This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2017.2683490, IEEE Wireless Communications Letters

A Hybrid Digital Analog Scheme for MIMO Multimedia Broadcasting

Bin Tan, Jun Wu, *Senior Member, IEEE,* Hao Cui, Rui Wang, Jian Wu, and Dian Liu

***Abstract*—A new scenario for wireless broadcasting is the antenna heterogeneity. In this letter, we proposed a new hybrid digital analog design to make the users attain the perceptual quality commensurate with their antenna numbers. A high resolution video is decomposed into a base layer and an enhancement layer. The base layer is transmitted by the conventional digital space time coding scheme. The enhancement layer is transmitted by the pseudo analog and spatial multiplexing that is applied to provide signal quality commensurate with antenna number of users. Compared with the conventional frameworks, the proposed system achieves a notable margin in the peak signal-to-noise ratio and much better performances for receivers with more antennas.**

***Index Terms*—Hybrid Digital Analog; MIMO Multimedia Broadcasting; Space Time Code; Antenna Heterogeneity**

I. Introduction

The broadcasting system is a typical point-to-multiple-point communication, which needs to guarantee a successful signal reception of all users in broadcasting coverage area. Thus, a Base Station (BS) generally utilizes the conservative source coding rate and modulation coding scheme (MCS) according to the users's signal qualities at cell edge [1] [2] [3]. Obviously, such kind of broadcasting systems are far from the optimality.

A new paradigm of pseudo analog transmission systems named SoftCast [4], which is able to deal with the signal to noise ratio (SNR) heterogeneity, tends to be optimal. The pseudo analog transmission removes the quantization, entropy coding and channel coding on the transform coefficients, and directly transmits the transform coefficients over the channel. Therefore, the channel noise is characterized by disturbance of transform coefficients. Due to the removal of entropy coding, the pseudo analog transmission is a little less efficient compared with the conventional digital video transmission. Thus, Hybrid Digital-Analog schemes (HDA) are proposed to attain both the compression efficiency and the graceful video quality adaptation [5] [6] [7].

With the development of wireless technology and the booming of all kinds of smart devices [8], the heterogeneous antenna configuration at hand-held devices has become more prevalent. The current broadcasting system does not consider the antenna heterogeneity, and is designed as if all the terminals have the minimum antenna, i.e., 1 antenna. Exploration of the antenna heterogeneity will be a right way to enhance capacities significantly in multiple-antenna broadcasting systems [9] [10] just as dealing with SNR heterogeneity.

In this letter, we design a superposition coding based HDA system which supports the antenna heterogeneity. A video is decomposed into a base layer and an enhancement layer. The base layer is transmitted by the conventional digital space time coding scheme to explore the transmit diversity gain, and all users with heterogeneous antennas can decode the base layer successfully. The enhancement layer is transmitted by the pseudo analog scheme. And the two layers are transmitted using multiple input multiple output (MIMO) techniques for spatial multiplexing. The users with two receiver antennas can decode the two layers successfully while the users with one receiver antenna can also decode limited information. In this way, users obtain signal qualities commensurate with their antenna numbers.

The main contribution of this paper is to propose a MIMO multimedia broadcast scheme to explore antenna heterogeneity and design a joint decoding algorithm for both MIMO and pseudo analog signal detection in one step.

II. System Overview

The proposed HDA system with the antenna heterogeneity is shown in Fig.1, where the BS transmitter is equipped with 2 antennas, and the terminal receivers may have 1 or 2 antennas.

At the sender side, a video sequence is firstly compressed by a H.264 encoder. The compressed source is channel encoded and modulated, then Alamouti space time block coding (STBC) is applied to exploit the transmit antenna diversity. The residue of H.264 can be represented as the difference between the original video source and the reconstructed video, which will be processed according to pseudo analog signal and mapped to 2 antennas in spatial multiplexing.

Two kinds of receivers are designed in the proposed scheme. For a receiver with 2 antennas, the received signals from 2 antennas are combined using maximal ratio combining (MRC) and decoded by a classic Alamouti decoder, where analog signals are regarded as noises. In order to decode the enhancement layer of video, the digital signals are reconstructed and subtracted from the received signals to obtain the noisy analog signals of 2 antennas. Then, a 2x2 MIMO decoding and a linear least square estimation (LLSE) are jointly designed to recover the analog signals. The video is finally reconstructed by the summation of the digital decoded image signals and the analog decoded residues.

B. Tan is with College of Electronics and Information Engineering, Tongji University, Shanghai 201804, and also with College of Electronics and Information Engineering, Jinggangshan University, Ji'an 343009, China (e-mail: 1310499@tongji.edu.cn).

J. Wu (Corresponding author), H. Cui, R. Wang, J. Wu and D. Liu are with College of Electronics and Information Engineering, and also with Key Lab of Ministry of Education in Embedded System and Service Computing, Tongji University (e-mail: wujun@tongji.edu.cn; hao.cui@live.com; {ruiwang,1410490,1310518}@tongji.edu.cn).

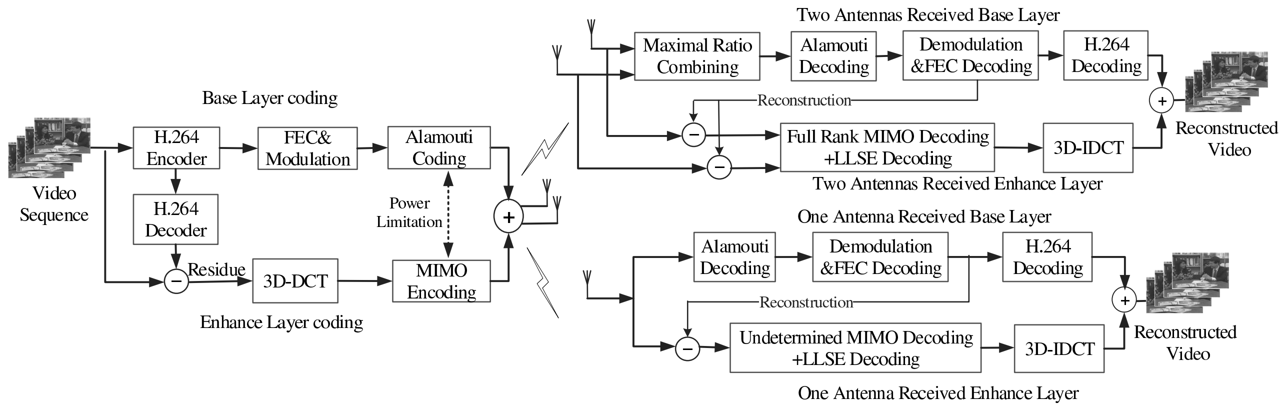

Fig. 1. Framework of the proposed hybrid digital-analog video transmission

For 1-antenna receiver, the digital signal processing chain is similar expect that the MRC is excluded. With the reconstructed noisy analog signal available for only 1 antenna, the decoding of 2 streams is a rank deficient MIMO decoding problem. Taking the sparsity of analog signal into consideration, MIMO decoding is jointly processed with LLSE decoding, which partially plays a role of the sparse signal detector.

III. HDA Scheme Design

A. Signal Design

The proposed system is a $2 \times \{1,2\}$ heterogeneous MIMO system, where the antenna heterogeneity is addressed by Alamouti STBC and 2×2 MIMO. In the digital part, Alamouti STBC encoding [11] is applied to the basic layer information. The channel coded and modulated symbols of a group of pictures (GoP) can be represented as a vector $X_d = \{D_1 D_2 D_3 \cdots, D_\text{n}\}$. By using Alamouti code, the complex symbols for one GoP are mapped to 2 antennas as

$$X_d = \begin{bmatrix} D_1 & -D_2^* & D_3 & -D_4^* & \cdots & D_{n-1} & -D_n^* \\ D_2 & D_1^* & D_4 & D_3^* & \cdots & D_n & D_{n-1}^* \end{bmatrix} \tag{1}$$

where $*$ denotes the complex conjugate.

The source of analog part is the residue of H.264. As shown in Fig.1, the analog source is first transformed by 3 Dimensional (3D)- Discrete Cosine Transform (DCT). The DCT coefficients are divided into equal-sized rectangular chunks, each chunk shares the same scaling factor. After power allocation among chunks, MIMO encoding is followed.

Let $\{a_i, i = 1 \cdots 4n\}$ be the scaled DCT components, every two scaled DCT coefficients are mapped to Inphase and Quadrature component to form a complex symbol. These complex symbols can be represented as a vector $\{A_l, l = 1 \cdots 2n\}$, two complex adjacent symbols are put onto the first and second antenna, respectively. The vector $\{A_l, l = 1 \cdots 2n\}$ will be reshaped to a $2 \times n$ matrix, each column corresponds to one time slot. Thus the MIMO analog data is written as

$$X_a = \begin{bmatrix} A_1 & A_3 & A_5 & A_7 & \cdots & A_{2n-1} \\ A_2 & A_4 & A_6 & A_8 & \cdots & A_{2n} \end{bmatrix} \tag{2}$$

where the complex symbols are designed as

$$A_l = a_{2l-1} + \sqrt{-1} \cdot a_{2l}, l = 1 \cdots 2n. \tag{3}$$

The scaled DCT components $\{a_i, i = 1 \cdots 4n\}$ are

$$a_i = g_j \cdot s_i, j = \lceil i/M \rceil \tag{4}$$

where s_i denotes the i^{th} DCT components in a GoP, M is the chunk size. The power scaling factor g_j for j^{th} chunk [4] is

$$g_j = \lambda_j^{-1/4} \left(\sqrt{\frac{P_a}{\sum_j \sqrt{\lambda_j}}} \right), j = 1 \cdots J \tag{5}$$

where λ_j is a variance of j^{th} chunk, P_a is the analog power.

Finally, the hybrid digital-analog signal X generated as

$$X = \alpha * X_d + \beta * X_a \tag{6}$$

where α is the digital power factor, β is the analog power factor. The further power allocation factors between digital and analog part will be elaborated in the next subsection.

B. HDA Power Allocation

Assume the total power budget for the digital and analog is limited to P_t, the power allocated to analog part and digital part is P_a and P_d, respectively. Obviously, the signal-interference and noise ratio (SINR) is used to describe the digital part

$$\begin{gathered} SINR = 10\log_{10} \frac{P_d}{P_a + N_0} \\ s.t. P_t = P_a + P_d \end{gathered} \tag{7}$$

where N_0 is the channel noise.

We denote SNR_t as SNR threshold when the digital decoding BER is smaller than 10^{-6}. Thus digital SINR should be no less than SNR_t. According to Eq.(7), the HDA power allocation should meets

$$\begin{aligned} P_d &\geq \frac{10^{(SNR_t/10)} \cdot (P_t + N_0)}{1 + 10^{SNR_t/10}} \\ P_a &\leq \frac{N_0 \cdot \left(10^{SINR/10} - 10^{SNR_t/10}\right)}{1 + 10^{SNR_t/10}}. \end{aligned} \tag{8}$$

The optimal power allocation between digital and analog parts should take equality in Eq.(8). Therefore, the power factors can be calculated as

$$\alpha = \sqrt{\frac{P_d}{(\sum X_d {X_d}^T)}}, \beta = \sqrt{\frac{P_a}{(\sum X_a {X_a}^T)}}. \tag{9}$$

IV. Receiver Design

At the receivers, we first decode the base layer data X_d which is encoded by Alamouti STBC. Note that X_a is treated as noise when decoding the digital part. Whether the number of receiver's antenna is 1 or 2, the receiver can successfully decode the base layer with Alamouti decoder to gain the transmit diversity. The noisy received analog signal can be obtained by the received signal minus the reconstructed digital signal. The following subsections focus on how to decode the reconstructed noisy received analog signals.

A. Two antennas

We design a joint receiver to fulfil MIMO decoding and LLSE in one step, which minimizes mean square error globally among all the $2 \times n$ received signals. In separate design, the performance of MIMO decoder with 2 received signals is not good enough. When the MIMO decoded signals are aggregated together for LLSE, it is hard to estimate the equivalent noise variance for LLSE. Without loss of generality, the reconstructed noisy received analog signal $Y_a\left[k\right]$ in the k^{th} time slot can be written as

$$Y_a\left[k\right] = H\left[k\right] \cdot X_a[k] + N'[k] \tag{10}$$

where

$$H\left[k\right] = \begin{bmatrix} h_{11}\left(k\right) & h_{12}\left(k\right) \\ h_{21}\left(k\right) & h_{22}\left(k\right) \end{bmatrix} \tag{11}$$

$$X_a\left[k\right] = \begin{bmatrix} A_{2k-1} & A_{2k} \end{bmatrix}^T \tag{12}$$

$h_{rt}(k)$ denote the path gain from transmit antenna $t = \{1, 2\}$ to receive antenna $r = \{1, 2\}$, $X_a[k]$ and $Y_a[k]$ are the sent and received 2×1 complex vector for time slot k, and $N'[k]$ is the equivalent Gaussian noise.

For the joint decoding design, we rewrite the complex-valued 2×2 MIMO model by its real-valued representation

$$\begin{aligned} Y_r\left[k\right] &= H_r\left[k\right] \cdot X_r[k] + {N'}_r[k] \\ &= H_r\left[k\right] \cdot G[k] \cdot S[k] + {N'}_r[k] \end{aligned} \tag{13}$$

where the real-valued channel matrix is given as

$$H_r\left[k\right] = \begin{bmatrix} R\left(h_{11}\left(k\right)\right) & R\left(h_{12}\left(k\right)\right) & -I\left(h_{12}\left(k\right)\right) & -I\left(h_{21}\left(k\right)\right) \\ R\left(h_{21}\left(k\right)\right) & R\left(h_{22}\left(k\right)\right) & -I\left(h_{21}\left(k\right)\right) & -I\left(h_{22}\left(k\right)\right) \\ I\left(h_{11}\left(k\right)\right) & I\left(h_{12}\left(k\right)\right) & R\left(h_{11}\left(k\right)\right) & R\left(h_{12}\left(k\right)\right) \\ I\left(h_{21}\left(k\right)\right) & I\left(h_{22}\left(k\right)\right) & R\left(h_{21}\left(k\right)\right) & R\left(h_{22}\left(k\right)\right) \end{bmatrix} \tag{14}$$

$\mathrm{R}\left(\cdot\right)$ and $\mathrm{I}\left(\cdot\right)$ represent the real and imaginary parts.

The real-valued scaled coefficient matrix is

$$X_r\left[k\right] = \begin{bmatrix} R(X_a\left[k\right])^T & I(X_a\left[k\right])^T \end{bmatrix}^T. \tag{15}$$

Substitute Eq.(3) into Eq.(15)

$$X_r\left[k\right] = \begin{bmatrix} a_{4k-3} & a_{4k-1} & a_{4k-2} & a_{4k} \end{bmatrix}^T. \tag{16}$$

Applying Eq.(4) into Eq.(16)

$$X_r\left[k\right] = G[k] \cdot \begin{bmatrix} s_{4k-3} & s_{4k-1} & s_{4k-2} & s_{4k} \end{bmatrix}^T \tag{17}$$

where

$$\begin{aligned} G\left[k\right] = diag(&g_{\lceil (4k-3)/M \rceil}, g_{\lceil (4k-1)/M \rceil}, \\ &g_{\lceil (4k-2)/M \rceil}, g_{\lceil (4k)/M \rceil}). \end{aligned} \tag{18}$$

Expand the Eq.(13) from one time slot to n time slots, the received analog signal for one GoP can be expressed as

$$Y_r = H_r \cdot G \cdot S + N_r' \tag{19}$$

where

$$\hat{Y}_r = \begin{bmatrix} Y_r\left[1\right] \\ Y_r\left[2\right] \\ \vdots \\ Y_r\left[n\right] \end{bmatrix}, \hat{S} = \begin{bmatrix} S\left[1\right] \\ S\left[2\right] \\ \vdots \\ S\left[n\right] \end{bmatrix}, \hat{N}_r' = \begin{bmatrix} {N'}_r\left[1\right] \\ {N'}_r\left[2\right] \\ \vdots \\ {N'}_r\left[n\right] \end{bmatrix}, \tag{20}$$

block diagonal matrix

$$\hat{H}_r = diag\left(H_r\left[1\right], H_r\left[2\right], ..., H_r\left[n\right]\right) \tag{21}$$

and $4n \times 4n$ diagonal matrix G is to expand Eq.(18) to $diag(\cdots, g_{\lceil (4k-3)/M \rceil}, g_{\lceil (4k-1)/M \rceil}, g_{\lceil (4k-2)/M \rceil}, g_{\lceil (4k)/M \rceil}, \cdots), k = 1 \cdots n$.

Thus the enhancement layer can be decoded by the LLSE

$$\hat{S} = \Lambda C^T (C \Lambda C^T + N_r') Y_r \tag{22}$$

where Λ is covariance matrix of the received analog signal Y_r, N_r' is the channel noise, the encoding matrix C is

$$C = H_r \cdot G. \tag{23}$$

The LLSE decoded enhancement layer data are added to the digital part, then the original video sequences are reconstructed at the receiver.

B. One antenna

Considering the receiver with 1 antenna, the receiver tries to decode two unknown variables with only one equation, which is under-determine problem. The Eq.(11) can be rewritten as

$$H\left[k\right] = \begin{bmatrix} h_{11}\left(k\right) & h_{12}\left(k\right) \end{bmatrix}. \tag{24}$$

The real-valued 2×4 channel matrix $H_r\left[k\right]$ is

$$H_r\left[k\right] = \begin{bmatrix} R\left(h_{11}\left(k\right)\right) & R\left(h_{12}\left(k\right)\right) & -I\left(h_{11}\left(k\right)\right) & -I\left(h_{12}\left(k\right)\right) \\ I\left(h_{11}\left(k\right)\right) & I\left(h_{12}\left(k\right)\right) & R\left(h_{11}\left(k\right)\right) & R\left(h_{12}\left(k\right)\right) \end{bmatrix}. \tag{25}$$

Then we construct Eq.(19) using Eq.(25) which is an under-determined equation. Considering the DCT coefficients have certain sparsity, the receiver with 1 antenna is possible to solve partial DCT coefficients. We design the similar joint received as Eq.(22) to solve this under-determined equations globally.

V. Simulation

Our experiments are carried out based on 7 standard video test sequences including $Intotree$, $Sheriff$, $Stockholm$, $City$, $ShuttleStart$, $Shields$ and $Jets$, with pixel sizes of

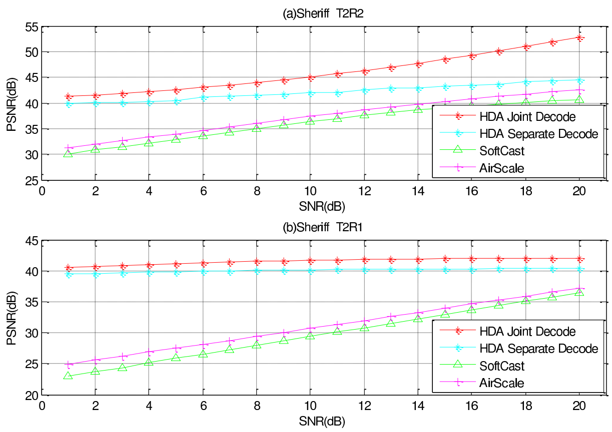

Fig. 2. PSNR performance of $sheriff$ using different schemes in a $2 \times \{1, 2\}$ heterogeneous MIMO system

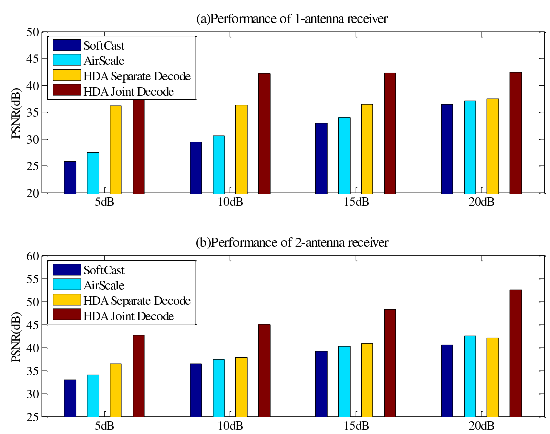

Fig. 3. Average PSNR achieved in a $2 \times \{1, 2\}$ heterogeneous MIMO system

1280×720. In our simulations, the GoP size is set to 16, the DCT coefficients are divided into 64 chunks. Perceived video quality is assessed by the subjective metric PSNR in dB. The Turbo code rate 1/3 and QPSK modulation is employed in digital processing, then QP can be calculated according to the bandwidth, i.e., the QP is set to 28 for $Shields$ sequence. The corresponding SNR threshold in digital part is set to guarantee the digital decoding BER is smaller than 10^{-6} .

We make comparisons with two reference schemes: SoftCast [4] and AirScale [12]. The performances of separate coding for the proposed scheme are also presented. The performance of $Sheriff$ by 2-antenna receiver in Fig.2(a) is constantly about 10dB better than SoftCast and AirScale. The gain comes from two aspects: the base layer in digital coding and the enhancement layer in MIMO coded pseudo analog coding. With the good performance from full rank MIMO decoding, the PSNR is steadily increasing with channel SNR. The performance of $Sheriff$ by 1-antenna receiver in Fig.2(b) is obvious better than reference schemes, achieving about 15dB gain at low SNR and the gain is narrowed down to about 6-7dB with channel SNR increasing to 20dB. The gain is mainly brought by digital part of HDA scheme, the digital gain does not increase with channel SNR since QP and MCS are fixed in digital broadcasting. The rank deficient MIMO decoding with 1-antenna bring a little gain with increase of channel SNR. As we expected, the performances of joint decoding is much better than that of the separate decoding.

We further evaluate the average PSNR performance for 1 and 2-antenna receiver, shown in Fig.3(a) and Fig.3(b), respectively. The PSNR gain is up to 10 dB and 6-12 dB for 1 and 2-antenna receiver. It is obvious that the proposed scheme outperforms the reference schemes.

VI. Conclusion

In this letter, we propose a HDA video broadcasting scheme to make use of the advantages brought by receive antenna heterogeneity. The base layer utilizes space time coding and the enhancement layer is transmitted in pseudo-analog with 2×2 MIMO techniques. With our proposed joint decoding algorithms, the 2-antenna receiver can decode the two streams successfully while the users with one receiver antenna can decode limited information. The simulation results show significantly gain for $2 \times \{1, 2\}$ heterogeneous MIMO system over two reference schemes, respectively. Even the rank deficient MIMO receiver also benefits from the proposed scheme.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (NSFC) under Grant Nos. 61390513, 61502341 and 61571329. This project is also supported by Huawei Innovation Research Plan funding (No. Y-B2015110117).

References

[1] H. Xu, K. Hua, and H. Wang, "Adaptive FEC coding and cooperative relayed wireless image transmission," *Digital Communications and Networks*, vol. 1, no. 3, pp. 213–221, 2014.
[2] M. X. Gong, S. E. Midkiff, and S. Mao, "Design principles for distributed channel assignment in wireless Ad Hoc networks," in *IEEE International Conference on Communications.*, vol. 5, May 2005, pp. 3401–3406 Vol. 5.
[3] M. N. Anjum and H. Wang, "Dynamic scheduling and analysis of real time systems with multiprocessors," *Digital Communications and Networks*, vol. 2, no. 3, pp. 130–138, 2016.
[4] S. Jakubczak and D. Katabi, "A cross-layer design for scalable mobile video," 2011.
[5] L. Yu, H. Li, and W. Li, "Wireless Scalable Video Coding Using a Hybrid Digital-Analog Scheme," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 24, no. 2, pp. 331–345, 2014.
[6] Y. Huang, S. Mao, and S. F. Midkiff, "A control-theoretic approach to rate control for streaming videos," *IEEE Transactions on Multimedia*, vol. 11, no. 6, pp. 1072–1081, Oct 2009.
[7] X. L. Liu, W. Hu, Q. Pu, F. Wu, and Y. Zhang, "ParCast: soft video delivery in MIMO-OFDM WLANs," in *International Conference on Mobile Computing and NETWORKING*, 2012, pp. 233–244.
[8] M. X. Gong, R. Stacey, D. Akhmetov, and S. Mao, "A Directional CSMA/CA Protocol for mmWave Wireless PANs," in *IEEE Wireless Communication and Networking Conference*, April 2010, pp. 1–6.
[9] Z. Zhao, Z. Ding, M. Peng, and Y. Li, "A full-cooperative diversity beamforming scheme in two-way amplify-and-forward relay systems," *Digital Communications and Networks*, vol. 1, no. 1, pp. 57–67, 2015.
[10] H. Cui, C. Luo, C. Chen, and F. Wu, "Scalable Video Multicast for MU-MIMO Systems with Antenna Heterogeneity," *IEEE Transactions on Circuits Systems for Video Technology*, pp. 1–1, 2015.
[11] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct 1998.
[12] H. Cui, C. Luo, C. W. Chen, and F. Wu, "Robust uncoded video transmission over wireless fast fading channel," in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, pp. 73–81.