# Precoded OFDM with Adaptive Vector Channel Allocation for Scalable Video Transmission over Frequency-Selective Fading Channels

Hong Zhang*, Xiang-Gen Xia*, Qian Zhang[1], Wenwu Zhu[1]


## Abstract

The precoded orthogonal frequency division multiplexing (OFDM) system with adaptive vector channel allocation can provide more protection against errors to more important layers in the layered bit stream structure of scalable video. The vector channel allocation information feedback to the transmitter end is the vector channel indexes. Simulation results have demonstrated the robustness of the proposed scheme to noise and fading inherent in wireless channel.


**Keywords**: precoded OFDM, vector channel allocation, scalable video, video transmission over wireless channel, unequal error protection

## 1 Introduction

The Internet is now incorporating broadband wireless communication technology to provide wireless access to multimedia service. The third generation (3G) wireless communication technology has made it feasible for video transmission over wireless channels by affording data rate from 384kbps to 2Mbps. However, video has to be efficiently compressed to save the precious bandwidth, which at the same time renders the compressed video bit stream extremely vulnerable to the error-prone wireless channel.

It is well known that different portions of the compressed video bit stream have different importance to the quality of the reconstructed video. Therefore, the idea of unequal error protection (UEP) fits well into the framework of video transmission [1]. Several schemes combining scalable video coding with UEP for wireless application have been proposed recently [2, 3]. In [2], Cheung proposed an optimal bit allocation scheme for joint source channel coding of scalable video. Rate compatible punctured convolutional (RCPC) code and UEP are combined in this scheme. The frequency-selective fading characteristics of wireless channels were not seriously considered in this case. In [3], Wang proposed to periodically allocate the available protection bits among different video layers based on channel condition feedback. In practice, the channel condition might not be readily available.

OFDM scheme has been known as an efficient way to deal with multipath in frequency-selective channel [4]. It is also possible to enhance the capacity by optimizing the data rate per subcarrier according to the SNR condition. It has been widely used in wireline broadband communication systems such as ADSL. For frequency-selective Rayleigh fading channel, the pre-coded OFDM scheme [5] has been proposed recently to offer more robustness to spectral nulls and

* The authors are with Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, {hzhang, xxia}@ee.udel.edu. Their work was performed when they were visiting Microsoft Research, Beijing, China. Hong Zhang was a summer intern.

[1] The authors are with Microsoft Research, 49 Zhichun Road, Sigma Building, Beijing 100080, China. {zqian, wwzhu}@microsoft.com

reduction of the length of cyclic prefix. The precoded OFDM system is now enhanced with adaptive vector channel allocation information being feedback to the transmitter end. The feedback information consists of channel index sets rather than values of channel coefficients. When applied to the transmission of multimedia over wireless channel, it can provide more protection against errors to more important layers in the layered bit stream structure of scalable video, and thus renders the proposed scheme more robust to noise and fading inherent in wireless channels.

The rest of this paper is organized as follows. Section 2 briefly describes the bit stream structure of several scalable video coding schemes and emphasizes the need to provide unequal protection to different layers in the layered structure. Section 3.1 summarizes the precoded OFDM scheme in [5]. In Section 3.2, adaptive vector channel allocation is proposed to fit into the precoded OFDM scheme and to provide unequal protection to the layered bit stream of scalable video. Simulation results are given in Section 4 to demonstrate the effectiveness of the proposed scheme.

## 2 Scalable Video Coding and UEP

Several scalable video coding schemes [6, 7] have been developed in response to the growing need on streaming video over the Internet, where the network channel capacity varies over a wide range depending on the type of connections and the network traffic at any given time. The fine granularity scalable (FGS) video coding scheme in MPEG-4 encodes video frames into multiple layers, including a base layer of video of relatively lower quality and several enhancement layers containing increasingly more detailed video data to enhance the base layer and thus resulting in video of increasingly higher quality. By predicting all enhancement layers form the base layer, losses or corruption of one or more enhancement layers during transmission have no effect on the frames followed. The recently proposed progressive fine granularity scalable (PFGS) video coding scheme [7] uses as many predictions as possible from the enhancement reference layers to improve coding efficiency over FGS. It keeps a prediction path from the base layer to the highest quality layers across several frames and thus guarantees graceful recovery from any enhancement layer errors. Erroneous higher quality enhancement layers can be automatically reconstructed from lower layers gradually over a few frames with such a prediction path.

Both of these scalable video coding schemes feature a layered bit stream structure with different importance at different layers. The base layer bit stream is very sensitive to channel errors. In the decoder, if any error were detected in the lower layers, the higher layers of the current frame would be discarded whether they are correct or not. The undetected errors in the base layer would propagate to the start of the next group of picture (GOP) and would cause serious drifting problems in the higher layers of the following predicted frames. Therefore the base layer is essential for decoding of the video streams and has to be encoded in a way to fit in the minimum channel bandwidth and transmitted in a well-controlled channel. Otherwise, the quality of the reconstructed video would be severely degraded. On the other hand, the enhancement layers can tolerate channel errors. Errors in the enhancement layers would corrupt video quality only in the current frame and several subsequent frames and the deterioration is not as appreciable as that caused by errors in the base layer. In this layered bit stream structure, the more important layers should be provided with more protection against errors compared with the less important layers.

## 3 Precoded OFDM with Adaptive Vector Channel Allocation

In this section, we first briefly describe the precoded OFDM system proposed in [5] and then present our adaptive vector channel allocation for video transmission.

### 3.1 Review of Precoded OFDM

In the precoded OFDM system, the data stream $x(n)$ is first blocked into $K \times 1$ vector sequence

$$\mathbf{x}(n) = [x_0(n), x_1(n), \cdots, x_{K-1}(n)]^T$$

where $[\ ]^T$ denotes transpose and $x_k(n) = x(Kn + k)$, $k = 0,1,\cdots,K-1$. Then each vector $\mathbf{x}(n)$ is precoded simply by appending $M - K$ zeros to its end to form a $M \times 1$ vector $\hat{\mathbf{x}}(n)$. The precoded $M \times 1$ vector sequence $\hat{\mathbf{x}}(n)$ is further blocked into N-vector subsequences

$$\hat{\mathbf{X}}(n) = [\hat{\mathbf{x}}_0(n), \hat{\mathbf{x}}_1(n), \cdots, \hat{\mathbf{x}}_{N-1}(n)]$$

where $\hat{\mathbf{x}}_k(n) = \begin{bmatrix}\mathbf{x}_k^T(n) & \mathbf{0}_{(M-K)\times 1}\end{bmatrix}^T$ and $\mathbf{x}_k(n) = \mathbf{x}(Nn + k)$. Let the output of the N-point IFFT of $\hat{\mathbf{X}}(n)$ be denoted by

$$\hat{\mathbf{Z}}(n) = [\hat{\mathbf{z}}_0(n), \hat{\mathbf{z}}_1(n), \cdots, \hat{\mathbf{z}}_{N-1}(n)]$$

then

$$\hat{\mathbf{z}}_l(n) = \tfrac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{\mathbf{x}}_k(n) \exp(j2\pi kl / N),\ l = 0,1,\cdots,N-1$$

The cyclic prefix is to append the last $\hat{\Gamma}$ vectors $\hat{\mathbf{z}}_l(n)$, $l = N - \hat{\Gamma}, N - \hat{\Gamma} + 1, \cdots, N-1$ to the front of $\hat{\mathbf{Z}}(n)$ and thus form $(N + \hat{\Gamma})$-vector subsequences

$$\mathbf{Z}(n) = [\hat{\mathbf{z}}_{N-\hat{\Gamma}}(n), \cdots, \hat{\mathbf{z}}_{N-1}(n), \hat{\mathbf{z}}_0(n), \cdots, \hat{\mathbf{z}}_{N-1}(n)]$$

Pack $\mathbf{Z}(n)$ into $M(N + \hat{\Gamma}) \times 1$ vector

$$\mathbf{z}(n) = [\hat{\mathbf{z}}_{N-\hat{\Gamma}}^T(n), \cdots, \hat{\mathbf{z}}_{N-1}^T(n), \hat{\mathbf{z}}_0^T(n), \cdots, \hat{\mathbf{z}}_{N-1}^T(n)]^T$$

which is called an OFDM block later on. The vectors $\mathbf{z}(n)$ are transmitted as scalar sequence $z(n)$ through parallel-to-serial conversion.

At the receiver, inverse operation is taken to extract the original data. Let $r(n)$ be the received scalar sequence at the receiver, which is converted into $M(N+\hat{\Gamma})\times 1$ vectors and unpacked into subsequences of $N+\hat{\Gamma}$ $M\times 1$ vectors, similar to the structure of $\mathbf{Z}(n)$. The last $N$ vectors $\hat{\mathbf{r}}_k(n), k=0,1,\cdots,N-1$ are inputs to the N-point FFT and the outputs are

$$\hat{\mathbf{y}}_k(n)=\tfrac{1}{\sqrt{N}}\sum_{l=0}^{N-1}\hat{\mathbf{r}}_l(n)\exp(-j2\pi kl/N),\ k=0,1,\cdots,N-1$$

Since a single-input-single-output (SISO) linear time-invariant system with transfer function $H(z)$ is equivalent to a multiple-input-multiple-output (MIMO) system by blocking of the input and output of the SISO system into $M\times 1$ vectors. The equivalent MIMO transfer function is the blocked version of $H(z)$, denoted by $\mathbf{H}(z)$ [5]. Let $\mathbf{x}(z)$ and $\mathbf{y}(z)$ denote the z-transforms of the input vectors and output vectors resulted from blocking the input and the output of the SISO system, respectively, then $\mathbf{y}(z)=\mathbf{H}(z)\mathbf{x}(z)$. When the order of $H(z)$ is $L$, the order of $\mathbf{H}(z)$ is $\tilde{L}=\lceil \tfrac{L}{M}\rceil$.

When the cyclic prefix length $\hat{\Gamma}$ satisfies $\hat{\Gamma}\geq\hat{L}$, it can be derived, following the same approach as that of conventional OFDM, that in the precoded OFDM system

$$\hat{\mathbf{y}}_k(n)=\mathbf{H}_k(n)\hat{\mathbf{x}}_k(n)+\hat{\boldsymbol{\xi}}_k(n),\ k=0,1,\cdots,N-1$$

where $\mathbf{H}_k(n)=\mathbf{H}_n(z)\big|_{z=\exp(j2\pi k/N)}$, $\mathbf{H}_n(z)$ is constructed from the estimated channel response vector $\mathbf{h}(n)$ for the transmitted OFDM block $\mathbf{z}(n)$ and $\hat{\boldsymbol{\xi}}_k(n)$ is the noise component vector. The noise component is made up of the additive white Gaussian channel noise term and the bias term induced by representing a time-varying channel with a constant estimate $\mathbf{h}(n)$ [5]. Thus its elements might not be identically and independently Gaussian distributed even when the channel noise term could be. When the precoder is used, the above relation actually becomes

$$\hat{\mathbf{y}}_k(n)=\hat{\mathbf{H}}_k(n)\mathbf{x}_k(n)+\hat{\boldsymbol{\xi}}_k(n)$$

where $\hat{\mathbf{H}}_k(n)$ denotes the submatrix consisting of the first $K$ columns of $\mathbf{H}_k(n)$.

Assuming knowledge of $\hat{\mathbf{H}}_k(n)$ is available at the receiver end, which is obtained from the estimate of the channel, a simple and efficient approach to estimate $\mathbf{x}_k(n)$ is to first obtain its least square approximation from

$$\tilde{\mathbf{x}}_k(n)=\hat{\mathbf{H}}_k^+(n)\hat{\mathbf{y}}_k(n)=\mathbf{x}_k(n)+\hat{\mathbf{H}}_k^+(n)\hat{\boldsymbol{\xi}}_k(n)$$

where $\hat{\mathbf{H}}_k^+(n) = (\hat{\mathbf{H}}_k^H(n)\hat{\mathbf{H}}_k(n))^{-1}\hat{\mathbf{H}}_k^H(n)$, then find maximum-likelihood (ML) solution for each element of $\mathbf{x}_k(n)$. When $M$ and $K$ are small, the computational complexity and performance of this approach have been shown to be comparable to those of convolutional coding schemes.

### 3.2 Adaptive Vector Channel Allocation for Scalable Video

From an MIMO system viewpoint, each $\hat{\mathbf{H}}_k(n)$, $k = 0,1,\cdots,N-1$, provides a vector channel for the vector $\mathbf{x}_k(n)$. These channels can be characterized and put into an order in terms of symbol error rate so that when it comes to the transmission of layered bit stream with different importance at different layers, layers with higher importance can be assigned channels with lower symbol error rates. It can be expected that by offering UEP to multiple layers in scalable video bit stream in this way, the received video quality may be more robust to noise and fading. However, the symbol error rate of a particular channel characterized by $\hat{\mathbf{H}}_k(n)$ is not readily related to $\hat{\mathbf{H}}_k(n)$ by a simple formula, due to the bias induced by using constant $\mathbf{h}(n)$ for the whole block of $\mathbf{z}(n)$ and mathematical difficulties in performance analysis of the ML solution. A simplified but reasonable approach is to use Frobenius norm of $\hat{\mathbf{H}}_k^+(n)$, i.e. $\left\|\hat{\mathbf{H}}_k^+(n)\right\|_F$, to characterize the symbol error rate, considering its relation with the energy of the noise component $\hat{\mathbf{H}}_k^+(n)\hat{\boldsymbol{\xi}}_k(n)$. At the receiver end, the channel estimate is used to derive $\hat{\mathbf{H}}_k(n)$ and then $\left\|\hat{\mathbf{H}}_k^+(n)\right\|_F$, $k = 0,1,\cdots,N-1$, are calculated and put into increasing order. The channel indexes corresponding to the ordered norms are feedback to the transmitter end periodically. There the base layer of the scalable video stream is assigned the set of vector channels with the smallest $\left\|\hat{\mathbf{H}}_k^+(n)\right\|_F$'s, the first enhancement layer is assigned the set of vector channels with the smallest $\left\|\hat{\mathbf{H}}_k^+(n)\right\|_F$'s among the unallocated channels, and so on. The number of vector channels to be assigned to each layer depends on the rate control algorithm, which is another topic not to be discussed in this paper.

## 4 Simulation Results

The CIF Foreman sequence is encoded with PFGS coding scheme in the simulation, and the encoding frame rate is 10 Hz. Only the first frame is encoded as I frame, and the other 49 frames are encoded as P frames or B frames. Because of the importance of the base layer, Reed-Solomon (RS) coding and interleaving is applied to offer more protection against burst errors. RS (255, 205) code over GF (256) is used such that every 205 bytes of the base layer information is protected by 50 redundancy bytes, providing error-correcting capability of 25 bytes. The interleaving depth is 200 codewords, not counting into the interleaving effect of FFT. The PFGS decoder has included some basic error-resilience tools and error-concealment tools, such as resynchronization marker and copying from motion-compensated areas in the previous frame, etc. The elements of $x(n)$ are selected from the QPSK constellation. There are six enhancement layers out of the PFGS encoder but only the lowest three would be transmitted. The number of vectors allocated to each layer is calculated offline and ad hoc, based on the average rate of multiple layers. Since the focus of the simulation is the effect of UEP, it is reasonable to take this simplified approach. The sampling rate or the symbol rate at the transmitter is $245973$ /s and $346560$ /s for $K/M = 1/1$ and $K/M = 3/4$, respectively. These would decide the channel bandwidth. The frequency-selective Rayleigh fading channel is simulated by the modified Jake's method [8]. The length of multipath spread is assumed to be a quarter of $MN$, which is 32 or equivalently, $L = 31$ when $MN = 128$. The tap energy is shaped with a peak-like profile featuring lots of scenarios, climbing up fast and tapering off slowly both in an exponential manner, and the sum of the average energies of all taps equals one. Maximum Doppler frequency shift $f_m = 100Hz$ and the key parameter $f_m T_s$ is around the order of $10^{-4}$, which characterizes the channel fading rate over samples [8] and where $T_s$ is the sampling period, the inverse of the sampling rate, decided by video data rate, multipath spread, precoded OFDM scheme parameters $M$ and $N$. The channel matrix $\hat{\mathbf{H}}_k(n)$ is computed from the channel response at the instant when $z(nM(N+\hat{\Gamma}) - MN/2)$ is transmitted. The set of indexes of $\hat{\mathbf{H}}_k(n)$, ordered in terms of $\left\|\hat{\mathbf{H}}_k^{+}(n)\right\|_F$, is used to decide the sets of vector channels for transmitting different layers. This information is feedback to the transmitter end every $F$ OFDM blocks. The smaller $F$ is, the more frequently vector channel allocation information is updated at the transmitter end and thus, the better protection is provided to lower layers. For $K/M = 1/1$, i.e., the conventional OFDM system, the detector is the ML detector.

In Fig. 1, the precoder code rate is $K/M = 3/4$ and similar to convolutional code, the smaller it is the more powerful the precoder is [5]. The PSNR upper bound is obtained when the original bit stream is cut up to the lowest three enhancement layers and decoded without any transmission error. It can be observed that UEP helps to keep the system more robust at low SNR or at high channel fading rate. From Fig. 1, it can also be seen that the precoded OFDM system outperforms the conventional OFDM system. In Fig. 2, the symbol error rate versus layer index for several feedback intervals is shown to demonstrate the effectiveness of the simplified approach to characterize symbol error rate of the vector channel. An interesting observation is that at high SNR, when the symbol error rate is already negligible for the base layer with RS coding protection, it is the enhancement layers that are worth more protection against errors, considering the overall quality of the video and the available vector channels. In this case, schemes with larger UEP information update interval or even without UEP might give more protection to enhancement layers and thus might result in higher, though quite marginal, average PSNR at high SNR. More feedback information might help to solve this problem, which is left for further research.

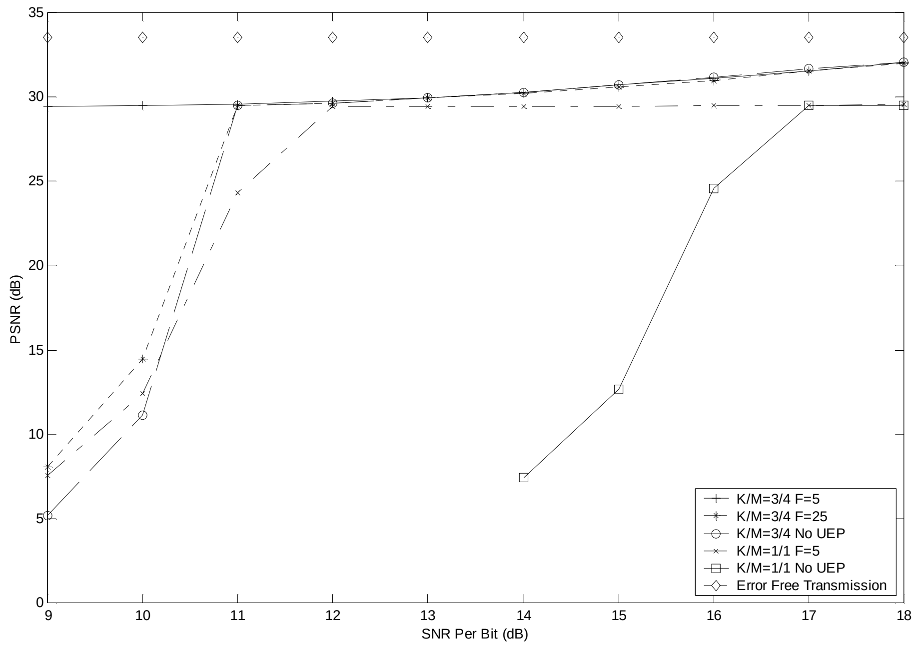


Fig. 1 Average PSNR of Foreman Y component versus SNR per information bit, MN=128 for all and the vector channel allocation information feedback interval is denoted by F (OFDM block).

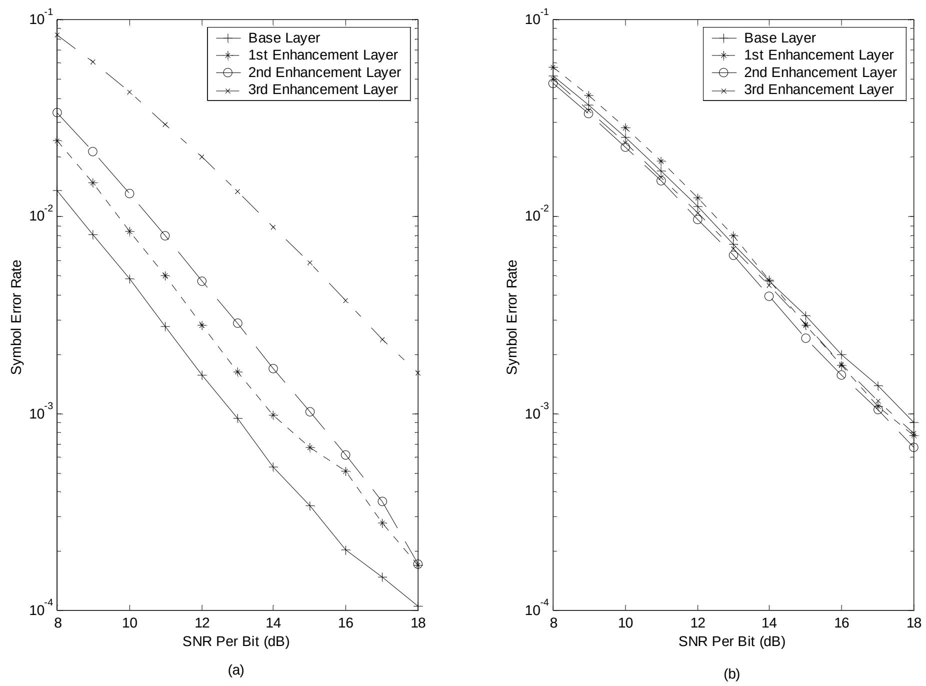


Fig. 2 Symbol error rate of different layers for precoded OFDM scheme with (a) and without (b) adaptive vector channel allocation, F=5 for (a) and K/M=3/4 for both.

## 5 Conclusion

The precoded OFDM system with adaptive vector channel allocation can provide more protection against errors to more important layers in the layered bit stream structure of scalable video. The vector channel allocation information feedback to the transmitter end is vector channel indexes but not coefficient values. Simulation results have demonstrated the robustness of the proposed scheme to noise and fading inherent in wireless channels.

## References


[1] J. F. Cai, Q. Zhang, W. Zhu and C. W. Chen, “An FEC-based error control scheme for wireless MPEG-4 video transmission”, *IEEE WCNC’2000*, July, 2000.

[2] G. Cheung and A. Zakhor, “Bit allocation for joint source/channel coding of scalable video”, *IEEE Trans. Image Processing*, March, 2000.

[3] G. J. Wang, Q. Zhang, W. Zhu and Y.-Q. Zhang, “Channel-adaptive unequal error protection for scalable video transmission over wireless channel”, MSR China Tech. Report .

[4] J. G. Proakis, *Digital Communications*, $4^{th}$ ed. New York: McGraw-Hill, 2001.

[5] X.-G. Xia, "Precoded and vector OFDM systems robust to channel spectral nulls and with reduced cyclic prefix length in single transmit antenna systems," *IEEE Trans. Commun.*, vol.49, pp.1363-1374, Aug. 2001.

[6] W. Li, “Overview of fine granularity scalability in MPEG-4 video standard”, *IEEE Trans. Circuits Syst. Video Technol*., March, 2001.

[7] F. Wu, S. Li and Y.-Q. Zhang, “A framework for efficient progressive fine granularity scalable video coding”, *IEEE Trans. Circuits Syst. Video Technol*., March, 2001.

[8] G. L. Stuber, *Principles of Mobile Communication*, $2^{nd}$ ed. Boston: Kluwer Academic Publishers, 2001.