

# 一种基于深度学习的 FDD 大规模 MIMO 系统 CSI 反馈方法

廖 勇<sup>1</sup>, 姚海梅<sup>1</sup>, 花远肖<sup>1</sup>, 赵 砚<sup>2</sup>

(1. 重庆大学通信与测控中心, 重庆 400044; 2. 中国人民解放军 61212 部队, 北京 100043)

**摘 要:** 针对频分双工(Frequency Division Duplexing, FDD)大规模多入多出(Multiple-Input Multiple-Output, MIMO)系统中现有信道状态信息(Channel State Information, CSI)反馈方法复杂度高、反馈精度低的问题, 本文提出一种基于深度学习的 CSI 压缩反馈方法. 该方法首先采用卷积神经网络(Convolutional Neural Network, CNN)提取信道特征矢量, 然后利用最大池化(Maxpooling)网络压缩 CSI, 最后考虑到大规模 MIMO 信道存在空间相关性的特点, 分别对单用户和多用户场景使用双向长短期记忆(Bidirectional Long Short-Term Memory, Bi-LSTM)网络和双向卷积长短期记忆(Bidirectional Convolutional Long Short-Term Memory, Bi-ConvLSTM)网络对 CSI 进行重构. 本文利用大规模 MIMO 信道数据对所提的深度学习网络进行离线训练, 该网络学习到的信道信息能充分表征信道的状态. 仿真结果表明, 与已有的典型 CSI 反馈方法相比, 本文所提方法反馈精度更高, 运行时间更短, 系统性能提升明显.

**关键词:** 频分双工; 大规模多入多出; 信道状态信息; 深度学习; 压缩反馈; 空间相关性

**中图分类号:** TN911.72 **文献标识码:** A **文章编号:** 0372-2112 (2020)06-1182-08

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2020.06.020

## CSI Feedback Method Based on Deep Learning for FDD Massive MIMO Systems

LIAO Yong<sup>1</sup>, YAO Hai-mei<sup>1</sup>, HUA Yuan-xiao<sup>1</sup>, ZHAO Yan<sup>2</sup>

(1. Center of Communication and TT&C, Chongqing University, Chongqing 400044, China;

2. 61212 Unit of the People's Liberation Army, Beijing 100043, China)

**Abstract:** Existing channel state information (CSI) feedback methods for frequency division duplexing (FDD) multiple-input multiple-output (MIMO) systems have high complexity and low feedback accuracy. In this paper, a deep learning-based CSI compression feedback method is proposed. The method first uses the convolutional neural network (CNN) to extract the channel feature vector, and then uses the maximum pooling (Maxpooling) network to compress the CSI. Finally, considering the spatial correlation of the massive MIMO channel, bidirectional long short-term memory (Bi-LSTM) network and bidirectional convolution long-term memory (Bi-ConvLSTM) network are used for single-user and multi-user scenarios respectively to recover the CSI. In this paper, the deep learning network is trained offline using massive MIMO channel data, the channel information learned by the network can fully characterize the states of the channel. The simulation results show that compared with the existing typical CSI feedback methods, the proposed method has higher feedback accuracy, shorter running time and better system performance.

**Key words:** FDD; massive MIMO; channel state information; deep learning; compression feedback; spatial correlation

### 1 引言

大规模(Multiple-Input Multiple-Output, MIMO)技术具有频谱效率高、系统容量大等诸多优点<sup>[1]</sup>. 在频分双

工(Frequency-Division Duplexing, FDD)大规模 MIMO 系统中, 基站(Base Station, BS)通过接收端反馈得到下行的信道状态信息(Channel State Information, CSI)<sup>[2]</sup>, 但大规模天线阵列的使用导致信道反馈量急剧增加<sup>[3]</sup>.

在大规模 MIMO 系统中,由于使用大规模天线阵列且天线间密集排列,从而使信道具有很强的空间相关性.近年来许多学者提出利用压缩感知(Compressed Sensing, CS)来实现 CSI 压缩反馈.文献[4]中提出了卡洛南-洛伊变换(Karhunen-Loeve Transform, KLT)和离散余弦变换(Discrete Cosine Transform, DCT)相结合的两种自适应压缩方法.文献[5]同样利用 DCT 和 KLT 作为稀疏基,对信道进行独立主元变换后进行压缩.但是文献[4,5]中当信道变化较快时,BS 很难及时地获取信道相关信息,从而造成反馈精度的下降,并且当天线数目增大时,导致反馈量开销增大.文献[6]提出基于主成分分析(Principal Component Analysis, PCA)的压缩反馈方法,其使用的压缩矩阵是通过前一长时阶段估计得到的,但 BS 难以得到及时、准确的 CSI.

近几年来已有研究人员将深度学习应用于无线通信系统物理层.文献[7]对深度学习用于物理层的研究进行了综述.文献[8]提出一种面向高速移动场景的基于深度学习的信道估计方法.文献[9]提出基于深度学习的联合信道估计与信号检测方法.文献[10]提出一种基于深度学习联合的信道均衡和译码.文献[11]提出基于深度学习的大规模 MIMO 的 CSI 反馈,其利用 CsiNet 网络来实现大规模 MIMO 系统的 CSI 反馈,与传统的 CS 方法相比, CsiNet 网络能获得较好的性能,但是该网络训练参数多,并且只利用少量的卷积层和全连接层对数据进行简单地特征提取来完成 CSI 的压缩和恢复,在大规模 MIMO 条件下并未充分利用天线之间的相关性.

因此,针对已有 CSI 反馈方法中计算复杂度高、反馈精度低的问题,考虑到大规模 MIMO 信道数据具有空间相关性的特点,本文提出了一种基于深度学习的具有低反馈开销、高反馈精度的 CSI 压缩反馈方法.该方法中使用的深度学习网络模型适用于单用户和多用户场景,信号处理包括压缩处理和解压处理两部分.所提方法采用卷积神经网络(Convolutional Neural Network, CNN)来提取信道特征矢量,通过用最大池化(Maxpooling)网络来对卷积后的 CSI 进行降维,然后利用大规模 MIMO 信道数据具有空间相关性的特点,使用双向长短期记忆(Bidirectional Long Short-Term Memory, Bi-LSTM)网络或双向卷积长短期记忆(Bidirectional Convolutional Long Short-Term Memory, Bi-ConvLSTM)网络来精确地恢复原始的 CSI.本文对比了所提方法与一些已有的 CSI 反馈方法的运行时间,并分析了它们在不同大规模天线数目下的反馈性能.

## 2 系统模型

考虑 FDD 模式下多用户大规模 MIMO 无线通信系

统,发射端配置  $N_t \gg 1$  根发射天线,用户端配置  $N_r \geq 1$  根接收天线,小区中有  $K$  个移动用户,信道为瑞利平坦衰落信道.系统模型如图 1 所示.

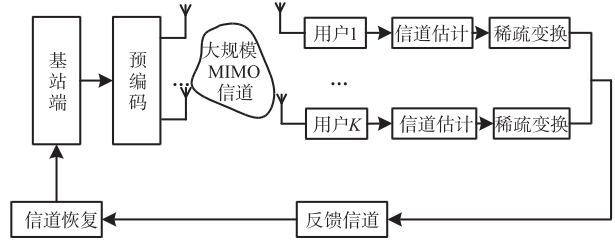


图1 多用户大规模MIMO系统信道反馈框图

在 FDD 系统中,首先考虑下行方向,BS 端对数据流进行预编码处理,然后通过无线信道将信号发送出去.用户端接收到信号后首先通过信道估计方法获得 CSI,然后对 CSI 进行处理,最后通过上行信道反馈给 BS,BS 恢复出 CSI 并利用获得的 CSI 进行预编码等操作.

假设小区中一共有  $K$  个用户,从 BS 到用户  $k$  的信道矩阵用  $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$  表示,令  $\mathbf{V}$  和  $\mathbf{s}$  分别为  $N_t \times KN_t$  的预编码矩阵和  $KN_t \times 1$  的发送信号向量.第  $k$  个用户的接收信号为:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{V} \mathbf{s} + \mathbf{n}_k \quad (1)$$

其中  $\mathbf{n}_k$  表示用户  $k$  所接收到的均值为零,方差为 1 的加性高斯白噪声.所有  $K$  个用户接收到的信号为:

$$\mathbf{y} = \mathbf{H} \mathbf{V} \mathbf{s} + \mathbf{n} \quad (2)$$

其中  $\mathbf{y} = (\mathbf{y}_1^T, \dots, \mathbf{y}_K^T)^T$ ,  $\mathbf{H} = (\mathbf{H}_1^H, \dots, \mathbf{H}_K^H)^H$ ,  $\mathbf{n} = (\mathbf{n}_1^T, \dots, \mathbf{n}_K^T)^T$ ,  $(\cdot)^T$  表示矩阵的转置.假设  $E[\mathbf{s} \mathbf{s}^H] = \mathbf{I}$ ,  $E[\mathbf{n} \mathbf{n}^H] = \mathbf{I}$ ,BS 端发射功率限制为  $P$ ,则有  $\text{tr}(\mathbf{V} \mathbf{V}^H) \leq P$ ,  $\text{tr}(\cdot)$  表示矩阵的迹.

在大规模 MIMO 系统中,由于发射端和接收端的天线数较多,天线间距较小,因此天线之间存在一定的相关性,具有空间相关性的  $N_r \times N_t$  信道矩阵可以建模为<sup>[4]</sup>:

$$\mathbf{H} = \frac{1}{\sqrt{\text{tr}(\mathbf{R}_t)}} \mathbf{R}_r^{\frac{1}{2}} \mathbf{H}_{\text{iid}} \mathbf{R}_t^{\frac{1}{2}} \quad (3)$$

其中,  $\mathbf{R}_r \in \mathbb{R}^{N_r \times N_r}$  表示接收相关矩阵,  $\mathbf{R}_t \in \mathbb{R}^{N_t \times N_t}$  表示发送相关矩阵,  $\mathbf{H}_{\text{iid}} \in \mathbb{C}^{N_r \times N_t}$  表示均值为 0、方差为 1 的独立同分布(Independently Identically Distribution, IID)的复高斯随机矢量.

假设均匀线性阵列天线安装在无线链路的两端,因此空间相关矩阵可通过 Clarke<sup>[12]</sup> 模型得到,  $\mathbf{R}_t$  或  $\mathbf{R}_r$  里的任意一个元素都为  $r_{ij}$ ,  $r_{ij}$  为发送端或接收端第  $i$  根天线与第  $j$  根天线之间的相关系数,其表达式如下:

$$r_{ij} = J_0\left(\frac{2\pi d_{ij}}{\lambda}\right) = J_0\left(\frac{2\pi d}{\lambda} |i-j|\right) \quad (4)$$

其中,  $J_0(\cdot)$  表示第一类零阶贝塞尔函数,  $i$  和  $j$  表示天线阵元序号,  $d_{ij}$  表示第  $i$  根与第  $j$  根发射或者接收天线之间的距离,  $d$  表示相邻发射或者接收天线之间的距离,  $\lambda$  表示载波波长.

### 3 基于深度学习的大规模 MIMO 信道反馈方法

#### 3.1 网络结构

本文提出的基于深度学习的网络结构如图 2 所示, 网络主要分为压缩处理和解压处理两部分. 对于单用户, 网络使用的是 2D CNN、2D Maxpooling 和 Bi-LSTM; 对于多用户, 网络使用的是 3D CNN、3D Maxpooling 和 Bi-ConvLSTM. 下面对网络各部分结构进行详细的介绍.

##### (1) CNN 网络与 Maxpooling 网络

在单用户中, 2D CNN 网络主要用于提取原始信道  $\mathbf{H}$  特征值. 通常, 一个 2D CNN 网络由多个卷积滤波器组成, 通过滑动窗口对数据进行卷积求和. 设  $\mathbf{W}$  为卷积滤波器, 对滑动窗口内的数据加权求和得到卷积输出. 因此 CNN 的变换公式为:

$$\mathbf{x} = f(\mathbf{H} * \mathbf{W} + \mathbf{b}) \quad (5)$$

其中,  $\mathbf{x}$  为卷积输出,  $\mathbf{b}$  为偏置,  $f$  为激活函数,  $*$  表示卷积操作. 本文卷积层使用的是 ReLU 激活函数.

2D Maxpooling 网络主要用于对卷积后的特征矢量降维, 通过一个池化窗口来寻找滤波器输出结果之间的最大值, 其变化表达式为  $\mathbf{x}' = \max(\mathbf{x})$ , 其中  $\mathbf{x}'$  表示池化后的特征矢量.

在多用户中, 3D CNN 与 2D CNN 卷积原理类似, 只是 3D CNN 的输入和输出比 2D CNN 多了一维. 同理,

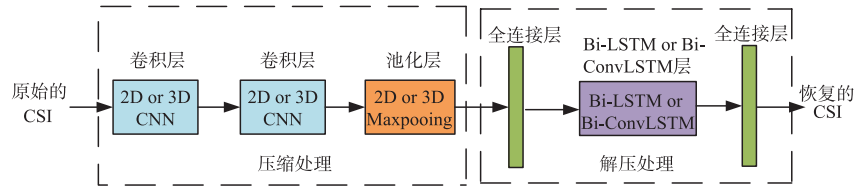


图2 所提的CSI反馈网络结构

#### 3.2 信号处理数据流

本文所提的CSI反馈网络的数据流程如图3所示, 下面分别对压缩处理和解压处理数据流进行详细的描述.

##### (1) 压缩处理

在本文中, 所提学习网络的输入数据为一个  $N_t \times N_r \times K$  大小的信道矩阵, 即其输入数据为  $\mathbf{H} \in \mathbb{C}^{N_t \times N_r \times K}$ . 因为信道数据开始为复数信号, 因此在对所提的学习网络进行输入之前需要对数据进行预处理, 这里把输入数据的实部和虚部提取出来然后将它们作为一个新的维度, 于是输入数据在单用户情况下为  $\mathbf{H}'_s \in \mathbb{R}^{N_t \times N_r \times 2}$ , 多

2D Maxpooling 是对 2D CNN 网络卷积后的特征进行降维, 而 3D Maxpooling 是对 3D CNN 网络卷积后的特征进行降维.

##### (2) Bi-LSTM 网络与 Bi-ConvLSTM 网络

Bi-LSTM 和 Bi-ConvLSTM 网络旨在利用大规模 MIMO 系统中天线之间的相关性来对 CSI 进行预测. 对于正向预测, 后一天线的 CSI 通过前一天线的 CSI 预测得到, 对于反向预测, 前一天线的 CSI 通过后一天线预测得到. 该网络充分利用前向和后向天线的 CSI 来进一步地提升 CSI 的恢复精度.

Bi-LSTM 网络主要用于单用户的 CSI 预测, Bi-LSTM 为两个单向 LSTM 网络的组合, 其中一个 LSTM 网络进行数据的正向预测, 另一个 LSTM 网络进行反向预测. LSTM 网络主要由若干个 LSTM 单元组成, 网络使用双正切激活函数. 将 Bi-LSTM 网络每个时间步的输出变换简写成下式:

$$\mathbf{a}^{(t)} = \text{Bi-LSTM}(\mathbf{I}_l^{(t-1)}, \mathbf{I}_r^{(t-1)}, \mathbf{x}_l^{(t)}, \mathbf{x}_r^{(t)}, \Theta) \quad (6)$$

其中  $\mathbf{I}_l^{(t-1)}$  和  $\mathbf{I}_r^{(t-1)}$  分别为 Bi-LSTM 网络第  $t-1$  根天线的正向隐藏层矢量和反向隐藏层矢量,  $\mathbf{x}_l^{(t)}$  和  $\mathbf{x}_r^{(t)}$  分别为 Bi-LSTM 网络第  $t$  根天线的正向输入和反向输入.  $\Theta$  和  $\text{Bi-LSTM}(\cdot)$  分别表示 Bi-LSTM 网络中的所有参数和变换函数.

Bi-ConvLSTM 是用于多用户的 CSI 预测, 并且 Bi-ConvLSTM 与 Bi-LSTM 结构有点差别, Bi-ConvLSTM 将 Bi-LSTM 中输入与各个门之间的连接由前馈式替换成了卷积运算.

户情况下为  $\mathbf{H}'_m \in \mathbb{R}^{N_t \times N_r \times K \times 2}$ .

输入数据经过预处理之后就分别送入两个 2D CNN 或 3D CNN 网络. 在单用户条件下, 第一个和第二个 2D CNN 卷积核数目分别为 4 个和 2 个, 卷积核大小都为  $3 \times 3$ . 第一个和第二个 2D CNN 网络的输出分别为  $\mathbf{H}''_s$  和  $\mathbf{H}'''_s$ :

$$\mathbf{H}''_s = f(\mathbf{H}'_s * \mathbf{W}_{1,s} + \mathbf{b}_{1,s}) \quad (7)$$

$$\mathbf{H}'''_s = f(\mathbf{H}''_s * \mathbf{W}_{2,s} + \mathbf{b}_{2,s}) \quad (8)$$

其中  $\mathbf{H}''_s \in \mathbb{R}^{N_t \times N_r \times 4}$ ,  $\mathbf{W}_{1,s}$  和  $\mathbf{b}_{1,s}$  分别为第一个 2D CNN 的权重和偏置.  $\mathbf{H}'''_s \in \mathbb{R}^{N_t \times N_r \times 2}$ ,  $\mathbf{W}_{2,s}$  和  $\mathbf{b}_{2,s}$  分别为第二个 2D CNN 的权重和偏置.

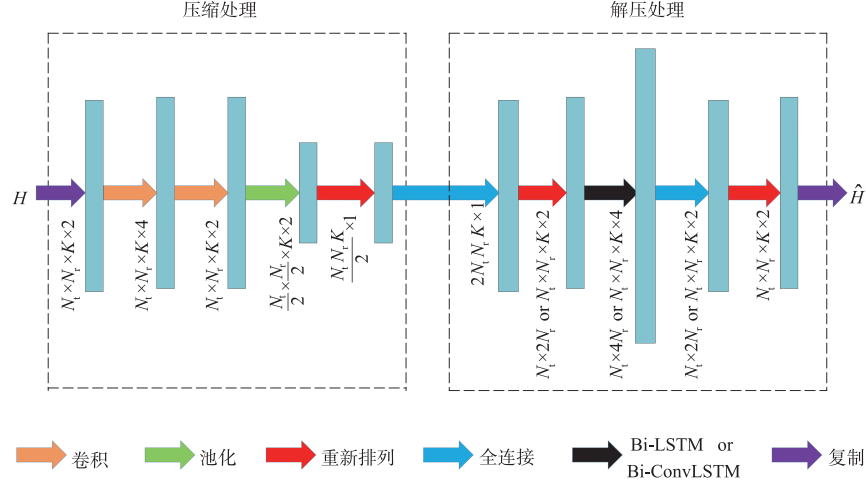


图3 所提的大规模MIMO CSI反馈网络中的数据流

在多用户条件下,第一个和第二个3D CNN 卷积核数目分别为4个和2个,卷积核大小都为 $3 \times 3 \times 3$ . 第一个和第二个3D CNN 网络的输出分别为 $\mathbf{H}'_m$ 和 $\mathbf{H}''_m$ :

$$\mathbf{H}'_m = f(\mathbf{H}'_m * \mathbf{W}_{1,m} + \mathbf{b}_{1,m}) \quad (9)$$

$$\mathbf{H}''_m = f(\mathbf{H}''_m * \mathbf{W}_{2,m} + \mathbf{b}_{2,m}) \quad (10)$$

其中 $\mathbf{H}'_m \in \mathbb{R}^{N_1 \times N_1 \times K \times 4}$ ,  $\mathbf{W}_{1,m}$ 和 $\mathbf{b}_{1,m}$ 分别为第一个3D CNN 的权重和偏置. $\mathbf{H}''_m \in \mathbb{R}^{N_1 \times N_1 \times K \times 2}$ ,  $\mathbf{W}_{2,m}$ 和 $\mathbf{b}_{2,m}$ 分别为第二个3D CNN 的权重和偏置.

第二个2D CNN 和第二个3D CNN 网络的输出将分别通过2D Maxpooling 和3D Maxpooling 网络将数据压缩到原来的1/4. 2D Maxpooling 和3D Maxpooling 网络的池化窗口大小分别为 $2 \times 2$ 和 $2 \times 2 \times 1$ , 池化后的输出分别为 $\widehat{\mathbf{H}}_s \in \mathbb{R}^{\frac{N_1}{2} \times \frac{N_1}{2} \times 2}$ 和 $\widehat{\mathbf{H}}_m \in \mathbb{R}^{\frac{N_1}{2} \times \frac{N_1}{2} \times K \times 2}$ . 然后将它们分别重新排列成一维向量, 分别得到 $\widehat{\mathbf{H}}'_s \in \mathbb{R}^{\frac{N_1 N_1}{2}}$ 和 $\widehat{\mathbf{H}}'_m \in \mathbb{R}^{\frac{N_1 N_1 K}{2}}$ .

### (2) 解压处理

在解压部分,首先利用全连接网络将数据增大4倍,分别得到 $\widehat{\mathbf{H}}''_s \in \mathbb{R}^{2N_1 N_1}$ 和 $\widehat{\mathbf{H}}''_m \in \mathbb{R}^{2N_1 N_1 K}$ . 然后将它们分别重新排列成适合Bi-LSTM 和Bi-ConvLSTM 网络输入的形式,得到 $\widetilde{\mathbf{H}}_s \in \mathbb{R}^{N_1 \times 2N_1}$ 和 $\widetilde{\mathbf{H}}_m \in \mathbb{R}^{N_1 \times N_1 \times K \times 2}$ .

将重新排列的CSI 通过Bi-LSTM 或Bi-ConvLSTM 网络来对CSI 进行预测. 对于CSI 反馈,Bi-LSTM 和Bi-ConvLSTM 网络的每个时间步都有输出,令

$$\widetilde{\mathbf{H}}_s = [\widetilde{\mathbf{h}}_s^{(1)}, \dots, \widetilde{\mathbf{h}}_s^{(t)}, \dots, \widetilde{\mathbf{h}}_s^{(N_t)}]^H,$$

$$\widetilde{\mathbf{H}}_m = [\widetilde{\mathbf{h}}_m^{(1)}, \dots, \widetilde{\mathbf{h}}_m^{(t)}, \dots, \widetilde{\mathbf{h}}_m^{(N_t)}]^H,$$

其中 $\widetilde{\mathbf{h}}_s^{(t)} \in \mathbb{R}^{2N_1}$ ,  $\widetilde{\mathbf{h}}_m^{(t)} \in \mathbb{R}^{N_1 \times K \times 2}$ . 我们可以得到单用户和多用户每个时间步CSI 预测序列分别如下:

$$\widehat{\mathbf{h}}_s^{(t)} = \text{Bi-LSTM}(\mathbf{I}_{\text{sf}}^{(t-1)}, \mathbf{I}_{\text{sr}}^{(t-1)}, \widetilde{\mathbf{h}}_s^{(t)}, \widetilde{\mathbf{h}}_s^{(t)}, \Theta_s) \quad (11)$$

$$\widehat{\mathbf{h}}_m^{(t)} = \text{Bi-ConvLSTM}(\mathbf{I}_{\text{mf}}^{(t-1)}, \mathbf{I}_{\text{mr}}^{(t-1)}, \widetilde{\mathbf{h}}_m^{(t)}, \widetilde{\mathbf{h}}_m^{(t)}, \Theta_m) \quad (12)$$

其中, $\mathbf{I}_{\text{sf}}^{(t-1)}$ 和 $\mathbf{I}_{\text{sr}}^{(t-1)}$ 分别为Bi-LSTM 网络中第 $t-1$ 根天线的正向和反向隐藏层矢量, $\widetilde{\mathbf{h}}_s^{(t)}$ 和 $\widetilde{\mathbf{h}}_s^{(t)}$ 分别为Bi-LSTM 网络中第 $t$ 根天线的正向和反向输入, $\Theta_s$ 和 $\widetilde{\mathbf{h}}_s^{(t)}$ 分别为Bi-LSTM 网络中的所有参数和第 $t$ 根天线的输出. $\mathbf{I}_{\text{mf}}^{(t-1)}$ 和 $\mathbf{I}_{\text{mr}}^{(t-1)}$ 分别为Bi-ConvLSTM 网络中第 $t-1$ 根天线的正向和反向隐藏层矢量, $\widetilde{\mathbf{h}}_m^{(t)}$ 和 $\widetilde{\mathbf{h}}_m^{(t)}$ 分别为Bi-ConvLSTM 网络中第 $t$ 根天线的正向和反向输入, $\Theta_m$ 和 $\widetilde{\mathbf{h}}_m^{(t)}$ 分别为Bi-ConvLSTM 网络中的所有参数和第 $t$ 根天线的输出.

Bi-LSTM 和Bi-ConvLSTM 网络的输出维度为输入的两倍,即 $\widehat{\mathbf{h}}_s^{(t)} \in \mathbb{R}^{4N_1}$ ,  $\widehat{\mathbf{h}}_m^{(t)} \in \mathbb{R}^{N_1 \times K \times 4}$ . 最后通过全连接网络将Bi-LSTM 网络每个时序的输出进行降维,使得最终的输出维度与输入维度一致,其表达式为:

$$\bar{\mathbf{h}}_s^{(t)} = f(\mathbf{W}_{\text{sl}}^{(t)} \widehat{\mathbf{h}}_s^{(t)} + \mathbf{b}_{\text{sl}}^{(t)}) \quad (13)$$

$$\bar{\mathbf{h}}_m^{(t)} = f(\mathbf{W}_{\text{ml}}^{(t)} \widehat{\mathbf{h}}_m^{(t)} + \mathbf{b}_{\text{ml}}^{(t)}) \quad (14)$$

其中 $\mathbf{W}_{\text{sl}}^{(t)}$ 和 $\mathbf{b}_{\text{sl}}^{(t)}$ 分别为单用户全连接层第 $t$ 根天线的权重和偏置, $\mathbf{W}_{\text{ml}}^{(t)}$ 和 $\mathbf{b}_{\text{ml}}^{(t)}$ 分别为多用户全连接层第 $t$ 根天线全连接层的权重和偏置, $\bar{\mathbf{h}}_s^{(t)} \in \mathbb{R}^{2N_1}$ ,  $\bar{\mathbf{h}}_m^{(t)} \in \mathbb{R}^{N_1 \times K \times 2}$ . 将 $\bar{\mathbf{h}}_s^{(t)} \in \mathbb{R}^{2N_1}$ 重新排列为 $\bar{\mathbf{h}}_s'^{(t)} \in \mathbb{R}^{N_1 \times 2}$ 形式后,再分别将 $\bar{\mathbf{h}}_s'^{(t)} \in \mathbb{R}^{N_1 \times 2}$ 和 $\bar{\mathbf{h}}_m^{(t)} \in \mathbb{R}^{N_1 \times K \times 2}$ 对应的实部和虚部加在一起得到最终输出:

$$\widehat{\mathbf{H}}_s = [\widehat{\mathbf{h}}_s^{(1)}, \dots, \widehat{\mathbf{h}}_s^{(t)}, \dots, \widehat{\mathbf{h}}_s^{(N_t)}]^H \quad (15)$$

$$\widehat{\mathbf{H}}_m = [\widehat{\mathbf{h}}_m^{(1)}, \dots, \widehat{\mathbf{h}}_m^{(t)}, \dots, \widehat{\mathbf{h}}_m^{(N_t)}]^H \quad (16)$$

其中, $\widehat{\mathbf{H}}_s \in \mathbb{C}^{N_1 \times N_1}$ ,  $\widehat{\mathbf{H}}_m \in \mathbb{C}^{N_1 \times N_1 \times K}$ ,  $\widehat{\mathbf{h}}_s^{(t)} \in \mathbb{C}^{N_1}$ ,  $\widehat{\mathbf{h}}_m^{(t)} \in \mathbb{C}^{N_1 \times K}$ .

### 3.3 离线模型训练和在线反馈

本文所提的方法采用离线模型训练和在线反馈的方式进行信号处理,通过MATLAB 收集数据和Python 对所提方法进行离线训练. 本文所使用训练设备的配

置为 GeForce GTX 1060 显卡, Intel Xeon® E3-1231 V3 处理器. 信道模型我们采用的是大规模 MIMO 瑞利平坦衰落信道模型<sup>[4]</sup>. 对于离线训练过程, 本文将大规模 MIMO 信道的 CSI 作为学习网络的输入数据和标签数据对学习网络进行训练, 使用的训练集、校验集和测试集的样本数分别为 20000, 5000 和 1000. 对于在线预测, 可以不需要重复训练, 直接使用训练好的模型即可. 本文使用端到端的方式获取网络中的所有权重和偏置, 所提网络的初始学习率设为 0.01, 训练 epoch 为 100, batch 大小为 100, 使用自适应矩估计 (ADAM, Adaptive Moment estimation) 优化方法训练网络. 本文所提网络的损失函数为均方误差 (Mean Squared Error, MSE), 通过最小化网络输出和标签数据之间的 MSE 来更新网络中的所有参数. MSE 损失函数的计算方法如下:

$$L(\Theta_{\text{csi}}) = \frac{1}{N} \sum_{i=1}^N (f(\mathbf{H}_i; \Theta_{\text{csi}}) - \mathbf{H}_i)^2 \quad (17)$$

其中,  $\mathbf{H}_i$  为第  $i$  个样本的原始输入,  $\Theta_{\text{csi}}$  为网络中所有的参数,  $f$  为整个网络的变换函数,  $f(\mathbf{H}_i; \Theta_{\text{csi}})$  为整个网络恢复的第  $i$  个样本数据,  $N$  为总共的样本数.

## 4 仿真和分析

为验证 FDD 大规模 MIMO 系统下所提方法的性能, 在本节中, 比较了所提方法与其他典型 CSI 压缩反馈方法的 MATLAB 仿真性能. 假设用户侧已获得理想的下行 CSI, 仿真系统的主要参数如表 1 所示.

表 1 仿真参数设置

仿真参数	设置
双工模式	FDD
用户个数	1/2
发射天线	$N_t = 32/64$
接收天线	$N_r = 2$
天线间距	0.5 $\lambda$
调试方式	QPSK
信道模型	大规模 MIMO 瑞利平坦衰落信道
工作频率	2.6GHz
功率分配	等功率分配

### 4.1 归一化均方误差

本节比较了数据压缩率为 1/4、发射天线分别为 32 和 64 根、用户数分别为 1 和 2 个时所提方法和 DCT、PCA 以及 CsiNet 的归一化均方差  $\text{NMSE} = E(\|\mathbf{H} - \hat{\mathbf{H}}\|_2^2 / \|\mathbf{H}\|_2^2)$ . 其中 NMSE 越小, 表示信道压缩反馈误差越小, 性能越好. 该 NMSE 采用了取对数形式表示, 结果如表 2 所示.

从表 2 可以看出, 不论是单用户还是多用户, 不论

BS 侧天线数为 32 还是 64, 本文所提方法都表现出优异的 NMSE 性能, 该方法更能准确地恢复出 CSI, 重构的 CSI 质量明显提高. 这主要是因为所提网络不需要知道信道的分布情况, 能够通过训练样本来训练网络, 而且该网络充分利用了大规模 MIMO 信道的空间相关性, 能更好地学习到信道的结构特征. 值得一提的是, 本小节对比方法中没有引入 KLT, 是因为当 BS 和用户都知道瞬时相关矩阵时, KLT 基具有最佳的稀疏表示, 只需要很少的测量数据就能准确的恢复信息, 并且其误差几乎一直为 0. 但在 KLT 在 BS 不知道瞬时相关矩阵的情况下, KLT 不具有稀疏压缩效果.

表 2 各种 CSI 反馈方法的 NMSE (dB) 对比

用户个数	$N_t$	DCT	PCA	CsiNet	所提方法
1	32	-14.32	-14.85	-16.12	-17.58
	64	-15.01	-16.28	-18.53	-20.48
2	32	-14.17	-14.70	-15.99	-17.45
	64	-14.86	-16.13	-18.40	-20.35

### 4.2 运行时间

为了更直观地对比上述方法的运行速率, 本小节记录了在单用户和多用户以及不同发射天线配置条件下进行完整 CSI 反馈的运行时间. 从表 3 中可以看出, 本文所提方法在运行时间上具有明显优势, 基于深度学习方法的运行时间比传统方法短. 这是因为深度学习网络只需要简单的矩阵相乘就完成完整的 CSI 反馈, 且所提网络的参数比 CsiNet 更少, 所以本文提出的方法运行时间比 CsiNet 更短.

表 3 各种 CSI 反馈方法的运行时间对比 (单位: s)

用户个数	$N_t$	DCT	PCA	KLT	CsiNet	所提方法
1	32	0.0124	0.0092	0.0352	0.0018	0.0012
	64	0.0176	0.0143	0.2113	0.0021	0.0015
2	32	0.0211	0.0153	0.0676	0.0032	0.0020
	64	0.0313	0.0258	0.3931	0.0039	0.0027

### 4.3 误码率

在大规模 MIMO CSI 反馈中, 误码率 (Bit Error Ratio, BER) 性能是衡量 CSI 反馈方法对系统整体性能影响的宏观指标. 本文反馈压缩比都设定为 1/4, 在保证反馈压缩比相同的情况下, 分别仿真了单用户和多用户大规模 MIMO 系统的 BER 性能.

图 4 表示 BS 端分别配置 32 根、64 根天线, 接收端配置 2 根天线的单用户系统不同反馈方法 BER 性能随信噪比 (SNR, Signal to Noise Ratio) 变化的曲线. 由图 4 可知, 不论 BS 端是 32 根、64 根天线, 所提方法比 DCT、PCA、KLT 和 CsiNet 方法的 BER 性能都要好. 当 BS 天线为 32 时, 所提方法与传统的 DCT、PCA、KLT 方法相比 SNR 提升约 2~4dB 增益, 与 CsiNet 方法相比 SNR

提升约 0.5dB 增益. 当 BS 天线为 64 时, 所提方法与传统的 DCT、PCA、KLT 方法相比 SNR 提升约 2 ~ 6dB 增益, 与 CsiNet 方法相比 SNR 提升约 0.7dB 增益. 并且由

图 4 可知, 基于空间相关性的大规模 MIMO 系统 BS 端天线数越多 BER 性能越好.

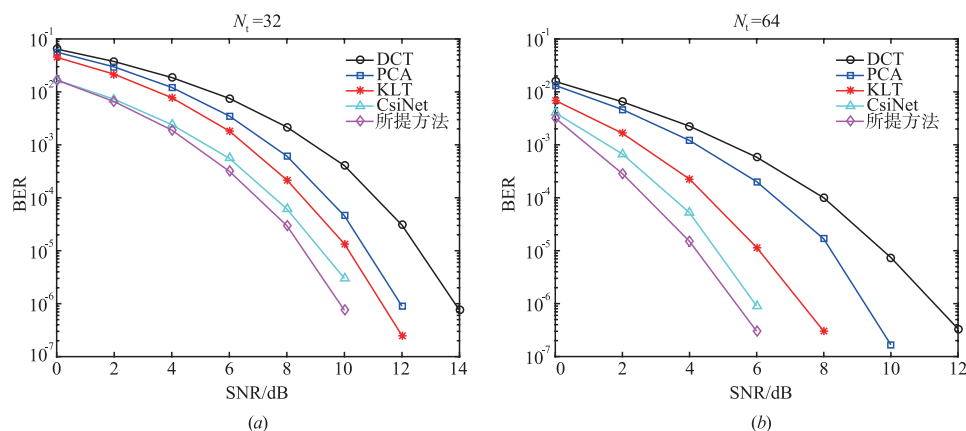


图4 单用户系统各种反馈方法BER比较

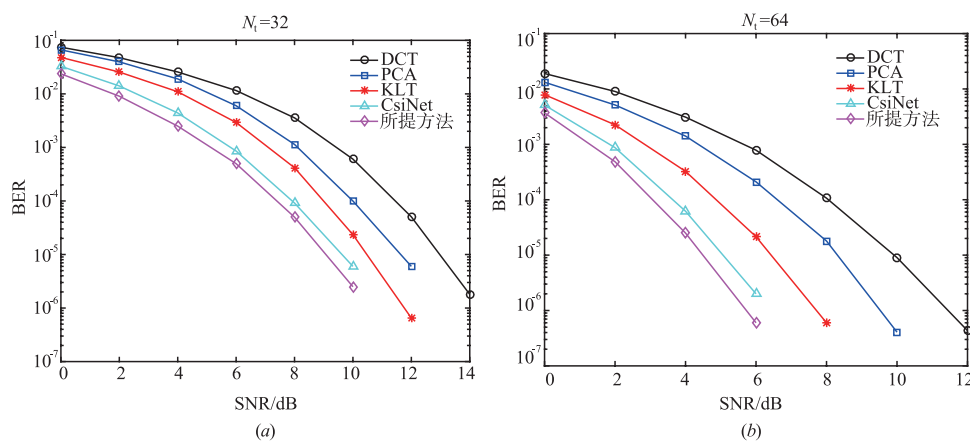


图5 多用户系统各种反馈方法BER比较

图 5 表示 BS 端分别配置 32 根、64 根天线, 接收端配置 2 根天线, 用户数为 2 的多用户系统不同反馈方法 BER 性能随 SNR 变化的曲线. 由图 5 不难得到, 不论 BS 端是 32 根、64 根天线, 多用户系统下的 BER 性能与单用户系统下的 BER 性能的趋势一致, 所提方法比其他的 CSI 反馈方法的 BER 性能都要好.

结合图 4 和图 5 可知, 不论 BS 端是 32 根、64 根天线, 不论是单用户还是多用户, 所提方法比 DCT、PCA、KLT 和 CsiNet 方法的 BER 性能都要好, 这是因为 DCT 需要下行信道的相关矩阵来生成压缩矩阵和恢复矩阵, 终端需要根据信道的变化情况周期性的向 BS 反馈该信息, 当信道变化较快时, BS 很难准确地获取信道相关信息, 从而造成反馈精度的下降. PCA 是通过提取主要成分进行压缩和恢复信道矩阵, 但其使用的压缩矩阵是通过前一长时阶段估计得到的, 因此 BS 难以得到准确的 CSI. KLT 方法当 BS 和移动端都知道瞬时相关矩阵时, KLT 基具有最佳的稀疏表示, 它只需要很少的

测量数据就能准确地恢复信道矩阵, 但是 KLT 基具有信号依赖性, 需要移动端反馈信道瞬时矩阵, 这反而增加了反馈量. 此外, CsiNet 方法只利用卷积层和全连接层对数据进行简单地特征提取来完成 CSI 的压缩和恢复, 并没有考虑大规模 MIMO 系统中天线的相关性, 而所提方法充分利用了大规模 MIMO 系统中天线的相关性来精确地恢复 CSI.

#### 4.4 系统容量

图 6 给出了 BS 端分别配置 32 根、64 根天线, 接收端配置 2 根天线的单用户系统不同反馈方法系统容量随 SNR 变化的曲线. 从图 6 可以看出, 随着 SNR 的增大所提方法获得的系统容量也随之增大, 且所提方法的系统容量都比其他方法高. 当 BS 端为 32 根天线时, 相比于传统方法, 所提方法系统容量提升在 1.6 ~ 2.8bps/Hz 左右, 相比于 CsiNet 方法, 系统容量提升在 0.3bps/Hz 左右. 当 BS 端为 64 根天线时, 所提方法系统容量提升在 2 ~ 3bps/Hz 左右, 相比于 CsiNet 方法, 系

统容量提升在 0.5bps/Hz 左右. 在 BS 配置不同天线下, 本文所提方法都保持着较高的系统性能增益.

图 7 给出了 BS 端分别配置 32 根、64 根天线, 接收端配置 2 根天线, 用户数为 2 的多用户系统不同反馈方法系统容量随 SNR 变化的曲线. 由图 7 可知, 对于多用户大规模 MIMO 系统, 在高 SNR 情况下, 本文所提方法的有明显的性能提升, 具有更优的信道容量. 在 SNR = 20dB 时, 当 BS 端为 32 根天线, 所提方法比传统方法的

系统容量提升在 1 ~ 1.8bps/Hz 左右, 比 CsiNet 方法系统容量提升 0.4bps/Hz 左右, 当 BS 端为 64 根天线时, 所提方法比传统方法的系统容量提升 1.2 ~ 2.0bps/Hz 左右, 比 CsiNet 方法系统容量提升 0.6bps/Hz 左右. 但是与单用户相比, 在相同天线数和 SNR 下, 多用户系统容量小于单用户系统容量, 这是因为多用户之间存在用户干扰, 所以多用户系统容量性能低于单用户系统.

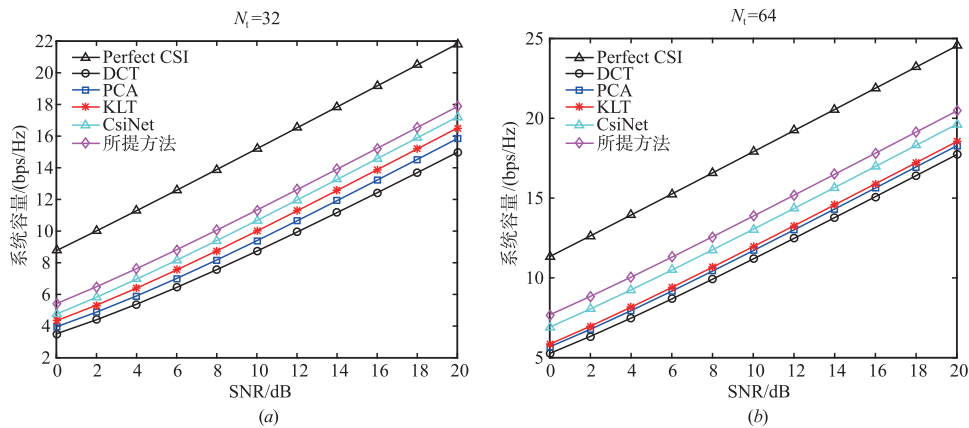


图6 单用户系统各种反馈方法系统容量比较

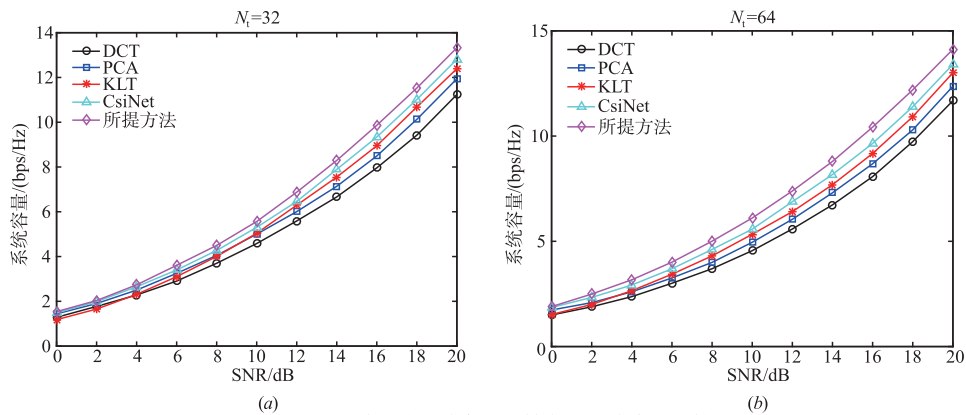


图7 多用户系统各种反馈方法系统容量比较

## 5 结论

本文研究了基于空间相关性的 FDD 大规模 MIMO 系统 CSI 压缩反馈方法. 考虑传统 CSI 压缩反馈方法复杂度高、反馈不准确以及 CsiNet 方法没有考虑大规模 MIMO 系统中天线的空间相关性的问题, 本文提出了一种新的基于深度学习的 CSI 压缩反馈方法, 并将其应用到单用户和多用户场景. 该方法考虑了大规模 MIMO 均匀线性天线阵列的空间相关性, 通过离线训练使模型充分学习信道的结构特征, 从而提升 CSI 压缩反馈的准确度. 通过理论分析和仿真结果可知, 本文所提的方法与传统 CSI 压缩反馈方法以及 CsiNet 方法相比, 压缩反馈的运行时间更短, 并且 BER 和系统容量性能更优.

## 参考文献

- [1] Larsson E G, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems [J]. IEEE Communications Magazine, 2014, 52(2): 186 - 195.
- [2] Barriac G, Madhow U. Space-time communication for OFDM with implicit channel feedback [J]. IEEE Transactions on Information Theory, 2004, 50(12): 3111 - 3129.
- [3] Tseng C C, Wu J Y, Lee T S. Enhanced compressive downlink CSI recovery for FDD massive MIMO systems using weighted block  $\ell_1$ -minimization [J]. IEEE Transactions on Communications, 2016, 64(3): 1055 - 1067.
- [4] Kuo P H, Kung H T, Ting P A. Compressive sensing based

- channel feedback protocols for spatially-correlated massive antenna arrays[A]. Wireless Communications and Networking Conference [C]. Shanghai, China: IEEE, 2012. 492 – 497.
- [5] Min S S, Chae C B. Compressed channel feedback for correlated massive MIMO systems[A]. Globecom Workshops [C]. Austin, TX, USA: IEEE, 2014. 327 – 332.
- [6] Ge A, Zhang T, Hu Z, et al. Principal component analysis based limited feedback scheme for massive MIMO systems [A]. Personal, Indoor, and Mobile Radio Communications [C]. Hong Kong: IEEE, 2015. 326 – 331.
- [7] Wang T Q, Wen C K, Wang H Q, et al. Deep learning for wireless physical layer: opportunities and challenges [J]. China Communications, 2017, 14(11): 92 – 111.
- [8] Liao Y, Hua Y X, Yao H M, Yang X Y. ChanEstNet: a deep learning based channel estimation for high-speed scenarios [A]. International Conference on Communications [C]. Shanghai, China: IEEE, 2019. 1 – 6.
- [9] Ye H, Li G Y, Juang B H. Power of deep learning for channel estimation and signal detection in OFDM systems [J]. IEEE Wireless Communications Letters, 2018, 7(1): 114 – 117.
- [10] Xu W, Zhong Z, Bei Y, et al. Joint neural network equalizer and decoder [A]. International Symposium on Wireless Communication Systems [C]. Lisbon, Portugal: IEEE, 2018. 1 – 5.
- [11] Wen C K, Shih W T, Jin S. Deep learning for massive MIMO CSI feedback [J]. IEEE Wireless Communications Letters, 2017, 7(5): 748 – 751.
- [12] Iqbal R, Abhayapala T D, Lamaheva T A. Generalised clarke model for mobile-radio reception [J]. IET Communications, 2009, 3(4): 644 – 654.

#### 作者简介



**廖 勇 (通信作者)** 男, 1982 年出生于四川省自贡市. 现为重庆大学副研究员, 博士生导师, 主要研究方向为下一代无线通信与网络, 人工智能、区块链、量子计算及其在无线通信中的应用等.

E-mail: liaoy@cqu.edu.cn



**姚海梅** 女, 1992 年生于江西省吉安市. 现为重庆大学微电子与通信工程学院研究生, 主要研究方向为人工智能算法及其在无线通信中的应用.

E-mail: meimei@cqu.edu.cn



**花远肖** 男, 1994 年生于四川省阆中市. 现为重庆大学微电子与通信工程学院研究生, 主要研究方向为人工智能算法及其在无线通信中的应用.

E-mail: huayx@cqu.edu.cn



**赵 砚** 男, 1982 年出生于山东省诸城市. 现为中国人民解放军 61212 部队助理研究员, 主要研究方向为通信信号处理与系统总体设计.

E-mail: chym\_zhaoyan@163.com