scheme of antenna mapping remains unchanged and the increased BER deteriorate the protections of JSCC.



(a) Four sequences in the same MIMO channels



(b) *akiyo* under various network conditions

Fig. 2. Quality comparison of the three schemes.

5. CONCLUSION

In this paper, we have investigated the problem of joint source-channel coding and antenna selection for scalable video transmission over MIMO wireless networks, and a novel JSCC scheme is proposed. In this scheme, the source rate is control by selectively transmitting, i.e., dropping part of video layers in the source. Then, for the selected layers, we find the optimal solution that which antenna will the layers are transmitted over and how much redundancy bits should be added in the channel. This optimal problem is an integer optimization problem, and a low-complexity branch-and-bound algorithm is proposed to solve it. The experimental results demonstrate our algorithm's efficiency.

References

[1] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An Architecture for Realizing very High Data Rates over the Rich-Scattering Wireless Channel," *IEEE Int. Symp. on Signals, Syst. Electron.*, pp. 295–300, Sep. 1998.

[2] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *Eur. Trans. Commun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[3] M. K. Jubran, M. Bansal, and L. P. Kondi, "Low-Delay Low-Complexity Bandwidth-Constrained Wireless Video Transmission Using SVC Over MIMO Systems," *IEEE Trans. Multimedia*, vol. 10, no. 8, pp. 1698 –1707, Dec. 2008.

[4] Daewon Song and Chang Wen Chen, "Scalable H.264/AVC Video Transmission Over MIMO Wireless Systems With Adaptive Channel Selection Based on Partial Channel Information," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 17, no. 9, pp. 1218 –1226, Sep. 2007.

[5] Tao Fang and Lap-Pui Chau, "GOP-Based Channel Rate Allocation Using Genetic Algorithm for Scalable Video Streaming over Error-Prone Networks," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1323–1330, Jun. 2006.

[6] Tracy D. Braun, Howard Jay Siegel, and Noah Beck, "A Comparison of Eleven Static Heuristics for Mapping a Class of Independent Tasks onto Heterogeneous Distributed Computing Systems," *J. Parallel Distrib. Comput.*, vol. 61, no. 6, pp. 810–837, Jun. 2001.

[7] Yao Wang, Zhan Ma, and Yen-Fu Ou, "Modeling Rate and Perceptual Quality of Scalable Video as Functions of Quantization and Frame Rate and its Application in Scalable Video Adaptation," *in Proceedings of Packet Video Workshop*, pp. 1–9, May. 2009.

[8] Christos H. Papadimitrious and Kenneth Steiglitz, *Combinatorial Optimization: algorithm and complexity*, Dover Publications Inc., Mineola, New York, 1998.

[9] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 17, no. 9, pp. 1103 –1120, Sep. 2007.