

Journal of Electronic Imaging

JElectronicImaging.org

Hybrid digital-analog video transmission in wireless multicast and multiple-input multiple-output system

Yu Liu
Xiaocheng Lin
Nianfei Fan
Lin Zhang

Hybrid digital-analog video transmission in wireless multicast and multiple-input multiple-output system

Yu Liu,* Xiaocheng Lin, Nianfei Fan, and Lin Zhang

Beijing University of Posts and Telecommunications, Key Lab of Universal Wireless Communication, Ministry of Education of PRC, No. 10 Xitucheng Road, Haidain District, Beijing 100876, China

Abstract. Wireless video multicast has become one of the key technologies in wireless applications. But the main challenge of conventional wireless video multicast, i.e., the cliff effect, remains unsolved. To overcome the cliff effect, a hybrid digital-analog (HDA) video transmission framework based on SoftCast, which transmits the digital bitstream with the quantization residuals, is proposed. With an effective power allocation algorithm and appropriate parameter settings, the residual gains can be maximized; meanwhile, the digital bitstream can assure transmission of a basic video to the multicast receiver group. In the multiple-input multiple-output (MIMO) system, since nonuniform noise interference on different antennas can be regarded as the cliff effect problem, ParCast, which is a variation of SoftCast, is also applied to video transmission to solve it. The HDA scheme with corresponding power allocation algorithms is also applied to improve video performance. Simulations show that the proposed HDA scheme can overcome the cliff effect completely with the transmission of residuals. What is more, it outperforms the compared WSVC scheme by more than 2 dB when transmitting under the same bandwidth, and it can further improve performance by nearly 8 dB in MIMO when compared with the ParCast scheme. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JEI.25.1.013006](https://doi.org/10.1117/1.JEI.25.1.013006)]

Keywords: hybrid digital-analog; SoftCast; ParCast; H.264/AVC; multicast; multiple-input multiple-output.

Paper 15462 received Jun. 4, 2015; accepted for publication Dec. 10, 2015; published online Jan. 13, 2016.

1 Introduction

With the boom of wireless digital facilities (e.g., smart phone, laptop, and tablet PC) and the development of wireless video communication technology, wireless video multicast has become one of the key technologies in wireless local area networks. The main challenge of wireless video multicast is to supply multiple receivers with different resolution videos fit to their channel qualities. In the traditional digital video transmission scheme, one essential requirement is to guarantee that all the receivers can decode video. Therefore, the digital coding scheme must encode the video stream at the lowest bit-rate. This results in all the receivers having to degrade their decoding video performance to be the same as that of the receiver with the worst channel in a multicast group, and receivers with higher channel qualities cannot get higher video performance. This is the so-called “cliff effect,”^{1,2} where pure digital coding schemes cannot realize graceful performance degradation. Traditionally, video frames are first compressed, encoded into a bit stream (e.g., MPEG4³ and H.264),^{4,5} then transmitted over MIMO-OFDM channels. Similarly, in the MIMO-OFDM system, the multiple-input multiple-output (MIMO) channel is divided into a set of spatial subchannels with different gains. Therefore, error behavior differs across subchannels. Further, if we adopt a unique code rate to support conventional digital transmission over all the subchannels, the received video qualities on different antennas will be degraded to be the same as that of the antenna with the worst gain. Thus, this

can also be regarded as the cliff effect problem in digital video multicast.

Recently, some schemes have been proposed to surmount the cliff effect. The typical approaches are scalable video coding (SVC)^{6–8} with hierarchical modulation (HM)^{9–12} schemes, e.g., multiple description coding (MDC)^{13–16} schemes and multiple resolution coding (MRC)^{17–19} schemes. MDC encodes a video stream into multiple substreams called descriptions. Receivers can decode a low-quality video with any description, and the decoding performance improves with the number of received descriptions. When applied to video multicast, descriptions should be transmitted at special bit-rates to meet receivers’ channel qualities. Then, receivers can receive different numbers of descriptions with particular bit-rates that can be supported by the receiver’s channel. Thus, receivers with different channel qualities can decode videos with different resolutions. However, MDC is still a purely digital coding scheme, and the encoding of MDC is complex. The cliff effect in MDC degenerates to the “staircase effect.”²⁰ What is more, the upper bound of MDC’s performance is still decided by its digital coding scheme, where quantization residuals are discarded. MRC that divides video into a base layer and a few enhancement layers is also applied to video multicast. In contrast to the MDC schemes, the basic video can be obtained by decoding the base layer, and the overall performance can be improved by the enhancement layers in MRC. Therefore, the base layer is coded with the lowest bit-rate and the enhancement layers are coded with multiple bit-rates. Then, receivers with better channels can decode more enhancement layers besides the base layer, and better videos can be achieved. However, similarly to MDC, MRC still cannot overcome

*Address all correspondence to: Yu Liu, E-mail: liuy@bupt.edu.cn

the cliff effect completely, and the encoding of MRC is also complex.

Another popular approach is SoftCast,^{21–23} which transmits analog signals directly. SoftCast broadcasts one video to all the receivers without allocating different bit-rates to different receivers, and, at the same time, the decoded video on each receiver matches up to their channel. Therefore, it can completely solve the existing cliff effect problem in traditional digital video multicast schemes. But when it comes to video performance, SoftCast does not behave well at low channel qualities. It may perform worse than the traditional H.264/AVC scheme when transmitted under the same bandwidth. Based on the SoftCast scheme, which can solve the cliff effect completely and achieve graceful degradation, some hybrid digital-analog (HDA) video delivery schemes that integrate the advantages of digital coding and analog coding are proposed, e.g., WSVC,²⁰ WaveCast,²⁴ and DCast.²⁵ WSVC is an HDA joint source-channel coding scheme, whereas SoftCast is used to code and transmit high-frequency layers of discrete wavelet transform coefficients. SoftCast is just used as the unequal error protection strategy in WaveCast. SoftCast in DCast is utilized to transmit the least important bits of coset. All these HDA schemes can solve the cliff effect effectively based on the utilization of SoftCast, and the video performances can be improved.

With the development of antenna technology, MIMO is used widely, and it is attracting more and more developers. Based on the SoftCast scheme, Ref. 26 proposes the ParCast scheme, which matches more important source components with higher-gain subchannels to solve the cliff effect and further improve video performance in MIMO. Through allocating power with joint considerations of the source and channel, minimum mean square error (MMSE) of the video transmission in MIMO-OFDM channel can be achieved. Unfortunately, the shortcoming of SoftCast, i.e., the poor performance at low SNRs, remains in ParCast.

As the successor of H.264/MPEG-4 or AVC standard, the emerging high-efficiency video coding (HEVC)²⁷ achieves more than 50% improvement in video compression over H.264/AVC standard, keeping comparable image quality at the expense of increased computational complexity. Although HEVC has a very good performance in data compression, it is a lossy compression because of quantization errors compared to analog video delivery. To deal with error-prone wireless channels, wireless HEVC²⁸ applies the reference picture set (RPS) feature to improve robustness to data losses compared to the predecessor H.264. Based on Refs. 29 and 30, HEVC can employ redundant Predicted (Marionette) frames to prevent temporal error propagation and fit the available bandwidth in real time. However, we cannot generate streams to handle every different channel quality, which means HEVC still has the staircase effect and these streams need extra storage. Above all, HEVC is lacking in computational complexity, flexibility, and storage compared to analog delivery.

While the initial idea was briefly introduced in our conference papers,^{31,32} we present a complete design for HDA video transmission and perform extensive evaluations in this work. This paper proposes to transmit the residuals of H.264/AVC encoding to further improve video quality. The H.264/AVC stream is encoded by forward error correction (FEC), and both the compression ratio of H.264/AVC and

bit error rate (BER) are taken into account to guarantee effective and trustworthy digital transmission.^{4,5} It is crucial that residuals of H.264/AVC are coded by SoftCast in video multicast and by ParCast in MIMO-OFDM video transmission. Additionally, with effective designs of the compression ratio and channel code rate, any receiver can decode a basic video when enough power is assigned to the digital signals. Power allocation strategies in video multicast and MIMO video transmission are designed separately, depending on the principle that the power strategy should guarantee accurate digital decoding and maximize analog gains. In addition, power allocation algorithms in digital coding and in analog coding are proposed, respectively. In order to make full use of channel bandwidth, an HDA mapping scheme is also introduced. Simulations show that the proposed HDA scheme can realize graceful degradation with the application of SoftCast. Moreover, the proposed HDA scheme in video multicast outperforms the WSVC scheme, and the proposed HDA scheme in MIMO can achieve almost perfect video quality, which outperforms the ParCast scheme significantly.

To summarize, the major contributions of this paper are for two cases: multicast and MIMO. First, we proposed a new HDA scheme in multicast to completely overcome the cliff effect with the transmission of residuals. While WSVC only transmits the low-frequency layers of DWT coefficients, our scheme transmits all the digital data to guarantee reconstructed quality at some bad channels. Unlike WSVC, the digital and analog signals in our HDA scheme do not interfere with each other, and more power will be allocated to the analog part, which means analog gains can be higher. Therefore, our scheme performs about 2 dB better than WSVC. Second, the difference and relation between multicast delivery and unicast over a MIMO-OFDM link are analyzed. Then, to handle video unicast over a MIMO-OFDM link, we proposed the HDA scheme in MIMO. Compared with ParCast, the proposed scheme utilizes the advantage of digital delivery to ensure important data can be correctly reconstructed. The simulation results show the overall performance of our scheme is 2 to 8 dB better than ParCast.

The rest of this paper is organized as follows. Section 2 introduces the related works. Then, the HDA scheme in video multicast is proposed in Sec. 3, and the HDA scheme in MIMO video transmission is introduced in Sec. 4. Next, Secs. 5 and 6 illustrate the performance of the proposed schemes. Finally, Sec. 7 concludes this paper.

2 Related Works

The proposed HDA schemes consist of three parts: a digital part coded by H.264/AVC,^{4,5} an analog part coded by SoftCast in multicast or ParCast in MIMO, and an HDA mapping scheme that superimposes the digital signals with analog signals.

2.1 H.264/AVC

H.264/AVC^{4,5} is the de facto video coding standard used currently. H.264/AVC contains a series of new features, and it can compress video more effectively. It has been widely used in almost all forms of digital video coding. In this paper, H.264/AVC is used in the digital coding of video sequences. Since the quantization error and the compression rate increase with the quantization parameter (QP), QP is used

as a variable to adjust the compression ratio based on the worst channel quality of the receivers group, and the basic digital video performance is also decided by QP.

2.2 SoftCast

Katabi et al.^{21–23} have proposed the SoftCast scheme, which can broadcast one coded video signal to multiple receivers, while all the receivers can decode videos matching their channel qualities. SoftCast mainly consists of four parts: decorrelation, power allocation, whitening, and linear least square estimator (LLSE) decoding.

Similar to the digital video coding scheme, the first step of SoftCast is decorrelation. Discrete cosine transform (DCT) or discrete wavelet transform (DWT) have been used to remove space correlation of video frames in SoftCast. Moreover, 3D-DCT²³ or 2D-DWT+DCT²⁰ can also be applied in SoftCast to remove the space correlation and the time correlation simultaneously. In addition, similar to the digital decorrelation, DCast²⁵ removes the interframe correlation by inter prediction.

Another important step of SoftCast is power allocation, which minimizes the reconstruction errors by scaling the magnitudes of signals. Let \mathbf{x} be a decorrelated signal, and it is divided into N chunks based on its coefficient structure. Assume that the chunked signal is $\tilde{\mathbf{x}}$ and the i 'th row of $\tilde{\mathbf{x}}$ corresponds to the i 'th chunk. Next, the diagonal matrix Λ_x , whose i 'th diagonal element λ_i is the variance of the i 'th chunk, can be obtained. Assuming that the total power budget of SoftCast is P , we can get the power allocation matrix \mathbf{G} , which is a diagonal matrix with the i 'th diagonal elements

$$g_i = \lambda_i^{-1/4} \sqrt{\frac{P}{\sum_{j=1}^N \sqrt{\lambda_j}}} \quad (1)$$

Thus, power allocation can be formulated to be

$$\tilde{\mathbf{y}} = \mathbf{G}\tilde{\mathbf{x}}, \quad (2)$$

where $\tilde{\mathbf{y}}$ is the powered signal.

Next, a Hadamard matrix \mathbf{H} is employed as the whitening matrix to balance packets with equal importance. We can get the whitening signal $\hat{\mathbf{y}} = \mathbf{H}\mathbf{G}\tilde{\mathbf{x}}$. Therefore, if these whitening chunks are assigned to packets, the packets will have the same power and it can provide better packet loss protection.²¹ After being transmitted in an interference environment, the received signal is

$$\mathbf{y} = \mathbf{H}\mathbf{G}\tilde{\mathbf{x}} + \mathbf{n} = \mathbf{C}\tilde{\mathbf{x}} + \mathbf{n}, \quad (3)$$

where \mathbf{n} is the channel noise and $\mathbf{C} = \mathbf{H}\mathbf{G}$.

At the decoder, the chunked signal can be reconstructed by an LLSE decoder, and the reconstructed signal is

$$\mathbf{x}_{\text{LLSE}} = \Lambda_x \mathbf{C}^T (\mathbf{C} \Lambda_x \mathbf{C}^T + \Sigma)^{-1} \mathbf{y}, \quad (4)$$

where \mathbf{x}_{LLSE} refers to the LLSE estimation of chunked signal $\tilde{\mathbf{x}}$, and Σ is a diagonal matrix with the noise power of each chunk in the diagonal element. Then, the reconstructed chunked signal with MMSE, i.e., \mathbf{x}_{LLSE} , can be achieved. Next, the original signals can be obtained by inverse decorrelation operation.

Different from the traditional digital video coding scheme, signals are coded without quantization and entropy coding in SoftCast, and videos are transmitted by analog transmission. Therefore, it can completely avoid the cliff effect. Thus, it can achieve effective compression, data protection, and transmission performance. What is more, with the power allocation algorithm and the LLSE decoder,³³ SoftCast can reach the MMSE of video multicast.

2.3 ParCast

ParCast²⁶ is a video coding scheme designed for single-user MIMO video transmission, while SoftCast is a video coding scheme designed for multiple users' multicast. The MIMO channel is first divided into subchannels by singular value decomposition (SVD) in ParCast. Then, different from the SoftCast²¹ scheme, which allocates different source chunks with different power, ParCast allocates the better subchannels to the more important data chunks and discards some bad chunk-subchannel pairs to reallocate more power to the important pairs before power allocation.

Assuming that the $M \times N$ MIMO channel state matrix is \mathbf{H} , it can be decomposed into M orthogonal subchannels by SVD, i.e.,

$$\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^H, \quad (5)$$

where \mathbf{U} and \mathbf{V} are the unitary matrices, $(\cdot)^H$ means conjugate transpose or Hermitian, \mathbf{S} is a diagonal matrix with the channel gain of the i 'th orthogonal subchannel s_i at the diagonal of \mathbf{S} , and the gains of \mathbf{S} are decreasing.

In ParCast, the encoder removes the temporal and spatial redundancy by 3D-DCT³⁴ first. Each group of DCT coefficients is divided into chunks uniformly and all the chunks are sorted by variances. Then all the chunks are assigned to the available subchannels based on the principle that higher-variance chunks are transmitted over higher-gain subchannels. After that, the encoder allocates transmission power to each chunk based on joint consideration of source and channel. Next, data are precoded by matrix \mathbf{V} , then transmitted over the MIMO system. Correspondingly, the receiver will multiply the received signal by \mathbf{U}^H as precoding at the decoder to divide the MIMO channel into subchannels. The whole encoding procedure ensures that the important data in a frame gets more protection by matching them to high-gain subchannels. To achieve the MMSE of video transmission, the analogous water-filling power allocation strategy, i.e., the SoftCast power allocation strategy, is adopted. Therefore, the received signal is

$$\mathbf{Y} = \mathbf{U}^H \mathbf{H} \mathbf{V} \mathbf{G} \mathbf{M} \mathbf{X} + \mathbf{n} = \mathbf{S} \mathbf{G} \mathbf{M} \mathbf{X} + \mathbf{n} = \mathbf{C} \mathbf{X} + \mathbf{n}, \quad (6)$$

where \mathbf{G} is the power allocation diagonal matrix, \mathbf{M} is the random orthogonal matrix that has the same effect of whitening in SoftCast, \mathbf{n} is the channel noise, and $\mathbf{C} = \mathbf{S} \mathbf{G} \mathbf{M}$. Therefore, when these whitening chunks are assigned to OFDM symbols, the average power per OFDM symbol is the same. It results in the average power across transmit antennas being the same, and it can ease the requirement on the dynamic range of the power amplifier at the transmitter.²⁶ Assume that λ_i is the variance of the i 'th chunk and the power budget is P . With joint consideration of source and channel, the i 'th diagonal element of \mathbf{G} can be derived as

$$g_i = \sqrt{\frac{P}{\sqrt{\lambda_i s_i^2} \sum_{j=1}^N \sqrt{\lambda_j / s_j^2}}} \quad (7)$$

Therefore, the encoding and transmission procedure in MIMO system is equivalent to that of the SoftCast in multicast, as shown in Eq. (3).

Similarly, the receiver can decode the original signal using the LLSE decoder. The optimal decoder is

$$\mathbf{D}_{\text{LLSE}} = \mathbf{Q}\mathbf{C}^T(\mathbf{C}\mathbf{Q}\mathbf{C}^T + \mathbf{\Sigma})^{-1}, \quad (8)$$

where \mathbf{Q} and $\mathbf{\Sigma}$ are both the diagonal matrices. The i 'th diagonal coefficient of \mathbf{Q} is λ_i/s_i^2 , where λ_i is the variance of the i 'th chunk, and the i 'th diagonal coefficient of $\mathbf{\Sigma}$ is the noise power of the i 'th subchannel.

Based on ParCast, ParCast+ is presented in Ref. 35 to feature improved video source decorrelation and more flexible source-subchannel mapping. Specifically, ParCast+ adopts a motion-aligned three-dimensional (3-D) transform to decorrelate the source, which is generally considered more efficient at compacting the source energy than the 3D-DCT in ParCast. However, the analog data in the HDA scheme is the residual, which hardly has relevance in the time domain. Therefore, in our proposed HDA scheme in MIMO, we adopt ParCast to deal with the analog data.

Different from the SoftCast scheme, the source and channel are considered jointly in ParCast. The chunk-subchannel pairs are powered by the SoftCast power allocation strategy, so it can achieve the MMSE of video transmission in MIMO with the corresponding LLSE decoder.

2.4 Hybrid Digital-Analog Scheme

Apart from the digital schemes and the analog schemes, there are still many HDA schemes proposed in recent years, e.g., the WaveCast video delivery scheme in Ref. 24, the DCast video delivery scheme in Ref. 25, and the WSVC video delivery scheme in Ref. 20. These schemes are all designed for video multicast.

WaveCast first transforms the video by 3D-DWT, then these transformed coefficients are powered by the SoftCast power allocation algorithm. Different from the SoftCast scheme, which transmits analog signals directly, WaveCast modulates the analog coefficients by a very dense constellation (64K-QAM) and codes the metadata by FEC with binary phase shift keying (BPSK). After that, the modulated signals are transmitted over the OFDM channel.

In DCast, the original signals are first transformed by DCT. Then two successive coset coding modules are applied to classify each DCT coefficient into most important bits, least important bits, and the remaining middle bits. Next, the least important bits are coded by SoftCast; meanwhile, the middle bits are coded by FEC with BPSK to be transmitted in the OFDM channel, and the most important bits are discarded since they have high correlation with the decoder prediction and can be guessed at the decoder.

The HDA scheme in WSVC maps the digital signals and the analog signals to the same constellation. In WSVC, video is decomposed into LL, LH, HL, and HH bands by 2D-DWT first. Then, the LL band is coded by H.264/AVC, and the output bitstream is coded by convolutional code with BPSK. Next, the residuals of the LL band quantization and the other

three bands are decomposed by temporal DCT, and the output decomposed residuals are coded by SoftCast. Finally, the output digital encoding signals (as coding signals of base layer) and the output SoftCast coding signals (as coding signals of enhancement layer) are superposed, then transmitted over the OFDM channels. Specifically, these analog coefficients powered by SoftCast are divided into two states, i.e., the "big" state and the "small" state. Then, the coefficient feature of SoftCast is taken into account in WSVC and the digital bitstream is superposed with the "small" analog coefficients. Additionally, these superposed coefficients are mapped to the I component of the constellation diagram, while the "big" analog coefficients are mapped to the Q component. Therefore, though the digital bitstream is slightly interfered with by the "small" analog coefficients; it is orthogonal to the "big" analog coefficients.

In addition, an effective power allocation algorithm for the analog part and the digital part is designed in WSVC, where the digital signal can still be reconstructed correctly from the superposed signals interfered with by "small" coefficients and channel noise. With the application of the HDA scheme, WSVC integrates the advantages of digital coding and analog coding, i.e., high-coding efficiency and graceful degradation with channels. Therefore, it is able to broadcast one video to multiple receivers while each receiver can decode video performances matching its channel qualities. What is more, the WSVC scheme outperforms other hybrid schemes significantly, e.g., SVC + HM and DCast.

In Sec. 3, a new HDA scheme is proposed to further improve video performance. The HDA scheme is also applied to the video transmission in MIMO to certify the performance.

3 Hybrid Digital-Analog in Wireless Video Multicast

In conventional digital video coding schemes, e.g., MPEG, H.26x, and so on, video signal is transformed, quantized, and entropy coded, where the quantization error is ignored. In order to achieve an effective, robust, high-quality video multicast scheme, SoftCast is introduced to transmit the residuals of H.264/AVC quantization in wireless video multicast, as shown in Fig. 1. Since SoftCast can overcome the cliff effect, the proposed scheme can also achieve graceful degradation; meanwhile, it improves the performance by analog gains. Next, a power allocation algorithm is designed in this section to ensure that all the receivers can decode the basic video and to maximize the performance gains of residual signals. In addition, an HDA mapping scheme is proposed to superpose digital bitstream and analog signals.

3.1 Digital Encoding of Hybrid Digital-Analog Video Multicast

In the proposed video multicast scheme, a video sequence is first encoded by H.264/AVC. Then, a low-rate convolutional code is used as the FEC to adapt robustly to channel SNR variability. To ensure that the digital signals can be decoded by all the receivers with different channel qualities, the FEC code rate is set to guarantee that the BER is less than the target BER, even when the modulated digital signals are transmitted over the worst channel. Assuming that the digital signal coding rate equals the channel bandwidth, since the FEC code rate is decided by the channel qualities, the quantization parameter (QP) that decides the compression rate of

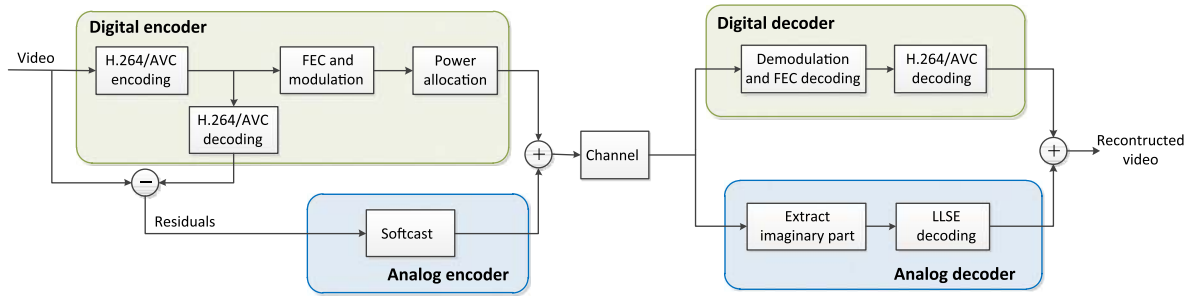


Fig. 1 Framework of the proposed HDA video multicast scheme.

H.264/AVC is also decided by channel qualities. When the channel quality is bad, a large QP will be adopted; a small QP will be adopted when the channel quality is good. Next, the BPSK modulator is employed to modulate the digital signal, and the H.264/AVC video information focuses on the real part of the modulated digital signals \mathbf{x}_{md} . Assuming that power allocated to the digital signal is P_d , the powered digital signal is $\mathbf{x}_d = \sqrt{P_d}\mathbf{x}_{md}$.

3.2 Analog Encoding of Hybrid Digital-Analog Video Multicast

The residuals of H.264/AVC quantization are sent to the analog encoder and powered by SoftCast to improve the performance of H.264/AVC. Information stored in the residuals depends on the QP set in H.264/AVC. The bigger the QP, the more information stored in the residuals, and the more performance gains provided by the residuals. In the analog part, the residual signals $\mathbf{x}_{\text{residual}}$ are first decomposed by 3D-DCT, then uniformly divided into n chunks. Next, it is powered by SoftCast, and the power factor of the i 'th chunk is

$$g_i = \lambda_i^{-1/4} \sqrt{\frac{P_a n}{\sum_{j=1}^n \sqrt{\lambda_j}}} \quad \text{s.t.} \quad \frac{\sum_{i=1}^n g_i^2 \lambda_i}{n} = P_a, \quad (9)$$

where P_a is the power allocated to the analog part and λ_i is the variance of the i 'th chunk. Thus, the powered analog signal is

$$\mathbf{x}_a = [g_1 \mathbf{x}_{\text{re},1}; g_2 \mathbf{x}_{\text{re},2}; \dots; g_n \mathbf{x}_{\text{re},n}], \quad (10)$$

where $\mathbf{x}_{\text{re},i}$ is the i 'th chunk of the residual signals.

3.3 Hybrid Digital-Analog Mapping and Power Allocation

To make full use of the channel resource, signal coding rates of the digital and analog parts are designed to be the same and equal to the channel bandwidth. Then, the transmitted HDA signal x is

$$x = x_d + j \cdot x_a, \quad (11)$$

where j is the symbol of the imaginary part. Different from the HDA scheme in WSVC, in our scheme, x_d is the I component of x and x_a is the Q component of x , as shown in Fig. 2. Therefore, the digital signals and the analog signals do not interfere with each other, more power will be allocated to the analog part, and the analog gains will be higher.

Once the received digital stream can be decoded, if the power allocated to the digital signals P_d is still increased, performance of the digital coding does not improve synchronously any longer. Therefore, P_d should be the power that can just support digital signal being decoded accurately under all the channel conditions. Assuming that the whole sending power is P , the power allocated to the analog part is $P_a = P - P_d$. Moreover, for the analog signals, the higher the P_a , the more analog gains can be obtained and the better the video performance that can be achieved. If we test binary signals coded by convolutional code with BPSK modulation, γ_0 , which is the minimum SNR where the decoding BER is smaller than the target BER (it is set to 10^{-6} in our scheme), can be obtained. Then, for the entire channel SNRs, the following equation should be satisfied

$$\frac{P_d}{N_0} \geq \gamma_0 \quad \text{s.t.} \quad N_0 = \frac{P}{10^{\text{SNR}/10}}. \quad (12)$$

Let N_m be the maximum noise variance and $\gamma_m = P/N_m$ be the minimum channel SNR of the receiver group. Then the minimum power allocated to the digital signal is

$$P_d = \gamma_0 N_m = \frac{\gamma_0}{\gamma_m} P. \quad (13)$$

Therefore, P_d is proportional to γ_0 and inversely proportional to γ_m . Therefore, for those channels whose SNRs are bigger than γ_m , the BER is lower than the target BER and the digital signals can be decoded accurately by all the receivers to provide a basic video. Then, the remaining power is all allocated to the analog signal to maximize the analog gains

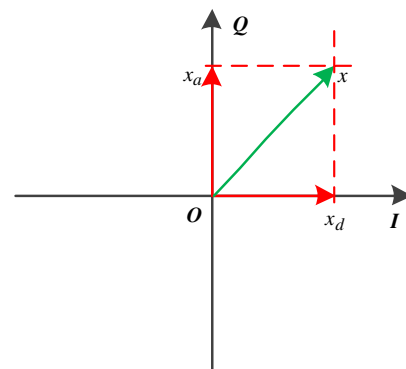


Fig. 2 Mapping of digital signals and analog signals to I/Q components of hybrid signals.

$$P_a = P - P_d = \left(1 - \frac{\gamma_0}{\gamma_m}\right)P. \quad (14)$$

Since P_a should be positive, then $\gamma_0 \leq \gamma_m$, where γ_m is the minimum SNR of multiple receivers.

3.4 Decoding of Hybrid Digital-Analog Video Multicast

Since x_d and x_a are mutually orthogonal in the hybrid signal, the digital signals and the analog signals will not interfere with each other. Therefore, the decoder can decode the digital signals from the hybrid signal by BPSK demodulator, Viterbi decoder, and H.264/AVC decoder⁷ directly. The analog signals can be obtained by extracting the imaginary part of the hybrid modulation signals, then the residual signals can be decoded by the LLSE decoder.³³ In the end, through superimposing the decoded digital signals and the decoded analog signals, the original video can be acquired. Since the H.264/AVC stream (the digital signals) can be decoded accurately by all the receivers with different channel qualities, the basic video can be obtained by all the receivers. In addition, receivers with different channel qualities can decode different amounts of information from the analog signals and obtain different analog gains. Therefore, receivers with better channels can receive better video, and the video quality is better than that transmitted in the H.264/AVC scheme. Thus, the proposed HDA video multicast scheme can achieve graceful degradation while improving video performance.

4 Hybrid Digital-Analog Transmission in Multiple-Input Multiple-Output System

4.1 Relation Between Multicast and Multiple-Input Multiple-Output

The proposed HDA scheme in multicast was designed with single-antenna links oblivious for multicast or broadcast. In fact, OFDM decomposes a wideband channel into a set of mutually orthogonal subcarriers, and the channel gains across these subcarriers are usually different.³⁶ With MIMO, each subcarrier is further divided into a set of spatial subchannels, again with different channel gains. In particular, error behavior differs across subchannels. If a one-size-fits-all code rate is used for a few or all subchannels, it generally needs to be conservative, and hence suboptimal, to accommodate the worst subchannels. Therefore, in this scene, unicast over a MIMO-OFDM link can assemble a multicast channel. With this in mind, we present the proposed HDA in MIMO.

However, there are some differences between unicast over a MIMO-OFDM link and a multicast channel. First, in MIMO-OFDM, a channel-dependent precoding operation is often necessary to make the spatial subchannels on the subcarrier mutually orthogonal so that the signals do not interfere with one another along different subchannels. Second, when some subchannels in MIMO have very poor performance, they can be discarded, and their power needs to be returned to the overall power budget for the other good subchannels. So the proposed HDA in multicast without precoding is not suitable for directly running over MIMO.

In this section, we propose the HDA scheme in MIMO to handle video unicast over a MIMO-OFDM link. The proposed HDA scheme is also applied to video transmission in MIMO, as shown in Fig. 3. Similar to the HDA video multicast scheme, videos are also coded by H.264/AVC first. In contrast, residuals of H.264/AVC in the MIMO system are coded by ParCast, which tailors the video unicast quality to the MIMO-OFDM channel. Compared with SoftCast, ParCast first employs SVD-based precoding to avoid interference between received signals on different subchannels. Then, ParCast allocates the better subchannels to the more important data chunks and discards some bad chunk-subchannel pairs to reallocate more power to the important pairs before power allocation. Moreover, in our proposed HDA scheme in MIMO, new power allocation algorithms are designed for the new case.

4.2 Digital Encoding of Hybrid Digital-Analog in Multiple-Input Multiple-Output

In the digital part, video frames are first encoded by H.264/AVC in groups, and the residuals of each group are sent to the analog encoder. Next, the H.264/AVC bitstream is coded by convolutional code with BPSK, then the modulated bitstream is divided into segments to be sent by multiple antennas. Note that the coding rate, which is decided by the convolutional code and the compression ratio of QP, cannot exceed the bandwidth, and the bitstream should be decoded by the MIMO receiver.

To guarantee that the received digital signals are interfered with by noise uniformly, a power allocation algorithm must be carried out before compounding digital signals and analog signals. Assuming that power allocated to the digital encoder is P_d , subchannel gains are s_i , $i = 1, 2, \dots, N$, where N is the number of subchannels used, and the digital power factor of the i 'th segment is $g_{d,i}$. Then, to reach equal error protection of the digital signal, Eq. (15) should be satisfied

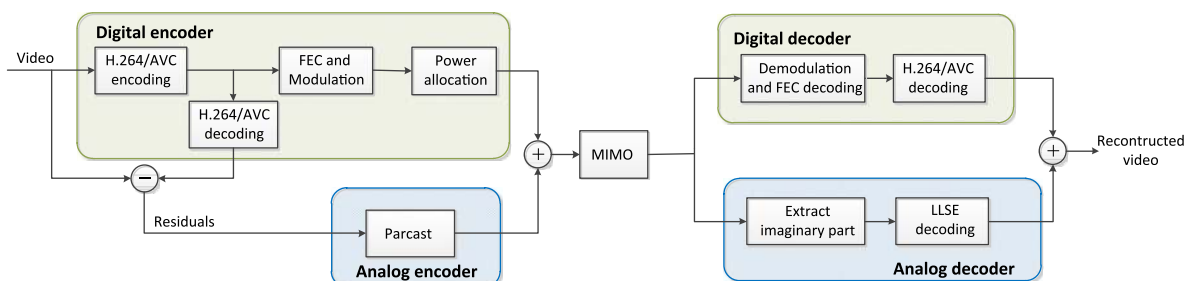


Fig. 3 Framework of the proposed hybrid digital-analog video scheme in MIMO.

$$\forall 1 \leq i, j \leq N, s_i^2 g_{d,i}^2 = s_j^2 g_{d,j}^2 \quad \text{s.t.} \quad \sum_{i=1}^N g_{d,i}^2 = P_d. \quad (15)$$

Then, the digital power factor can be calculated as

$$g_{d,i} = \sqrt{\frac{P_d}{s_i^2 \sum_{j=1}^N 1/s_j^2}}. \quad (16)$$

Therefore, though these digital segments are transmitted with different gains, BERs in each receiver antenna are the same. Here we call the powered digital signal x_d .

4.3 Analog Encoding of Hybrid Digital-Analog in Multiple-Input Multiple-Output

Residuals of H.264/AVC coding are sent to the analog encoder, then divided into chunks and coded by ParCast. Let power allocated to the analog encoder be P_a , and the analog power factor of the i 'th chunk is

$$g_{a,i} = \sqrt{\frac{P_a}{\sqrt{\lambda_i s_i^2} \sum_{j=1}^N \sqrt{\lambda_j / s_j^2}}}, \quad (17)$$

where λ_i is the variance of the i 'th chunk and s_i is the gain of the corresponding i 'th subchannel.

After being encoded by ParCast, the analog signal is decomposed into chunks to be transmitted by antennas. Here we call the coded signal x_a . With the gains of residuals, performance of H.264/AVC can be improved significantly and graceful degradation can be reached.

4.4 Hybrid Digital-Analog Mapping and Power Allocation in Multiple-Input Multiple-Output

To make full use of channel bandwidth, the signal coding rates of digital signal segments and those of analog signal chunks are also designed to be the same and equal to the channel bandwidth. Then the coded digital signals and coded analog signals are mapped to the HDA signals by the proposed HDA mapping scheme, as shown in Eq. (5) and Fig. 2. Thus, the digital signal and the analog signal are mutually orthogonal and do not interfere with each other.

Similar to the HDA scheme in multicast, power allocated to the digital signals P_d should be power that can just support the digital stream to be decoded when transmitted over the MIMO channel. The rest of the power is allocated to the analog signals. Let the whole sending power budget be $P = P_a + P_d$.

When the binary signal is tested by convolutional code with BPSK, γ_0 , which is the minimum SNR when the decoding BER is smaller than the target BER (it is 10^{-6} in our scheme), can be calculated. Then, the following equation should be satisfied to guarantee that the digital signal can be decoded when transmitted over the MIMO channel

$$\frac{P_{\text{rec-d}}}{N_0} \geq \gamma_0 \quad \text{s.t.} \quad N_0 = \frac{P}{10^{\text{SNR}/10}}, \quad (18)$$

where $P_{\text{rec-d}}$ is the received digital signal power and N_0 is the noise variance in each antenna. According to Eqs. (15) and (16), the received digital signal power can be derived as

$$P_{\text{rec-d}} = \frac{1}{N} \sum_{i=1}^N s_i^2 g_{d,i}^2 = \frac{P_d}{N \sum_{i=1}^N 1/s_i^2} \quad \text{s.t.} \quad P_{\text{rec-d}} \geq \gamma_0 N_0, \quad (19)$$

where N is the number of subchannels utilized by digital signals. Therefore, the power allocated to the digital encoder is

$$P_d \geq N_0 \gamma_0 N \sum_{i=1}^N 1/s_i^2. \quad (20)$$

Let $\gamma_m = P/N_0$ be the channel SNR. Then, to maximize the analog gains, equality should be reached and the power allocated to the digital signals is

$$P_d = \frac{\gamma_0 \sum_{i=1}^N 1/s_i^2}{\gamma_m} P. \quad (21)$$

As a result, digital signals received by all the antennas can be accurately decoded. Obviously, P_d is proportional to γ_0 and inversely proportional to γ_m . Thus, power allocated to the analog signals is

$$P_a = P - P_d = \left(1 - \frac{\gamma_0 \sum_{i=1}^N 1/s_i^2}{\gamma_m}\right) P. \quad (22)$$

Note that P_a should be positive, therefore

$$\gamma_m \geq \gamma_0 \sum_{i=1}^N 1/s_i^2 \triangleq \mathfrak{R}. \quad (23)$$

Since the diagonal coefficient of \mathbf{S} is decreasing, $1/s_i^2$ of some bad subchannels may be very big. Therefore, in our scheme, some bad subchannels with low-channel gains are discarded for digital signal transmission in order to reach a smaller \mathfrak{R} and improve the analog gains.

4.5 Decoding of Hybrid Digital-Analog in Multiple-Input Multiple-Output

Assume that the MIMO channel state matrix is \mathbf{H} and it is decomposed by SVD, i.e., $\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^H$. Then the precoded HDA signals are $\tilde{\mathbf{x}} = \mathbf{V}\mathbf{x}$ and the received HDA signals are

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{H}\tilde{\mathbf{x}} + \mathbf{n} = \mathbf{U}\mathbf{S}\mathbf{V}^H\mathbf{V}\mathbf{x} + \mathbf{n} = \mathbf{U}\mathbf{S}\mathbf{x} + \mathbf{n} \\ &= \mathbf{U}\mathbf{S}(\mathbf{x}_d + \mathbf{j} \cdot \mathbf{x}_a) + \mathbf{n}. \end{aligned} \quad (24)$$

Next, the received signals are multiplied by \mathbf{U}^H before decoding

$$\mathbf{y} = \mathbf{U}^H\tilde{\mathbf{y}} = \mathbf{S}(\mathbf{x}_d + \mathbf{j} \cdot \mathbf{x}_a) + \mathbf{U}^H\mathbf{n}. \quad (25)$$

Even though the HDA signals are transmitted over the complex MIMO channel, the digital signals and the analog signals in \mathbf{y} are still mutually orthogonal, as shown in Eq. (25). Thus, the receiver can also decode the digital bitstream from the hybrid signals directly and extract the analog signals from the imaginary part. With the digital decoders and the LLSE decoder, the analog and digital signals can be decoded, respectively, and the video can be obtained by

adding up the decoded digital signals and the decoded analog signals.

5 Simulation of Hybrid Digital-Analog Video Multicast Scheme

5.1 Simulation Environment

Test videos: we select a few representative video sequences, including *foreman*, *mobile*, *news*, *bus*, *football*, *akiyo*, *tennis*, *container*, and *coastguard*. They are in .cif format, with a frame size of 352×288 pixels. These videos have different motion characteristics, background textures, and energy distributions. For example, *news* and *akiyo* have a smooth background and little motion. They are therefore easily compressed with a conventional video codec. *Container* has many small moving objects, and there is continuous movement in the whole scene due to camera movement. *Foreman* contains an active close-up object and panning background. *Football* and *coastguard* both include large texture areas and complex motion. Consequently, both are challenging cases for conventional compression systems.

Performance metric: to evaluate the video delivery quality, we use the standard peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) index as performance metrics.

The PSNR is a standard objective measure of video quality and is defined as a function of the mean squared error (MSE) between all pixels of the decoded video and the original version as follows

$$\text{PSNR} = 10 \lg \left(\frac{2^L - 1}{\text{MSE}} \right), \quad (26)$$

where L is the number of bits used to encode luminance pixels for the original (uncompressed) video, typically 8 bits.

Unlike PSNR, the SSIM index is a full reference metric, in other words, the measuring of image quality based on an initial uncompressed or distortion-free image as reference. The SSIM emphasizes that the human visual system (HVS) is highly adapted to extract structural information from visual scenes. Therefore, a measurement of structural similarity (or difference) should provide a good approximation of perceptual image quality. The SSIM index is defined as a product of luminance, contrast, and structural comparison functions, and it is a decimal value between 0 and 1. A value of 0 means zero correlation with the original image, and 1 means the exact same image. Through this index, image and video compression methods can be effectively compared.

Schemes for comparison: we compared the proposed HDA video multicast scheme with three reference schemes, namely H.264/AVC, SoftCast, and WSVC. In our proposed HDA scheme, 3D-DCT is applied to remove space and time correlation of video residuals in one group of pictures (GOP). The GOP size is set to 16 in our simulation. Next, the transformed coefficients of residuals are divided into 44×36 chunks to be encoded by SoftCast. In addition, the hybrid signal is transmitted with a constant power budget P . γ_0 of different convolutional code rates has been tested as the prior information, as shown in Table 1.

In the WSVC scheme,²⁰ the parameter settings are mostly similar to the proposed HDA scheme; only the QPs and FEC

Table 1 γ_0 of different FEC code rate.

Convolutional code	Generator polynomial	γ_0
(3,1,4)	[13 11 14]	4.5 dB
(4,1,7)	[124 133 151 103]	1 dB
(5,1,8)	[377 233 254 216 301]	-0.5 dB

code rates, which are decided by the channel bandwidth and channel quality, are different. What's more, performances of the 3D-DCT SoftCast scheme²³ and the traditional H.264/AVC digital video scheme are also introduced. Except for the WSVC scheme, hybrid signals are transmitted with the same channel bandwidth and the signal coding rate that equals the bandwidth. On the other hand, analog signals are divided into two parts and only the low-frequency layer of DWT coefficients is coded by H.264/AVC in WSVC,⁴ and bandwidth used in WSVC is just half of the bandwidth used in other schemes. Thus, WSVC schemes with bandwidth equal to that of other schemes is also simulated and the signal data are repeated in this simulation with the same power budget P .

5.2 Parameter Analysis

In this subsection, we set different parameters in our proposed HDA scheme and analyze their roles. Since results for all tested video sequences exhibit similar trends, only the performances of *foreman* are shown in this subsection.

Figure 4 shows the PSNR performances of our scheme when processed at different convolutional code rates and different QPs. It is obvious that the performance of our scheme is much better than that of the pure H.264/AVC scheme when coded at the same digital coding environments, i.e., the same digital coding system and the same digital coding parameters. When the convolutional code rate is 1/4, the

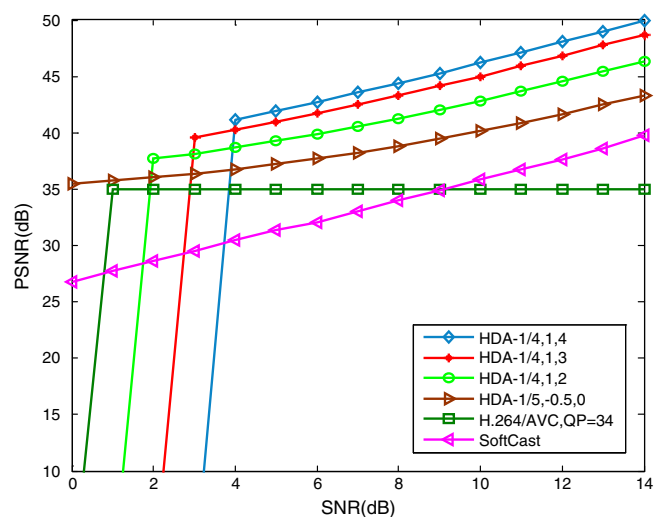


Fig. 4 PSNR of “foreman_cif.yuv” at different parameters and different schemes, where (1/4,1,4) refers to (FEC code rate, γ_0, γ_m), the “H.264/AVC, QP = 34” is coded by H.264/AVC with 1/4 convolutional code with BPSK and it equals the case “HDA-1/4, 1, 1.” QPs of HDA schemes are all 34 at code rate = 1/4, while QP of HDA scheme is 35 at code rate = 1/5.

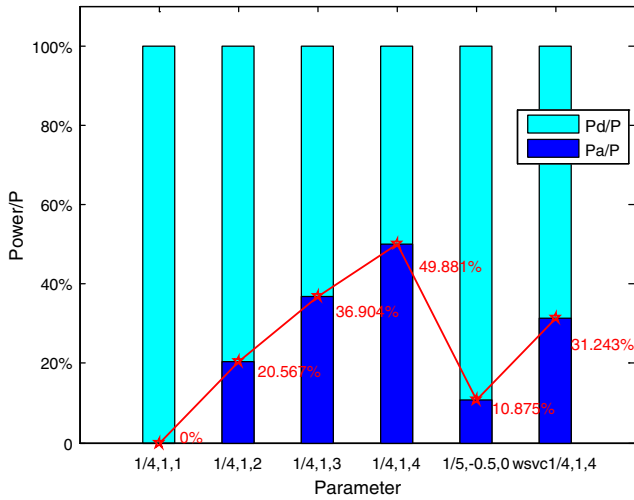


Fig. 5 Power allocated to digital part and analog part. Parameter refers to (FEC code rate, γ_0 , γ_m).

corresponding γ_0 is 1 dB and the corresponding QP is 34. Then, the curve “H.264/AVC, QP = 34” shows the performance of the conventional video multicast scheme, and it equals the case of our scheme when $\gamma_0 = \gamma_m = 1$ dB.

It also can be seen that the overall performance of our scheme increases with γ_m when coded at the same convolutional code rate 1/4. However, the cliff points where the digital signals can just be decoded also move with γ_m . Since γ_m is the minimum SNR of the receiver group, all the receivers of different cases can decode videos with different qualities. Since γ_0 of the FEC code rate 1/4 is 1 dB and $\gamma_m \geq \gamma_0$, a

higher code rate should be adopted when the minimum SNR of the receiver group is smaller than the γ_0 and the code rate is 1/4. For example, when the minimum SNR is 0 dB, the strategy whose code rate is 1/5 and QP is 35 is adopted; then this strategy can cover receivers whose SNRs range from 0 to 14 dB, as shown in Fig. 4. Therefore, when the worst channel SNR is low, a higher coding rate with a larger QP is adoptable. A relatively lower coding rate with a smaller QP can be adopted to improve the basic H.264/AVC digital video quality when the worst channel SNR is high.

The power allocated to the digital signals P_d and the power allocated to the analog signals P_a at different coding parameters and different coding schemes are shown in Fig. 5. Obviously, P_a increases with γ_m at the same parameters. Therefore, if the cliff point is bigger (i.e., the minimum SNR of all the receivers’ channels is bigger), more power will be allocated to the analog part, the performance gains based on analog signals will be higher, and the entire video performance will be better. Even though video is transmitted at the same FEC code rate (1/4) and the same γ_m (4 dB), the power allocated to P_a and P_d in the proposed scheme and that in WSVC are not the same. Because the digital signals in WSVC are not only interfered with by the channel noise, but also by the “small” analog signals, more power will be allocated to the digital part to ensure accurate decoding of digital signals. As a result, the analog gains of WSVC are lower than the proposed scheme. It can also be seen that the analog power at (1/5, -0.5, 0) is less than the analog power when the FEC code rate is 1/4, so the analog gains are much lower.

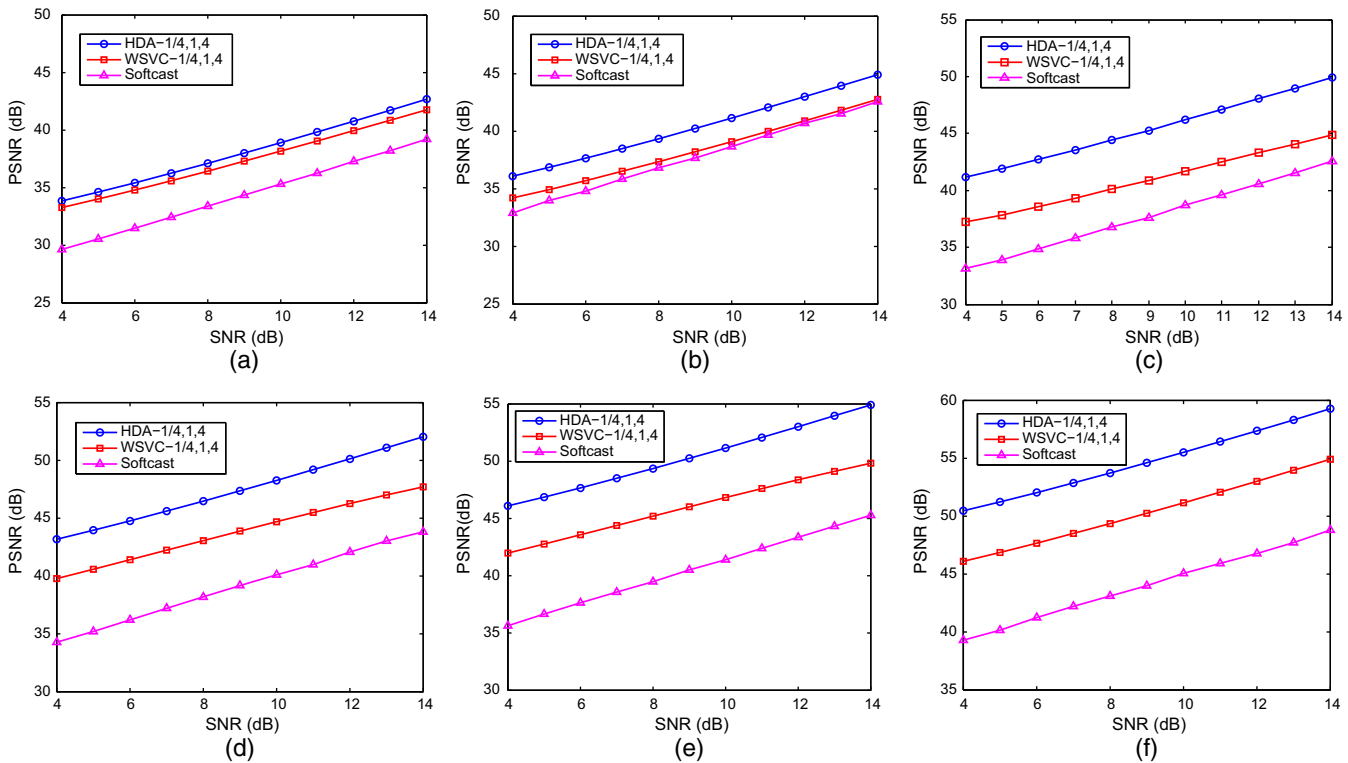


Fig. 6 PSNR comparison in multicast: (a) football, (b) coastguard, (c) foreman, (d) container, (e) news, (f) akiyo.

Table 2 Parameters when processing different sequences.

Scheme	CIF	QP	QPSNR(dB)	P_a	P_d
WSVC	News	25	40.51	40.489%	59.511%
	Container	26	38.95	42.686%	57.314%
	Coastguard	33	32.03	33.365%	66.635%
	Football	38	28.64	33.295%	66.705%
	Foreman	30	35.66	31.243%	68.757%
	Akiyo	24	42.86	32.573%	67.427%
HDA multicast	News	29	39.02	49.881%	50.119%
	Container	31	36.03	49.881%	50.119%
	Coastguard	38	29.55	49.881%	50.119%
	Football	43	27.01	49.881%	50.119%
	Foreman	34	34.64	49.881%	50.119%
	Akiyo	26	41.52	49.881%	50.119%

5.3 Simulation Results

Figure 6 shows performance of the proposed scheme when compared with other schemes. Assuming that SNRs of the receiver group range from 4 to 14 dB, parameters (1/4,1 dB,4 dB) can be adopted to cover all the receivers. Suitable QPs should be adopted first, as shown in Table 2.

As shown in Fig. 6, WSVC is about 3 to 8 dB better than SoftCast. PSNRs of our scheme are about 1 to 5 dB higher than WSVC. Since only the low-frequency layers of DWT coefficients in WSVC are encoded by H.264/AVC, PSNR performance of WSVC may be worse than the pure H.264/AVC digital transmission scheme with the same QP at some bad channels. Power allocation of the proposed HDA multicast scheme depends on γ_m and γ_0 ; P_a and P_d of different sequences are the same. However, power allocation of the WSVC scheme is not only based on γ_m and γ_0 , but also on variances of all the chunks. So power allocations of WSVC schemes change with parameters. Though quantization PSNRs (QPSNRs) of the proposed scheme are smaller than those of the WSVC scheme, the proposed HDA scheme still outperforms the WSVC scheme owing to the analog gains.

To certify the performance of the HDA scheme, Fig. 7 depicts the SSIM values for the tested video sequences. It can be seen that for different SNRs, the proposed HDA scheme outperforms the WSVC and SoftCast since it achieves better visual quality. The visual quality comparison

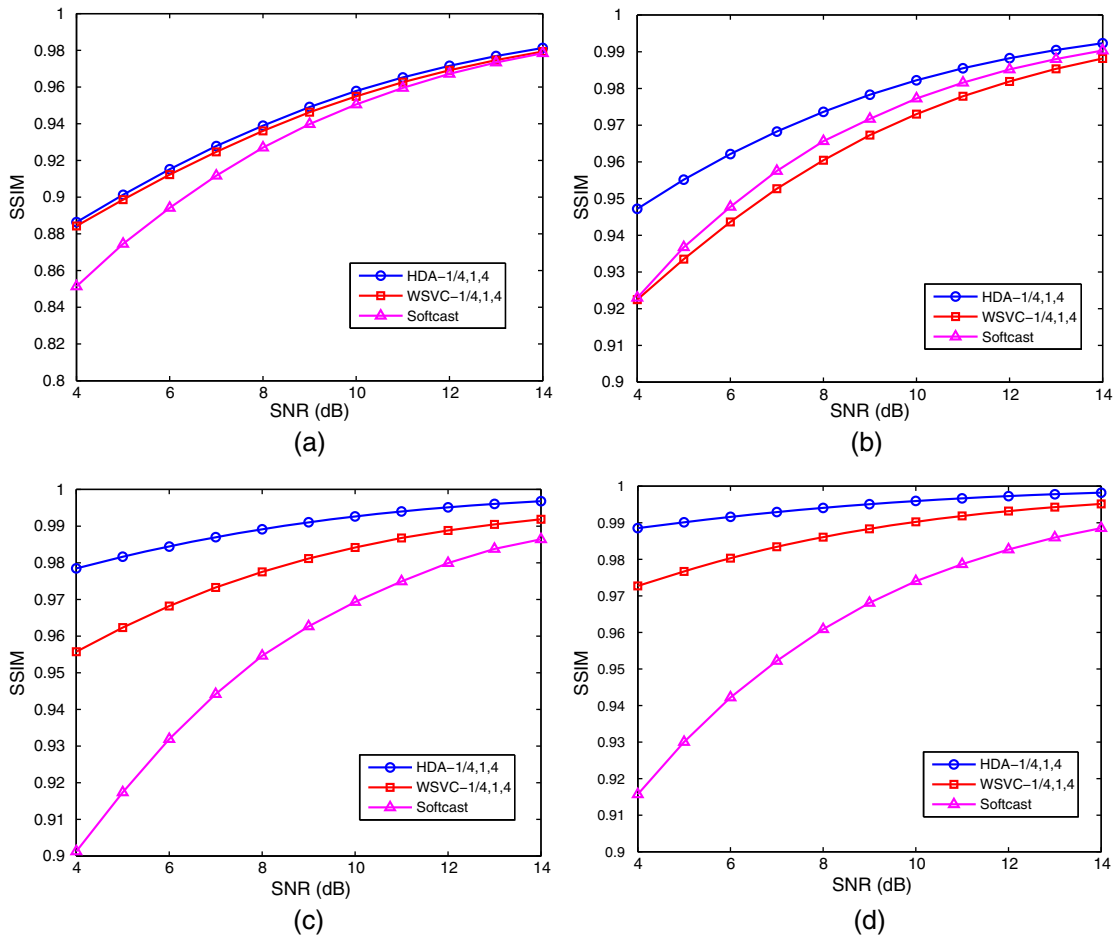


Fig. 7 SSIM comparison in multicast: (a) *football*, (b) *coastguard*, (c) *container*, (d) *news*.

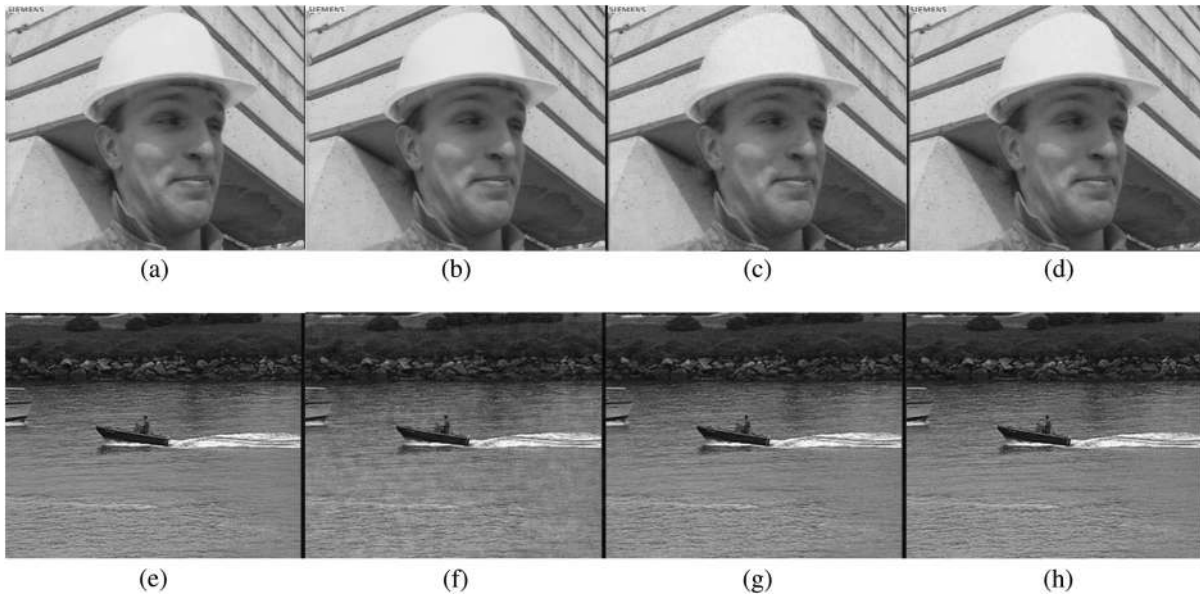


Fig. 8 Visual quality comparison; SNR is 4 dB. The columns from left to right: the original frame, SoftCast, WSVC, and the proposed HDA scheme: (a) PSNR, SSIM, (b) 32.5 dB, 0.84, (c) 37.2 dB, 0.92, (d) 41.2 dB, 0.96, (e) PSNR, SSIM, (f) 32.9 dB, 0.92, (g) 34.2 dB, 0.92, (h) 36.1 dB, 0.95.

is given in Fig. 8. The channel SNR is set to be 4 dB. For all schemes, the PSNRs and SSIMs are above 30 dB and 0.8, respectively; therefore, the figures are reconstructed well. However, there is still slight visual improvement in our scheme.

6 Simulation of Hybrid Digital-Analog Transmission in Multiple-Input Multiple-Output

6.1 Simulation Environment

The test videos are shown in Sec. 5.1. The GOP size is set to 16 in our simulation. We assume there is a 64×64 MIMO channel system²⁶ whose channel matrix \mathbf{H} is a Gaussian complex matrix with zero mean and unit variance in each row. We also assume that the channel estimation is perfect; then the matrix is divided into 64 subchannels by SVD. In our simulation, compression ratios of different QPs and BERs of different FEC code rates (i.e., convolutional code rates in our scheme) have been tested as the prior information. Signal coding rates of each chunk are designed to be the same and equal to the bandwidth.

We compared the proposed HDA video multicast scheme with the reference schemes, namely ParCast²⁶ and ParCast+.³⁵ Our proposed HDA scheme is first coded by H.264/AVC. The coefficients of residuals are divided into 64 $44 \times 36 \times 16$ chunks to be encoded by ParCast. Meanwhile, the modulated digital stream is also divided into chunks with $44 \times 36 \times 16$ values in each chunk. The hybrid signal is transmitted over the MIMO channel with a constant total power budget P . In the ParCast scheme, each GOP is processed by 3D-DCT³⁴ first to remove spatial and temporal correlations. However, ParCast+ adopts a motion-aligned 3-D transform to decorrelate the source since simple ParCast and ParCast+ transmission need only half bandwidth when compared with the HDA scheme. To utilize extra bandwidth efficiently and benefit from a subchannel diversity gain, they use only the high-gain subchannels and discard the low-gain subchannels.

6.2 Parameter Analysis

In this subsection, different parameters of our proposed HDA scheme are set and analyzed. Since results for all tested video sequences exhibit similar trends, only performances of *foreman* are shown. The target MIMO channel SNR ranges from 1 to 14 dB. The minimum γ_m is 1 dB. In order to ensure that the BER is smaller than the target BER 10^{-6} , the following equation should be reached

$$\gamma_0 \leq \frac{\gamma_m}{\sum_{i=1}^N 1/s_i^2}. \quad (27)$$

Thus, with the channel state information, $\gamma_0 = 1$ dB can be derived to cover all the SNRs, and the corresponding parameter \mathfrak{R} is 0.7453.

As shown in Table 1, the FEC code rate 1/4 can be adopted. Then QP is set to 37 without going beyond the channel bandwidth. Figure 9 shows the power assigned to

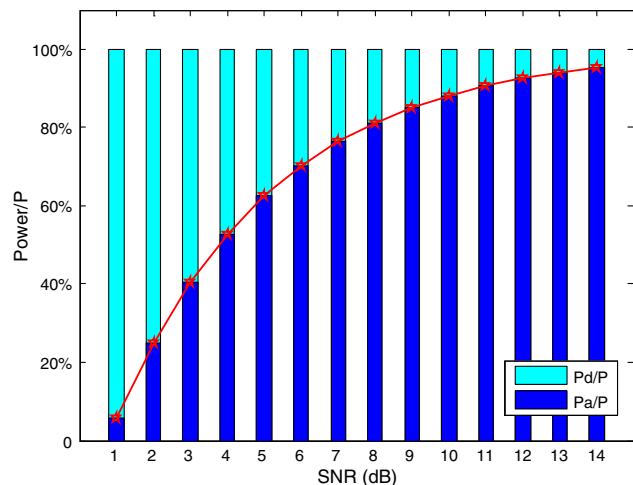


Fig. 9 Power allocated to digital part and analog part. The code rate is 1/4, QP is 37, γ_0 is 1 dB, and γ_m is set to be the same as SNR.

Table 3 Parameters as QP changes at SNR = 1 dB.

QP	37	38	39	40
QPSNR	33.28	32.54	32.05	31.45
Used subchannels	46	42	40	38
Parameter	0.7453	-0.5917	-1.2122	-1.8482
P_a	5.70%	30.68%	39.91%	48.10%
P_d	94.30%	69.32%	60.09%	51.90%
PSNR (dB)	33.83	34.86	34.89	34.83

the digital part P_d and the power assigned to the analog part P_a . It is obvious that P_a increases with SNR ($\gamma_m = \text{SNR}$), and the analog gains of our scheme improve accordingly. Therefore, in our scheme, when the SNR increases, the overall performance is improved by increasing γ_m rather than changing the FEC code rate. In addition, the performance increases gracefully while it is jumpy in traditional digital transmission schemes. Since P_a is less than 10% of P when SNR = 1 dB, the analog gain is much smaller than that of other cases when γ_m is big, and the overall PSNR performance at SNR = 1 dB does not outperform the traditional digital transmission scheme significantly.

As shown in Table 3, even the QPSNR of H.264/AVC decreases with QP, and the degradation is very small. What's more, a large QP results in a high-compression ratio of H.264/AVC, and more low-gain subchannels are discarded for digital transmission. Therefore, performance can be improved by increasing QP properly, as shown in Fig. 10. When we change QP from 37 to 38, the parameter β decreases very quickly since there are four bad subchannels that are discarded. Then, power allocated to the analog signal increases from 5.7% to 30.68%, and the overall performance increases about 1 dB even though there is a 0.7-dB degradation in quantization. However, when we increase QP continuously, the analog gains no longer increase quickly, while

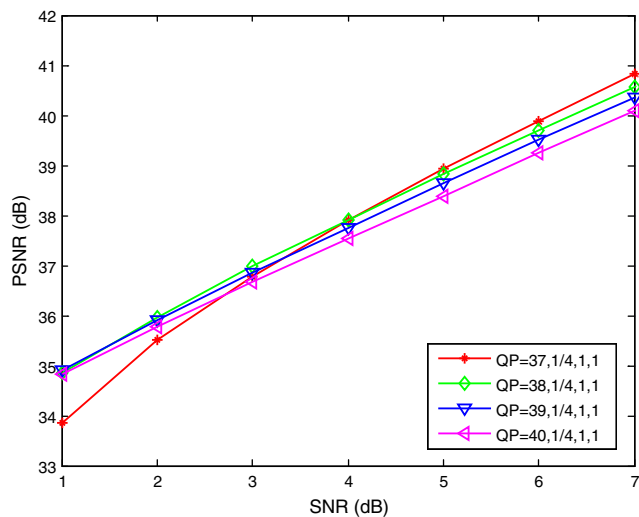


Fig. 10 PSNR of “foreman_cif.yuv” at different QPs. The FEC code rate is 1/4, γ_0 is 1 dB, and γ_m is set to be the same as SNR.

Table 4 Parameters as QP changes at SNR = 5 dB.

QP	37	38	39	40
P_a	62.457%	72.405%	76.079%	79.338%
P_d	37.543%	27.595%	23.921%	20.662%
PSNR (dB)	38.9208	38.8200	38.6386	38.3848

the QPSNR degrades continuously. The result is that the overall PSNR performance no longer increases with QP.

As shown in Table 4, even though P_a increases with QP, the PSNR performance decreases slightly at SNR = 5 dB. In conclusion, when P_a is very small, the PSNR performance can be improved by increasing QP slightly (increasing P_a to be more than 30%; meanwhile, the degradation of QSNR is small).

6.3 Simulation Results

In this subsection, we compare our proposed scheme with the reference schemes. First, the relationship between our proposed scheme H.264/AVC is illustrated. As shown in Fig. 11, if we allocate all the power P to the digital part in our HDA scheme, then the progress degenerates into H.264/AVC video transmission in the designed MIMO system and the PSNR is about 33.3 dB. In most cases, the

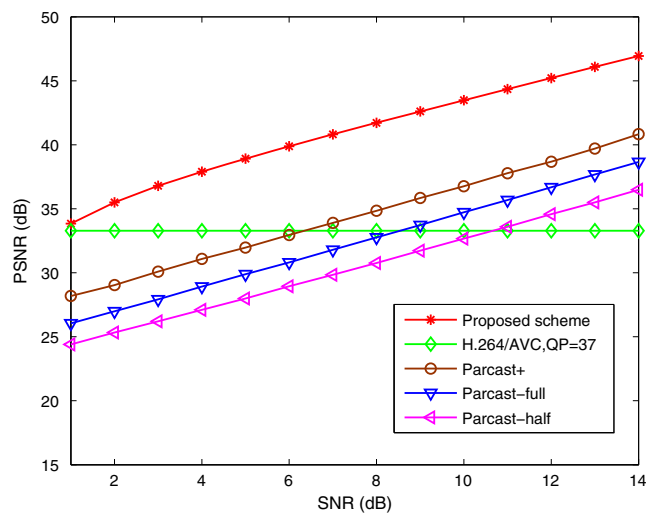


Fig. 11 PSNR of “foreman_cif.yuv” at different schemes. The red curve is the PSNR curve of the proposed scheme with QP = 37. “ParCast-half” is the analog curve with half bandwidth, and “ParCast-full” is the analog curve with bandwidth equal to that of our scheme.

Table 5 Parameters when processing different sequences.

CIF	Football	Coastguard	Akiyo
QP	50	45	31
P_a	63.17%	30.68%	44.08%
P_d	36.83%	69.32%	55.92%

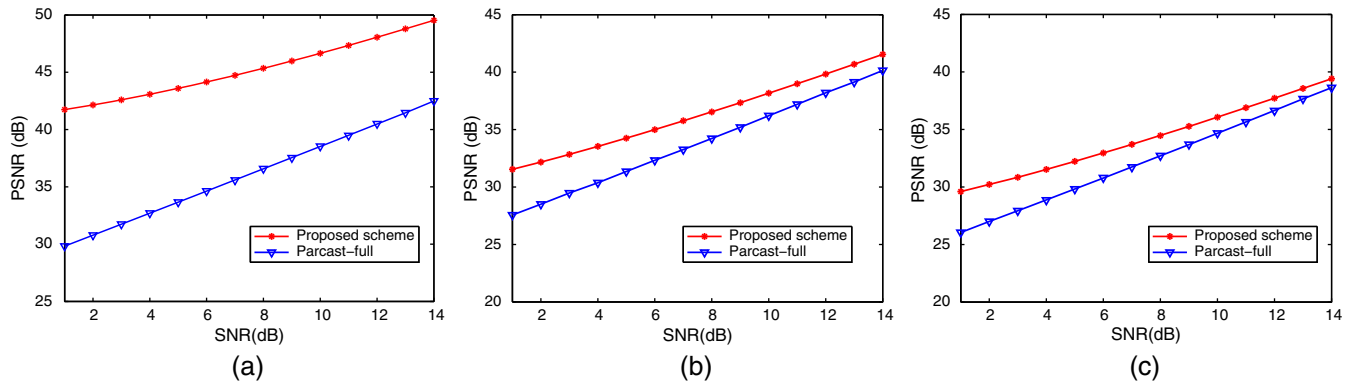


Fig. 12 PSNR comparison in MIMO: (a) *akiyo*, (b) *coastguard*, (c) *football*.

analog data in our scheme should be transmitted to improve the overall performance. Since the ParCast scheme with half bandwidth needs to employ all the subchannels, while the ParCast scheme with full bandwidth can just employ the half higher-gain subchannels, the “ParCast-full” curve outperforms the “ParCast-half” curve.²⁶ As shown in Fig. 11, performance of the proposed HDA scheme is much better than that of the pure ParCast and ParCast+ schemes in MIMO channel systems. What’s more, the proposed scheme can also overcome the cliff effect completely, and the PSNR performance increases with the channel quality.

Apart from the *foreman*, the other CIF video sequences mentioned above are also processed to certify performance of the proposed HDA scheme. Suitable QPs should be adopted, as shown in Table 5. As shown in Fig. 12, our scheme performs about 7 dB better than the pure ParCast scheme.

7 Conclusions

In this paper, an HDA transmission framework is proposed for wireless video multicast and video transmission in MIMO-OFDM systems. Then, the compression rate of H.264/AVC and the FEC code rate are considered jointly based on the channel qualities to guarantee that the video can be compressed effectively and decoded by all the receivers. To overcome the cliff effect in conventional digital transmission schemes, SoftCast is applied to wireless video multicasts and ParCast is applied to video transmissions in MIMO to transmit the H.264/AVC quantization residuals. Then, the residuals can be transmitted with the MMSE. What’s more, an HDA mapping scheme and the corresponding power allocation algorithms are proposed to ensure accurate decoding of the digital signals while maximizing the analog gains. Simulation shows that the proposed HDA schemes can overcome the cliff effect completely and achieve graceful degradation. Moreover, the HDA video multicast can achieve much better PSNR performance when compared with the WSVC scheme, and the HDA transmission in MIMO also outperforms the ParCast scheme significantly.

Acknowledgments

This work has been sponsored by National Science Foundation of China (No. 61201149), the Fundamental Research Funds for the Central Universities (No. 2014ZD03-01), and the Beijing Higher Education Young Elite Teacher

Project. The authors would also like to thank the reviewers for their constructive comments.

References

- B. Barmada et al., “Prioritized transmission of data partitioned H.264 video with hierarchical QAM,” *IEEE Signal Process. Lett.* **12**, 577–580 (2005).
- A. Schertz and C. Week, “Hierarchical modulation—the transmission of two independent DVB-T multiplexes on a single frequency,” *EBU Tech. Rev.* (2003).
- W. Li, “Overview of fine granularity scalability in MPEG-4 video standard,” *IEEE Trans. Circuits Syst. Video Technol.* **11**, 301–317 (2001).
- T. Wiegand et al., “Overview of the H.264/AVC video coding standard,” *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 560–576 (2003).
- T. Stockhammer, M. Hannuksela, and T. Wiegand, “H.264/AVC in wireless environments,” *IEEE Trans. Circuits Syst. Video Technol.* **13**, 657–673 (2003).
- M. Ghandi and M. Ghanbari, “Layered H.264 video transmission with hierarchical QAM,” *J. Visual Commun. Image Represent.* **17**(2), 451–466 (2006).
- Z. Reznic, M. Feder, and S. Freundlich, “Apparatus and method for applying unequal error protection during wireless video transmission,” U.S. Patent 8,006,168 (2011).
- 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.* **17**, 1103–1120 (2007).
- T. Kratochvil, “Hierarchical modulation in DVB-T/H mobile TV transmission,” *Lect. Notes Electr. Eng.* **41**, 333–341 (2009).
- H. Jiang and P. Wilford, “A hierarchical modulation for upgrading digital broadcast systems,” *IEEE Trans. Broadcast.* **51**, 223–229 (2005).
- M. Morimoto et al., “Study on power assignment of hierarchical modulation schemes for digital broadcasting,” *IEICE Trans. Commun.* **E77-B**(12), 1495–1500 (1994).
- M. K. Chang and S. Y. Lee, “Performance analysis of cooperative communication system with hierarchical modulation over rayleigh fading channel,” *IEEE Trans. Wireless Commun.* **8**, 2848–2852 (2009).
- N. Franchi et al., “Multiple description video coding for scalable and robust transmission over IP,” *IEEE Trans. Circuits Syst. Video Technol.* **15**, 321–334 (2005).
- C. Hellge et al., “Mobile TV with SVC and hierarchical modulation for DVB-H broadcast services,” in *IEEE Int. Symp. Broadband Multimedia Systems and Broadcasting, 2009 (BMSB ’09)*, pp. 1–5 (2009).
- A. Vitali, “Multiple description coding—a new technology for video streaming over the internet,” *EBU Tech. Rev.* (312) (2007), Correlating transforms; multiple description coding (MDC).
- V. Goyal, “Multiple description coding: compression meets the network,” *IEEE Signal Process. Mag.* **18**(5), 74–93 (2001).
- K. Ramchandran et al., “Multiresolution broadcast for digital HDTV using joint source/channel coding,” *IEEE J. Sel. Areas Commun.* **11**, 6–23 (1993).
- J. Vass and S. Zhuang, “Multiresolution-multicast video distribution over the Internet,” in *Wireless Communications and Networking Conf. (WCNC 2000)*, Vol. 3, pp. 1457–1461, IEEE (2000).
- C. Bouman and B. Liu, “Multiple resolution segmentation of textured images,” *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 99–113 (1991).
- L. Yu, H. Li, and W. Li, “Wireless scalable video coding using a hybrid digital-analog scheme,” *IEEE Trans. Circuits Syst. Video Technol.* **24**(2), 331–345 (2014).
- D. Katabi, H. Rahul, and S. Jakubczak, “Softcast: one video to serve all wireless receivers,” CSAIL Technical Reports (2009).

22. S. Jakubczak and D. Katabi, "Softcast: one-size-fits-all wireless video," in *SIGCOMM'10 – Proc. SIGCOMM 2010 Conf.*, pp. 449–450, New Delhi, India (2010).
23. S. Jakubczak and D. Katabi, "A cross-layer design for scalable mobile video," in *Proc. Annual Int. Conf. on Mobile Computing and Networking, MOBICOM*, pp. 289–300, Las Vegas, Nevada (2011).
24. X. Fan et al., "WaveCast: wavelet based wireless video broadcast using lossy transmission," in *Visual Communications and Image Processing (VCIP), 2012 IEEE*, pp. 1–6 (2012).
25. X. Fan, F. Wu, and D. Zhao, "D-cast: DSC based soft mobile video broadcast," in *Proc. 10th Int. Conf. on Mobile and Ubiquitous Multimedia (MUM'11)*, pp. 226–235, Beijing, China (2011).
26. X. L. Liu et al., "ParCast: soft video delivery in MIMO-OFDM WLANs," in *Proc. Annual Int. Conf. on Mobile Computing and Networking, MOBICOM*, pp. 233–244, Istanbul, Turkey (2012).
27. T. Schierl et al., "System layer integration of high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.* **22**(12), 1871–1884 (2012).
28. K. E. Psannis, "HEVC in wireless environments," *J. Real Time Image Process.* (2015).
29. K. Psannis, "Efficient redundant frames encoding algorithm for streaming video over error prone wireless channels," *IEICE Electron. Express* **6**(21), 1497–1502 (2009).
30. K. Psannis and Y. Ishibashi, "Enhanced h.264/avc stream switching over varying bandwidth networks," *IEICE Electron. Express* **5**(19), 827–832 (2008).
31. N. Fan, Y. Liu, and L. Zhang, "Hybrid digital-analog video transmission based on H.264/AVC and ParCast in MIMO-OFDM WLANs," in *21st Int. Conf. on Telecommunications (ICT 2014)*, pp. 395–399 (2014).
32. N. Fan et al., "Hybrid digital-analog video multicast scheme based on H.264/AVC and SoftCast," in *Int. Symp. on Wireless Personal Multimedia Communications (WPMC 2014)*, pp. 277–282 (2014).
33. K. H. Lee and D. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.* **24**, 1283–1290 (1976).
34. R. Chan and M. Lee, "3D-DCT quantization as a compression technique for video sequences," in *Proc. Int. Conf. on Virtual Systems and MultiMedia, 1997 (VSMM '97)*, pp. 188–196 (1997).
35. X. L. Liu et al., "ParCast+: parallel video unicast in MIMO-OFDM WLANs," *IEEE Trans. Multimedia* **16**(7), 2038–2051 (2014).
36. D. Halperin et al., "Predictable 802.11 packet delivery from wireless channel measurements," in *SIGCOMM'10 – Proc. SIGCOMM 2010 Conf.*, pp. 159–170, New Delhi, India (2010).

Yu Liu is currently an associate professor and a doctoral supervisor with Beijing University of Posts and Telecommunications (BUPT), China. She received her BS degree in automatic control and her PhD in signal and information processing in 2001 and 2006, respectively, both from BUPT, China. Her main research areas include compressive sensing, distributed source coding, and information processing in wireless networks.

Xiaocheng Lin received his BS degree from BUPT, China, in 2013. Currently, he is pursuing his MS degree at BUPT. His research interests include video coding and transmission, distributed source coding, and compressive sensing.

Nianfei Fan received his BS degree from BUPT in 2012 and his MS degree in 2015. He majored in signal and information processing and did research on compressed sensing and hybrid digital-analog signal processing during his postgraduate study. He worked in RICOH Software Research Center Beijing Co., Ltd., and did research on BLE in 2014. In 2015, he worked in 58.com in Beijing with search development responsibility.

Lin Zhang is the dean of School of Information and Communication Engineering, BUPT, China. He got his BS and PhD in 1996 and 2001, both from BUPT, China. After being a postdoctoral researcher at Information and Communications University, Republic of Korea, and holding a research fellow position in Nanyang Technological University, Singapore, he joined BUPT in 2004. His current research interests are mobile cloud computing and the internet of things.