Journal of Southeast University (English Edition) Vol. 21, No. 4, pp. 405 – 410 Dec. 2005 ISSN 1003—7985

Scalable transcoding for video transmission over space-time OFDM systems

Qian Tuanjie Sun Jun Xie Rong Su Pengcheng Wang Jia

(Institute of Image Communication and Information Processing Research, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: A new scheme combining a scalable transcoder with space time block codes (STBC) for an orthogonal frequency division multiplexing (OFDM) system is proposed for robust video transmission in dispersive fading channels. The target application for such a scalable transcoder is to provide successful access to the pre-encoded high quality video MPEG-2 from mobile wireless terminals. In the scalable transcoder, besides outputting the MPEG-4 fine granular scalability (FGS) bitstream, both the size of video frames and the bit rate are reduced. And an array processing algorithm of layer interference suppression is used at the receiver which makes the system structure provide different levels of protection to different layers. Furthermore, by considering the important level of scalable bitstream, the different bitstreams can be given different level protection by the system structure and channel coding. With the proposed system, the concurrent large diversity gain characteristic of STBC and alleviation of the frequency-selective fading effect of OFDM can be achieved. The simulation results show that the proposed schemes integrating scalable transcoding can provide a basic quality of video transmission and outperform the conventional single layer transcoding transmitted under the random and bursty error channel conditions.

Key words: scalable transcoding; down-sampling; space-time codes; orthogonal frequency division multiplexing

With the development of network and wireless communication technology, the demand for video streaming services grows rapidly. However, because various networks are heterogeneous with respect to available bandwidth and client capability, a successful access to a video server in a wired network from a mobile wireless user cannot be guaranteed. In contrast to the video content in server, the terminal device often has lower display size and bandwidth. High quality MPEG-2 bitstreams such as DVD existing in a video server often cannot be delivered to a mobile user in real time if the video bitstream is not transcoded to lower bit rate to fit the bandwidth constraint. In this case, a video transcoder can be used at the video proxy server or gateway to dynamically adjust the video bit rate according to channel bandwidth[1].

On the other hand, due to the limitations of wireless channels including limited bandwidth, channel loses, noise, interference, multi-path propagation, and uncorrected channel errors may result in significant image quality degradation at the decoder. This is particularly evident in standard coders, such as those based on MPEG-x or H. 26x, where variable length coding is used or where compression involves a predictive coding scheme[2]. To achieve high video quality at the decoder requires error control techniques for video transmission. The unequal error protection (UEP) is one of the useful and practical techniques for robust video transmission. It can provide different level error protection to different bits in the compressed stream[3–5]. In addition, the fine granularity scalable (FGS) video coding scheme adopted in the MPEG-4 streaming video profile provides a fine granularity scalable bitstream of different degrees of importance at different levels[6]. This property makes FGS bitstream suitable for transmitting over the error-prone channels with unequal error protection.

In wireless communications, frequency-selective fading in unknown dispersive channels is a dominant problem in high data rate transmission. The resulting multipath effects reduce the received power and cause inter symbol interference (ISI). Orthogonal frequency division multiplexing (OFDM) is often applied to combat this problem[7]. OFDM is a special case of multicarrier transmission, where a single datastream is distributed and transmitted over a number of lower transmission rate subcarriers. Therefore, OFDM in effect slices a broadband frequency-selective fading channel into a set of parallel narrow-band flat-fading channels.

In a flat-fading channel, an extra signal gain can be obtained by applying space-time block code (ST-

Received 2005-04-13.

Foundation item: The National Natural Science Foundation of China (No. 60202006, 60332030).

Biographies: Qian Tuanjie (1971—), male, graduate, tjqian@sjtu.edu.cn; Sun Jun (corresponding author), male, doctor, professor, sunjun@cdtv.org.cn.

BC), which brings in transmission diversity[8,9]. However, one report[10] shows that even with feedback from the decoder subsequent to the STBC decoder, the performance of the STBC decoder itself will not be improved by soft decoding since there is no new independent extrinsic information. Consequently it is necessary to concatenate an outer channel code with the STBC code to enhance the error correcting capability of the system. RS coding is a type of forward error correction (FEC) that is being widely used because of its relatively large error correction capability.

In this paper, we propose a UEP scheme to provide robust access to the pre-encoded high quality video server from mobile wireless terminals in dispersive fading channels. The scheme combines an MPEG-2 with an MPEG-4 fine grain scalable transcoder with STBC for an OFDM system. With the proposed system, the bandwidth-adaptive target bitstreams can be obtained by a scalable transcoder and the concurrent large diversity gain characteristic of STBC can be achieved. Furthermore, by considering the important level of video stream, the different coded video stream can be given different level protection by the system structure and channel coding. Experimental results show that our system can provide better performance of video transmission.

1 System Description

Assume that the mobile end user information such as display size, bandwidth availability and channel conditions be sent to the server over a feedback channel. Fig. 1 shows the proposed video transmission architecture. At the transmitter, the scalable transcoder makes the MPEG-2 input bitstreams with high resolution (D1 format) output MPEG-4 base layer stream (base stream) and FGS stream with lower resolutions (CIF, QCIF). Then two streams are transmitted over the layered space-time OFDM systems. At the receiver, MPEG-4 FGS decoder will reconstruct video sequence.

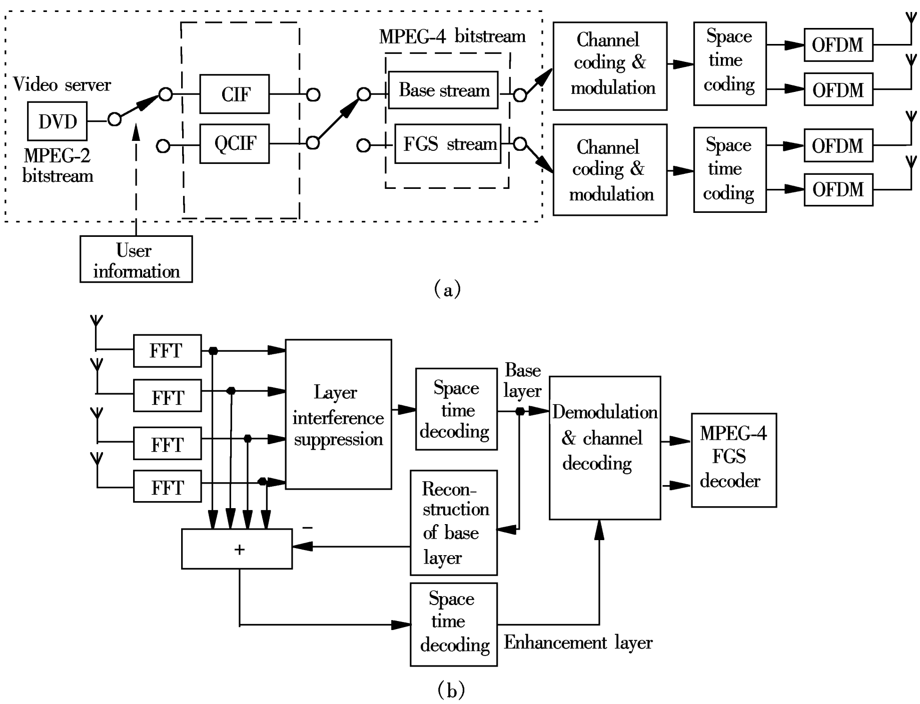

Fig. 1 Video transmission architecture. (a) Transmitter; (b) Receiver

2 Scalable Transcoding

Scalable transcoding is a process of converting a single compressed video bitstream into multiple streams. Scalable transcoding has been studied in Refs. [11, 12]. For bit rate adaptation, an SNR scalable transcoder using enhancement information construction is presented[11]. But such a transcoder cannot fit the terminal user with lower resolution. In Ref. [12], a cascaded decoder-encoder transcoder from MPEG-2 to MPEG-4 FGS is presented. However, this transcoder is too complex to be used in real-time applications because code information which comes from input bitstream is not taken into account. To keep tradeoffs between the quality and complexity, based on our previous scheme[13], an efficient MPEG-2 to MPEG-4 FGS transcoder with spatial downsizing is proposed, as shown in Fig. 2, in which the process of motion compensation of B frames is omitted.

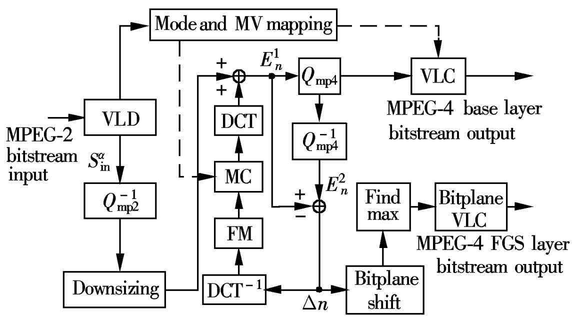

Fig. 2 Scalable transcoder architecture integrating down-sampling

2.1 MPEG-2 to MPEG-4 scalable transcoder

Motion estimation (ME), motion compensation (MC), DCT and IDCT are the most time-consuming processes[14]. So the main goals of transcoding design are to avoid the cascaded decoding and re-encoding process, and to skip these bottlenecks, while maintaining the video quality. In our transcoder, code information including motion vector (MV), macroblock mode, quantization parameter from the input bit-stream are re-used to decrease computation complexity. There are three types of pictures in an MPEG-2 bit stream: I, P, and B pictures. I picture is intra coded picture, which needs no ME and MC processes. Both P picture (predicted picture) and B picture (bidirectionally-interpolated picture) need MC process. Assuming that B pictures are not used as references for further prediction, and the computational complexity of transcoding without drift compensation on frames is far less than transcoding with MC, B pictures are treated similarly to I pictures. So, for I and B frames, the DCT coefficients are fed to perform MPEG-4 re-quantization directly. Since I pictures are the anchors for subsequent P and B frames, inverse DCT of MPEG-4 for I pictures is still needed to reconstruct the reference frames.

Now we will analyze the transcoding of these three types of pictures separately. S_{in}^{α} and $S_{\mathrm{out}}^{\alpha}$ ($\alpha = \mathrm{I}$, B, P) denote the variables of input and output residues, respectively.

For I and B pictures, transcoding is expressed as

$$S_{\mathrm{out}}^{\mathrm{I}} = Q_{\mathrm{mp4}}[\mathrm{DCT}(\mathrm{DCT}^{-1}(D(Q_{\mathrm{mp2}}^{-1}(S_{\mathrm{in}}^{\mathrm{I}}))))] = Q_{\mathrm{mp4}}[D(Q_{\mathrm{mp2}}^{-1}(S_{\mathrm{in}}^{\mathrm{I}}))] \tag{1}$$

where D refers to spatial downsizing, Q_{mp2} and Q_{mp4} denote the incoming quantization parameter of MPEG-2 and the output quantization parameter of MPEG-4, respectively. As shown in Eq. (1), the DCT and IDCT process can be omitted in I and B pictures transcoding.

For P pictures, the transcoding equation is

$$S_{\mathrm{out}}^{\mathrm{P}} = Q_{\mathrm{mp4}}\{\mathrm{DCT}[\mathrm{DCT}^{-1}(Q_{\mathrm{mp2}}^{-1}(S_{\mathrm{in}}^{\mathrm{P}})) + \mathrm{MC}\{D(p_{n-1}^{\mathrm{mp2}})\} - \mathrm{MC}(p_{n-1}^{\mathrm{mp4}})]\} = Q_{\mathrm{mp4}}\{D\{Q_{\mathrm{mp2}}^{-1}(S_{\mathrm{in}}^{\mathrm{P}}) + \mathrm{DCT}[\mathrm{MC}\{D(p_{n-1}^{\mathrm{mp2}})\} - \mathrm{MC}(p_{n-1}^{\mathrm{mp4}})]\}\} \approx Q_{\mathrm{mp4}}\{D\{Q_{\mathrm{mp2}}^{-1}(S_{\mathrm{in}}^{\mathrm{P}}) + \mathrm{DCT}[\mathrm{MC}\{D(p_{n-1}^{\mathrm{mp2}}) - p_{n-1}^{\mathrm{mp4}}\}]\}\} \tag{2}$$

where p_{n-1}^{mp2} and p_{n-1}^{mp4} denote the original resolution reference frame and the reduced resolution reference frame, respectively. Here, it is assumed that the MC is a nearly linear operation.

The FGS enhancement level residue error Δn can be denoted as

$$\Delta n = E_n^1 - E_n^2 = D(Q_{\mathrm{mp2}}^{-1}(S_{\mathrm{in}}^{\alpha})) - E_n^2 + \mathrm{DCT}(\mathrm{MC}(D(p_{n-1}^{\mathrm{mp2}}) - p_{n-1}^{\mathrm{mp4}})) \tag{3}$$

where E_n^2 is the output residue in DCT domain after quantization and de-quantization.

2.2 Motion vector mapping and mode decision

To obtain the new motion vectors after downsizing, Ref. [15] used the weighted average of candidate motion vectors. The drawback of this approach is that the result is prone to impulse noise in candidate motion vectors. A median filter exhibits good performance against impulse noises[16]. And the candidate motion vectors have different degrees of correlations with the target motion vector. Therefore, a weighted median is used to perform the composition. For simplicity, the ratio of the image resizing 2 : 1 is taken into account in this paper. The new motion vector $\boldsymbol{v}'$ is estimated as

$$\boldsymbol{v}' = \begin{bmatrix} R_x^{-1} & 0 \\ 0 & R_y^{-1} \end{bmatrix} \boldsymbol{v}_{\mathrm{mean}} \tag{4}$$

where $\boldsymbol{v}_{\mathrm{mean}}$ belongs to $\{\boldsymbol{v}_i, i = 1, 2, 3, 4\}$, and satisfies

$$\sum_{i=1}^{4} w_i \|\boldsymbol{v}_{\mathrm{mean}} - \boldsymbol{v}_i\| \leqslant \sum_{i=1}^{4} w_i \|\boldsymbol{v}_j - \boldsymbol{v}_i\| \quad j = 1, 2, 3, 4 \tag{5}$$

where $\boldsymbol{v}_i$ represents the i-th macroblock (MB) motion vector of the input video; R_x, R_y are the down-sampling ratio (1/2); w_i is the number of the nonzero AC coefficients as motion measurement.

The new MB code mode needs to be re-estimated. The rules for determining the mode are as follows: ① If all the four input MBs are intra-mode, then the code mode for the downsized MB is set to intra-mode; ② If all the four input MBs are skipped, the downsized MB will be skipped; ③ In all other cases, the mode for the downsized MB is set to inter-mode.

In addition, all of the intra-MBs in B pictures of the input MPEG-2 bitstream are converted to inter-MB with zero motion vectors.

3 Transmitter and Receiver

Suppose that the frequency-selective fading channel state is constant during two consecutive OFDM symbol intervals (one frame) and varies from one frame to another (quasi-static fading). The additive Gaussian noise at the k-th subcarrier of the n-th

OFDM symbol are samples of independent complex Gaussian random variables with zero mean and variance $N_0/2$ per complex dimension. Suppose that the frequency synchronization is perfect and channel state information $\boldsymbol{H}$ is available at the receiver.

At the transmitter, the space-time coded system is divided into two layers and each layer has two transmit antennas as shown in Fig. 1. Each video stream is divided into L packets, and every packet has B bits. Each packet is encoded using RS (N, K) over GF (2^m). And the resulting blocks of code words are modulated by QPSK, which map consecutive 2 bits to a specific symbol. And STBC is employed which maps two consecutive symbols $[S_{1,i}, S_{2,i}]$ into four symbols $[S_{1,i}, -S_{2,i}^*, S_{2,i}, S_{1,i}^*]$ ($i = 1, 2$), where $*$ denotes conjugation operation. Then these symbols are transmitted over the channels as OFDM symbols. Each OFDM symbol has 2 N_s/m RS code symbols, where N_s is the number of subcarriers in OFDM. Let a complex vector $\boldsymbol{x}(n)$ denote an OFDM symbol, where n is the n-th discrete time index of the OFDM symbol interval. There are N_s complex signals in vector $\boldsymbol{x}(n)$, which are represented as $\boldsymbol{x}(k, n)$, $k = 1, 2, \ldots, N_s$. For subcarrier k, signals $\boldsymbol{x}(k, n)$ and $\boldsymbol{x}(k, n+1)$ in two consecutive OFDM symbols $\boldsymbol{x}(n)$ and $\boldsymbol{x}(n+1)$ are encoded into four OFDM symbols $\boldsymbol{x}_1(n)$, $\boldsymbol{x}_1(n+1)$, $\boldsymbol{x}_2(n)$ and $\boldsymbol{x}_2(n+1)$. The symbols $\boldsymbol{x}_1(n)$, $\boldsymbol{x}_1(n+1)$ are transmitted from the first antenna of layer 1, and $\boldsymbol{x}_2(n)$, $\boldsymbol{x}_2(n+1)$ are transmitted from the second antenna of layer 1 in two consecutive time intervals. The sub-stream of the enhancement layer is processed as the same as that of base layer.

At the receiver, an array processing algorithm of layer interference suppression is used[17]. After suppressing the interference of the other layer, for layer 1, the transmission system is equivalent to a space-time coded system with two transmit antennas and two receive antennas while for layer 2, it is equivalent to a space-time coded system with two transmit antennas and four receive antennas. For these equivalent systems, $\boldsymbol{X}_{n,i}(k)$ is the transmitted signal vector from antennas of layer i at the k-th subcarrier of the n-th OFDM symbol. $\boldsymbol{R}_{n,i}(k)$ is the received signal vector, and the channel frequency response matrix is $\boldsymbol{H}_i(k)$. So the receiver computes the decision metric:

$$\sum_{t=n}^{n+1} |\boldsymbol{R}_{t,i}(k) - \boldsymbol{H}_i(k)\boldsymbol{X}_{t,i}(k)|^2$$

over all possible received code words $\boldsymbol{X}_{n,i}(k)$ and decides in favor of the code word that minimizes the sum.

The first experiment is to show the system performance of different layers. A typical Rayleigh fading channel model with five multiple paths is adopted. The channel bandwidth is 1 MHz, and there are 128 subcarriers. Four subcarriers on each end are zero-padded as guard tones and the rest 120 tones are employed to transmit data.

Fig. 3 illustrates the bit error rates (BER) of layer 1 and layer 2 of this system with the same channel code rate 0.7 (similar results can be achieved at other channel code rates). It shows that the performance of layer 1 is near to the single user systems with two transmit antennas and two receive antennas (2Tx & 2Rx). However, where the SNR is less than 6 dB, the performance of layer 2 is a little lower than that of the single user systems with two transmit antennas and four receive antennas (2Tx & 4Rx). With the increase of SNR, the performance of layer 2 becomes better. When SNR is big enough, it is near to that of the single user systems with two transmit antennas and four receive antennas. Also it shows that the performance of layer 1 is about 2 dB worse than that of layer 2 at the BER 10^{-3}. This shows that the system structure can provide a UEP capacity for data transmission. Hence, the stream transmitted over layer 2 can be protected better than that over layer 1.

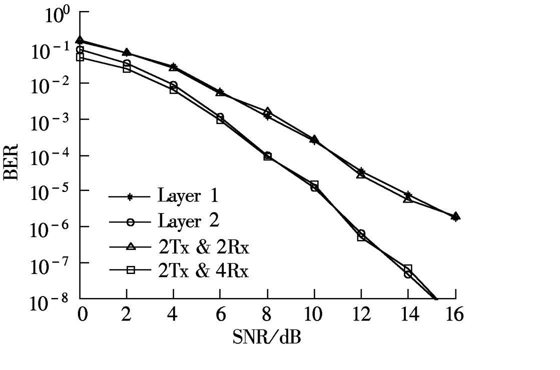

Fig. 3 Decoding performance comparison between layer 1 and layer 2 of our system and two single user systems 2Tx & 2Rx and 2Tx & 4Rx

4 Simulation Results

To demonstrate the effectiveness of our proposed system, the "mobile" sequence, which has a moderate amount of motion and high level of detail is considered. We code it at bit rate 3 Mbit/s, 30 frames, spatial resolution 720 × 576 and temporal resolution 25 frame/s with MPEG-2 codec. The structure of the GOP is (12, 3). The compressed video is transcoded into quarter spatial resolution whilst retaining the temporal resolution. The bit rate of the base layer is 267.4 kbit/s, and the bit rate of the truncated FGS layer is 267.5 kbit/s.

It is noted that because error concealment is unavailable in the decoder, an uncorrected channel error may result in the breakdown of the decoder. So in our experiments, minimum channel coding is needed to guarantee that the received streams are decodable. For comparison, the complete cascaded decoder-encoder single layer (nonscalable) transcoder provides a single bitstream with bit rate 535 kbit/s. And the single bitstream is divided into stream 1 and stream 2 to fit the layered STBC coded system. Simulation results are shown in terms of the BER performance versus the signal-to-noise ratio (SNR). The average PSNR of each frame after 10 times transmission is taken. Tab. 1 shows the different schemes.

Tab. 1 Different schemes

Schemes	Streams	Layers (equivalent system)	Channel codes
Our scheme (Scalable _ UEP)	Base stream	Layer 2(2T& 4R)	RS (30, 21)
	FGS stream	Layer 1(2T& 2R)	RS (30, 21)
Nonscalable _ UEP	Stream 1	Layer 2(2T& 4R)	RS (30, 21)
	Stream 2	Layer 1(2T& 2R)	RS (30, 21)
Scalable _ EEP	Base stream	Layer 2(2T& 2R)	RS (30, 21)
	FGS stream	Layer 1(2T& 2R)	RS (30, 21)
Nonscalable _ EEP	Stream 1	Layer 2(2T& 2R)	RS (30, 21)
	Stream 2	Layer 1(2T& 2R)	RS (30, 21)

Fig. 4 shows the comparison results of PSNR for mobile sequence using our scalable approach (scalable _ UEP) and the nonscalable scheme (nonscalable _ UEP) at a typical BER 10^{-3} in wireless environment with the same unequal error protection. By comparison, results of local decoding without transmission are also shown in Fig. 4. It can be seen that the PSNR of the scalable scheme is nearly 1. 7 dB worsethan that of the nonscalable without transmission. The reason is that the code efficiency of FGS is worse than that of the nonscalable encoder[6]. But we can see that the average PSNR of the scalable scheme is nearly 3. 2 dB better than that of the nonscalable. This is because the base stream of scalable scheme can achieve better protection by the system to guarantee a basic quality. And the error in the FGS stream cannot bring image drift because the FGS stream employs the intra-coding. On the other hand, for the nonscalable bitstream, a whole

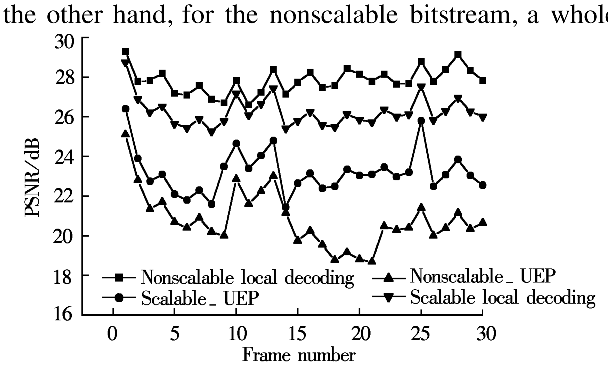

Fig. 4 Average PSNR of "mobile" video at SNR = 8 dB, BER = 10^{-3}

frame is partitioned into two equal parts with the same importance. Stream 1 is transmitted over layer 1, stream 2 is transmitted over layer 2. Because a strong relationship between the two streams, errors in stream 1 are easy to propagate to stream 2. So the average PSNR drops greatly. In addition, an interesting phenomenon is that there are three peaks at frames 1, 13, 25. This is because these frames are I frames which are not affected by the error propagation from B or P frames and, therefore, generally have higher PSNR values.

Fig. 5 shows the average PSNR of the reconstructed mobile sequence separately under different schemes as shown in Tab. 1. We can see clearly that our unequal error protection's performance is much better than the other schemes. For example, when SNR is 6 dB, the average PSNR of scalable _ UEP is higher than scalable _ EEP 9. 3 dB, nonscalable _ EEP 10. 1 dB, while the UEP _ nonscalable scheme is higher than EEP _ scalable 3. 8 dB, EEP _ nonscalable 5. 6 dB. The underlying reason for the UEP scheme outperforming the EEP scheme is that the UEP scheme gives different levels of protection according to different degrees of importance of the streams. The base stream is used to provide a minimally acceptable quality of video. So in the UEP scheme, it has the highest protection. While FGS frames are intra-coded, no error propagation takes place, thus weak protection is applicable. On the other hand, the EEP scheme cannot use this characteristic of the streams; the protection of the base stream is not enough while the FGS stream is overprotected, which leads to poor decoding performance.

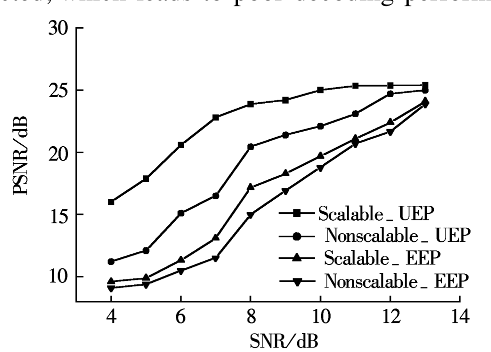

Fig. 5 Transmission performance of the proposed scheme with different SNR values

5 Conclusion

In this paper an efficient scheme integrating scalable transcoding with space-time OFDM systems is proposed for the mobile end-user accessing a wired video server. The MPEG-2 compressed video is transcoded by a scalable transcoder into MPEG-4 FGS bitstreams. An unequal error protection for different layers is employed which is given by channel code and the system structure. Experimental results show that the video quality obtained using our approach is higher the conventional single layer transcoding transmitted under the random and bursty error channel conditions.

References

[1] Wu Dapeng, Yi Wei, Hou Thomas, et al. Scalable video coding and transport over broadband wireless network [J]. *Proceedings of the IEEE*, 2001, **89**(1): 6 - 20.

[2] Zhang Q, Zhu W, Zhang Y Q. Channel-adaptive resource allocation for scalable video transmission over 3G wireless networks [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, **14**(8): 1049 - 1063.

[3] Pel Y, Modestino J W. Use of concatenated FEC coding for real-time packet video over heterogeneous wired to wireless IP Networks [A]. In: *Proceedings of International Symposium on Circuits and Systems* [C]. Bangkok, 2003, **2**: 840 - 843.

[4] Lisimachos Kondi P, Ishtiaq Faisal, Katsaggelos Aggelos K. Joint source channel coding for motion compensated DCT based SNR scalable video [J]. *IEEE Transactions on Image Processing*, 2002, **11**(9): 1043 - 1052.

[5] Po Keang, Kahn Joseph M. Transmission of analog signals using multicarrier modulation: a combined source channel coding approach [J]. *IEEE Transactions on Communications*, 1996, **44**(11): 1432 - 1443.

[6] Wu F, Li S P, Zhang Y Q. A framework for efficient progressive fine granularity scalable video coding [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, **11**(3): 332 - 344.

[7] Van Nee Richard, Prasad Ramjee. *OFDM for wireless multimedia communications* [M]. Boston: Artech House, 2000.

[8] Alamouti S M. A simple transmitter diversity scheme for wireless communications [J]. *IEEE Journal on Selected Areas in Communications*, 1998, **16**(10): 1451 - 1458.

[9] Tarokh V, Seshadri N, Calderbank A R. Space-time codes for high data rate wireless communication: performance criterion and code construction [J]. *IEEE Transactions on Information Theory*, 1998, **44**(2): 744 - 765.

[10] Bauch G. Concatenation of space-time block codes and "turbo"-TCM [A]. In: *IEEE International Conference on Communications* [C]. Vancouver, 1999, **2**: 1202 - 1206.

[11] Barrau Eric. MPEG video transcoding to fine granular scalable format [A]. In: *IEEE International Conference on Image Processing* [C]. New York, 2002. 717 - 720

[12] Yue Y, Chen C W. SNR scalable transcoding for video over wireless channels [A]. In: *IEEE Wireless Communications and Networking Conference* [C]. Chicago, 2000, **3**: 23 - 28.

[13] Xie Rong, Liu Jilin, Wang Xinguo. Efficient MPEG-2 to MPEG-4 compressed video transcoding [A]. In: *SPIE Visual Communications and Image Processing* [C]. San Jose, 2002, **4671**: 192 - 201.

[14] Assuncao P, Ghanbari M. A frequency-domain transcoder for dynamic bit-rate transcoder for dynamic bit-rate reduction of MPEG-2 bitstreams [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1998, **8**(8): 943 - 967.

[15] Shen B, Sethi I K, Bhaskaran V. Adaptive motion vector resampling for compressed video down-scaling [A]. In: *IEEE International Conference on Image Processing* [C]. Washington, DC, 1997. 771 - 774.

[16] Shanableh T, Ghanbari M. Heterogeous video transcoding to lower spatio-temporal resolutions and different encoding formats [J]. *IEEE Transactions on Multimedia*, 2000, **2**(2): 101 - 110.

[17] Su Pengcheng, Sun Jun, Wang Jia, et al. Studies on scalable video transmission techniques over layered space-time OFDM systems [J]. *Chinese Journal of Electronics*, 2004, **13**(3): 439 - 443.

一种具有可分级转码能力的空时 OFDM 视频传输方法

钱团结　孙　军　解　蓉　苏鹏程　王　嘉

(上海交通大学图像通信与信息处理研究所,上海 200030)

摘要:结合空时 OFDM 和转换编码技术,提出了一种无线衰落信道下具有可分级转码能力的鲁棒视频传输方法.采用可分级转码器将高质量的 MPEG-2 压缩视频转换为低码率、低分辨率 MPEG-4 可分级码流来满足网络带宽和终端设备显示的要求.在接收端采用一种层干扰抑制算法,使得分层空时 OFDM 系统不同层的传输性能存在差异,从而使系统具有不对等保护能力.根据分级码流的重要程度不同,将转码输出的可分级码流分别由分层空时编码 OFDM 系统的不同层来实现视频的鲁棒传输.实验结果表明:在典型的随机突发错误的无线环境下,提出的具有可分级转码能力的系统的视频传输性能优于传统的非分级转码的视频传输系统.

关键词:可分级转码;下采样;空时码;正交频分复用

中图分类号:TN919