Imperial College London

Space-Time Design for Deep Joint Source Channel Coding of Images over MIMO Channels

Chenghong Bian, Yulin Shao, Haotian Wu, and Deniz Gunduz

1. Introduction:

We consider novel deep joint source-channel coding (DeepJSCC) algorithms for wireless image transmission over multi-input multi-output (MIMO) Rayleigh fading channels, when channel state information (CSI) is available only at the receiver. We consider two different schemes; one exploiting the spatial diversity and the other exploiting the spatial multiplexing gain of the MIMO channel, respectively. For the former, we utilize an orthogonal space-time block code (OSTBC) to achieve full diversity and increase the robustness against channel variations. In the latter, we directly map the input to the antennas, where the additional degree of freedom can be used to send more information about the source signal.

2. System model:

We denote the input image by $\boldsymbol{X} \in \mathbb{R}^{C\times H\times W}$, where C, H and W are the number of color channels, height, and width of the image.

Transmitter:

DeepJSCC encoder, $f(\cdot)$:

$$\mathbf{z} = \mathbf{f}(\boldsymbol{S});\ \mathbf{z} \in \mathbb{R}^{\ell}$$

Space-time Mapping, $STM(\cdot)$:

$$\boldsymbol{S} = STM(\boldsymbol{z})$$

$\boldsymbol{S} \in \mathbb{C}^{N_t\times k}$; $||\boldsymbol{S}||_F^2 = N_t k$, where k denotes the number of channel uses.

power normalization: Given power P, transmit $\sqrt{P}\boldsymbol{S}$.

bandwidth ratio: $\rho = \frac{k}{CHW}$.

Input-output relation for MIMO:

$$\boldsymbol{Y} = \sqrt{P}\boldsymbol{H}\boldsymbol{S} + \boldsymbol{N};$$

With $\boldsymbol{Y} \in \mathbb{R}^{N_r\times k}$; $\boldsymbol{N} \in \mathbb{C}^{N_r\times k}$; $\boldsymbol{N}_{ij} \sim \mathcal{CN}(0, \sigma^2)$.

Receiver:

Assumes Perfect CSIR.

MIMO Detector: $Det(\cdot,\cdot)$

$$\widehat{\boldsymbol{S}} = Det(\boldsymbol{Y}, \boldsymbol{H});$$

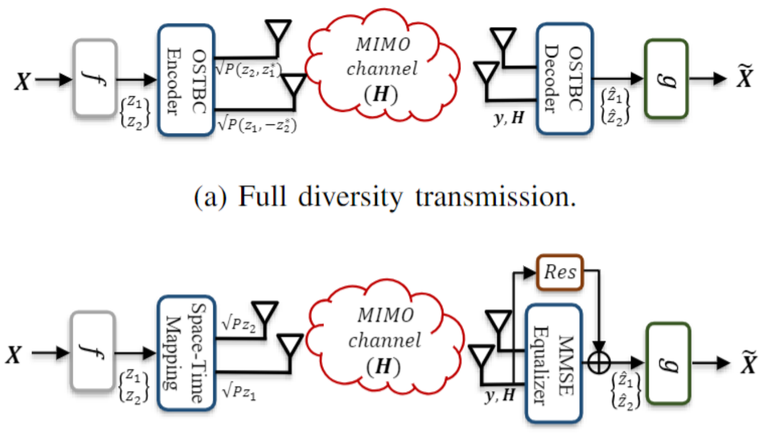
DeepJSCC decoder: $g(\cdot)$,

$$\widehat{\boldsymbol{X}} = g(\widehat{\boldsymbol{S}})$$

3. Proposed Method:

We give the exact implementation for $STM(\cdot)$ and the MIMO detector, $Det(\cdot,\cdot)$. Different implementation of the $STM(\cdot)$ and the corresponding $Det(\cdot,\cdot)$ modules lead to different diversity-multiplexing trade-off. We are particularly interested in exploring the full-diversity transmission (Fig. 1(a)) and the multiplexing case (Fig. 1(b)) detailed as follows.

Source code available at:
https://github.com/aprilbian/ST_JSCC



(a) Full diversity transmission.

(b) Multiplexing scheme.

Fig.1 Block diagrams for full-diversity scheme as well as the multiplexing scheme.

Full Diversity Transmission: We realize the full diversity transmission via OSTBC. The $N_t = 2$ and $N_t = 3$ cases are considered as follows:

OSTBC design for $N_t = 2$:

$$\mathcal{G}(z) = \begin{pmatrix} z_1 & -z_2^* \\ z_2 & z_1^* \end{pmatrix}$$

The $STM(\cdot)$ in this case partitions the input sequence into $\ell/2$ pairs where each pair, denote as (z_1, z_2), is organized as shown in the above matrix.

OSTBC design for $N_t = 3$:

$$\mathcal{G}(z) = \begin{pmatrix} z_1 & -z_2 & -z_3 & -z_4 & z_1^* & -z_2^* & -z_3^* & -z_4^* \\ z_2 & z_1 & z_4 & -z_3 & z_2^* & z_1^* & z_4^* & -z_3^* \\ z_3 & -z_4 & z_1 & z_2 & z_3^* & -z_4^* & z_1^* & z_2^* \end{pmatrix}$$

Similar with the $N_t = 2$ case, the $STM(\cdot)$ processes 4 elements in a row in a way shown above.

To efficiently detect/estimate the messages, $\mathbf{z} = [z_1, z_2]^T$ for $N_t = 2$, while $\mathbf{z} = [z_1, z_2, z_3, z_4]^T$ for $N_t = 3$, we notice $\boldsymbol{Y} = \sqrt{P}\boldsymbol{H}\mathcal{G}(\boldsymbol{z})$ can be re-organized into $\boldsymbol{y}_{eff} = \sqrt{P}\boldsymbol{H}_{eff}\boldsymbol{z} + \mathbf{n}_{\mathbf{eff}}$, and we give $\boldsymbol{H}_{eff}$ for $N_t = 3$ case as an example:

$$H_{eff} = \begin{pmatrix} h_1' & h_2' & h_3' & 0' & h_1^T & h_2^T & h_3^T & 0^T \\ h_2' & -h_1' & 0' & h_3' & h_2^T & -h_1^T & 0^T & h_3^T \\ h_3' & 0' & -h_1' & -h_2' & h_3^T & 0^T & -h_1^T & -h_2^T \\ 0' & -h_3' & h_2' & -h_1' & 0^T & -h_3^T & h_2^T & -h_1^T \end{pmatrix}$$

Note that $\boldsymbol{H}_{eff}'\boldsymbol{H}_{eff} = 2\sum_i |\boldsymbol{h}_i|^2\, \boldsymbol{I}_4$ implying z_i can be equalized independently as:

$$\hat{m}_i = 2\sqrt{P}\sum_{j=1}^{3} |\boldsymbol{h}_j|_2^2 z_i + \left(\boldsymbol{H}_{eff}\boldsymbol{n}_{eff}\right)_{\boldsymbol{i}}$$

$$\hat{z}_i = \hat{m}_i / (2P\textstyle\sum_{\boldsymbol{j}}^{\boldsymbol{3}} |\boldsymbol{h}_j|_2^2 + \sigma^2)$$

With the estimated $\hat{\mathbf{z}}$, the reconstructed image is simply $\widehat{\boldsymbol{X}} = g(\hat{\boldsymbol{z}})$.

Multiplexing scheme:

In this case, $\ell = N_t k$, i.e., no redundancy is introduced. $\boldsymbol{z}$ is mapped sequentially to form $\boldsymbol{S}$.

$$\boldsymbol{S} = Reshape(\boldsymbol{z}) = STM(\boldsymbol{z})$$

The residual-assisted MMSE equalization is applied:

MMSE equalization:

$$\boldsymbol{Z}_{MMSE}[i] = \boldsymbol{H}'\left(\boldsymbol{H}\boldsymbol{H}' + \frac{\sigma^2}{P}\boldsymbol{I}_{N_r}\right)^{-1}\boldsymbol{Y}[i]$$

NN-aided calibration

$$\widehat{\boldsymbol{Z}}[i] = \boldsymbol{Z}_{MMSE}[i] + Res(\boldsymbol{Y}[i], \boldsymbol{H})$$

Where $i \in [1, k]$. Then, $\widehat{\boldsymbol{X}} = \boldsymbol{g}(\hat{\boldsymbol{z}})$.

Loss function:

$$\mathcal{L} = ||\boldsymbol{X} - \widehat{\boldsymbol{X}}||_F^2$$

Identical to both schemes.

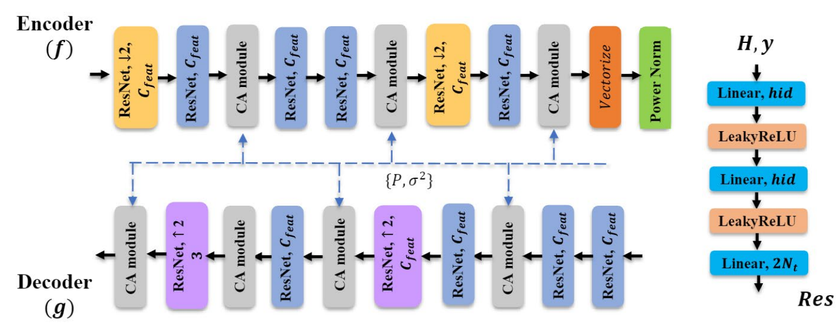


Fig.2 Neural Network Architecture for DeepJSCC encoder and decoder

Neural Network Architecture for $\boldsymbol{f}(\cdot), \boldsymbol{g}(\cdot)$:

The structures are shown in Fig. 2 where we adopt 2d CNN as the backbone and pixel shuffling is used inside the module $'ResNet, \uparrow 2'$ to upsample input features.

SNR-adaptive module: Self attention (SA) modules are adopted which take (P, σ^2) as additional inputs in both DeepJSCC encoder/decoder to ensure that the system is robust to different SNR regimes.

4. Experiments:

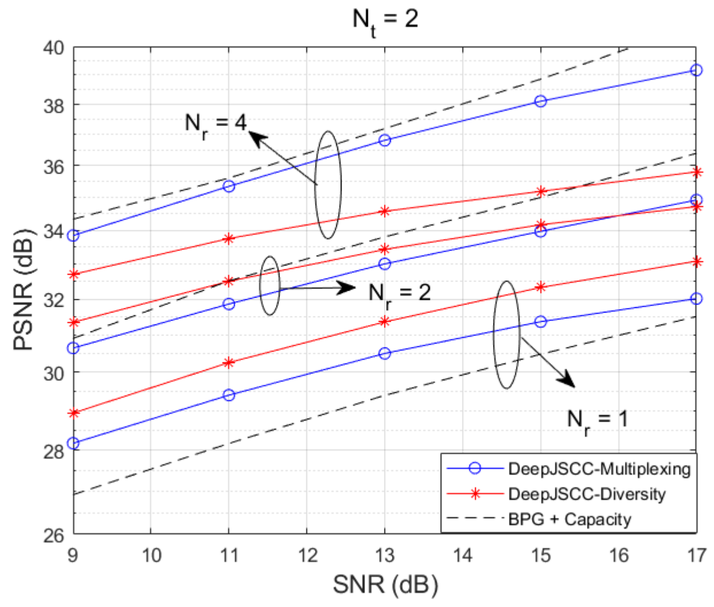


Fig. 3(a) Diversity v.s. Multiplexing with $N_t = 2$

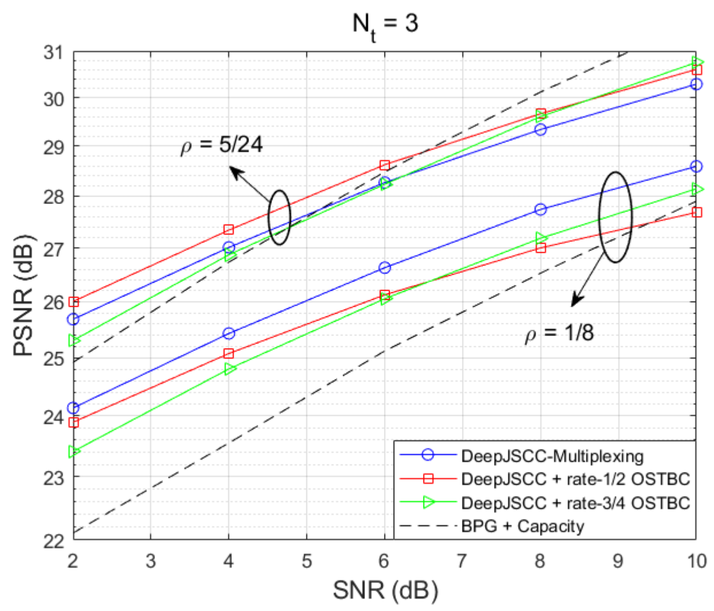


Fig. 3(b) Diversity v.s. Multiplexing with $N_t = 3$ and $N_r = 1$

Simulation results show that the diversity scheme outperforms the multiplexing scheme for lower signal-to-noise ratio (SNR) values and a smaller number of receive antennas at the AP. When the number of transmit antennas is greater than two, however, the full-diversity scheme becomes less beneficial. We also show that both the diversity and multiplexing schemes can achieve comparable performance with the state-of-the-art BPG algorithm delivered at the instantaneous capacity of the MIMO channel, which serves as an upper bound on the performance of separation-based practical systems.