

FFDNet-Based Channel Estimation for Massive MIMO Visible Light Communication Systems

Zhipeng Gao, Yuhao Wang, *Senior Member, IEEE*, Xiaodong Liu, Fuhui Zhou, *Member, IEEE*,
Kat-Kit Wong, *Fellow, IEEE*

Abstract—Channel estimation is of crucial importance in massive multiple-input multiple-output (m-MIMO) visible light communication (VLC) systems. In order to tackle this problem, a fast and flexible denoising convolutional neural network (FFDNet)-based channel estimation scheme for m-MIMO VLC systems was proposed. The channel matrix of the m-MIMO VLC channel is identified as a two-dimensional natural image since the channel has the characteristic of sparsity. A deep learning-enabled image denoising network FFDNet is exploited to learn from a large number of training data and to estimate the m-MIMO VLC channel. Simulation results demonstrate that our proposed channel estimation based on the FFDNet significantly outperforms the benchmark scheme based on minimum mean square error.

Index Terms—Channel estimation, m-MIMO, visible light communication, FFDNet, deep learning.

I. INTRODUCTION

WITH the ever increasing mobile data traffic, massive multiple-input multiple-output (m-MIMO) is promising to enhance the communication capacity and the spectrum efficiency of the existing visible light communication (VLC) systems [1]-[3]. However, such benefits may be limited due to the inaccuracy of channel estimation [4]. Thus, an accurate channel estimation is of crucial importance in the m-MIMO VLC systems. However, when the dimensions of the channel matrix become large, it is extremely challenging to design efficient channel estimation techniques and achieve an accurate channel estimation for the m-MIMO systems [5]-[6]. Moreover, the conventional channel estimation techniques for m-MIMO systems such as minimum mean square error (MMSE) are typically dependent on the statistical properties of the channel and specific prior information [7]-[8]. In practice, it is difficult to obtain those statistical properties, especially when the number of transmitters and that of receivers are large. Therefore, the performance of those conventional channel estimation methods are not always optimal.

In order to solve these problems, deep learning (DL) has attracted an increasing attention and has great potential to be applied for channel estimation of the wireless communication systems [9]-[11]. Specifically, in [9], a DL-based signal detector without requiring prior knowledge about the channel state

information or background noise was proposed for cooperative detection. And in [10], a DL-based scheme was proposed to achieve the super-resolution channel estimation and direction-of-arrival estimation in m-MIMO systems. Since the line-of-sight (LOS) link communication is dominant in VLC systems, the m-MIMO VLC channel has the characteristic of sparsity. This is similar to that in [11] where the mmWave m-MIMO channel matrix was regarded as a two-dimensional (2D) noise-free natural image by exploiting the sparsity of the channel. Then a convolutional neural network (CNN)-based image denoising network was exploited to estimate the m-MIMO mmWave channel. Motivated by those facts, the m-MIMO VLC channel matrix can be also identified as a 2D noise-free natural image.

Based on the above-mentioned analyses, it is feasible to utilize the theory of image denoising in DL for estimating the m-MIMO VLC channel. However, the state-of-the-art CNN-based image denoising methods are still limited in flexibility and efficiency [12]. To overcome the drawbacks of the existing CNN-based image denoising methods, a image denoising network called fast and flexible denoising convolutional neural network (FFDNet) was proposed in [12]. FFDNet has the superiority in terms of both denoising performance and computation efficiency. Moreover, FFDNet performs effectively and flexibly on the denoising of images that are corrupted by additive white gaussian noise (AWGN). Therefore, it is promising to apply FFDNet into the m-MIMO VLC systems for realizing the channel estimation.

Motivated by the above-mentioned facts, a channel estimation scheme is proposed based on FFDNet in a m-MIMO VLC system. The main contributions of our work are summarized as follows. A m-MIMO VLC system using a transceiver array which is composed of massive LEDs and PDs is considered. Then, the concept of image denoising is exploited and a trained FFDNet network is used to estimate the m-MIMO VLC channel. Moreover, the performance of channel estimation using FFDNet is evaluated and compared with that obtained by the traditional MMSE channel estimation scheme. Simulation results demonstrate that the performance of FFDNet with a fixed input noise level is better than that of MMSE at high real noise levels. Furthermore, the performance gain can be further improved by FFDNet with a tunable input noise level.

The rest of this paper is organized as follows. Section II comprehensively presents the m-MIMO VLC systems and the architecture of FFDNet image denoising network. In Section III, simulation results are presented to evaluate the scheme. Finally, the paper concludes with Section IV.

Z. Gao, Y. Wang and F. Zhou are with Nanchang University, China (e-mail: gdpeungee@email.ncu.edu.cn, wangyuhao@ncu.edu.cn, zhoufuhui@iee.org).

X. Liu is with Wuhan University, China (e-mail: xiaodongliu@whu.edu.cn). K. K. Wong is with University College London, UK (e-mail: kai-kit.wong@ucl.ac.uk).

This work was supported in part by the National Natural Science Foundation of China (61661028 and 61701214).

II. CHANNEL AND FFDNET NETWORK

A. System Scenario and Channel Model

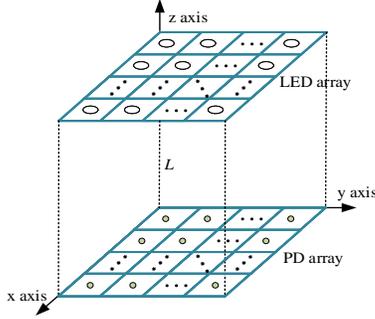


Fig. 1: The m-MIMO VLC system model.

In this paper, a m-MIMO VLC system is considered with N_t light emitting diodes (LEDs) as the transmitters and N_r photo detectors (PDs) as the receivers. It is assumed that the specifications and performance of all LEDs are identical, as are the PDs. As shown in Fig. 1, the LED-array plane is parallel to the PD-array plane, and all LEDs and PDs are arranged with an equal interval [13]. And the vertical distance between the LED-array and the PD-array is L .

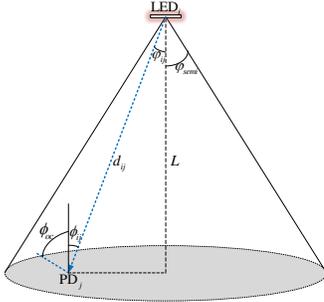


Fig. 2: The VLC channel model.

The current settings of the m-MIMO VLC channel is modeled as the LOS as shown in Fig. 2, since multipath delays resulting from reflections and diffuse refractions are typically negligible in the m-MIMO VLC scenario [14]. The channel gain h_{ij} between the i th LED and the j th PD is given by [15]

$$h_{ij} = \frac{A_r}{d_{ij}^2} R_o(\varphi_{ij}) T_s(\phi_{ij}) g(\phi_{ij}) \cos(\phi_{ij}), \quad (1)$$

where $i = 1, 2, 3, \dots, N_t$ and $j = 1, 2, 3, \dots, N_r$. A_r represents the effective collection area of the receiver PD. d_{ij} denotes the distance between the i th transmitting LED and the j th receiving PD. ϕ_{ij} and φ_{ij} represent the incidence angle and the irradiance angle, respectively. $T_s(\phi_{ij})$ is the gain of the optical filter, and $g(\phi_{ij})$ denotes the gain of the optical concentrator (OC), which can be calculated by

$$g(\phi_{ij}) = \begin{cases} \frac{n^2}{\sin^2(\phi_{oc})}, & \text{if } 0 \leq \phi_{ij} \leq \phi_{oc}, \\ 0, & \text{otherwise } \phi_{ij} > \phi_{oc}, \end{cases} \quad (2)$$

where ϕ_{oc} is the field of view (FOV) of the PD. n denotes the refractive index of OC. Moreover, $R_o(\varphi_{ij})$ represents the Lambertian radiant intensity of the transmitting LEDs, which can be expressed as

$$R_o(\varphi_{ij}) = \frac{m+1}{2\pi} \cos^m(\varphi_{ij}), \quad (3)$$

where m is the order of Lambertian emission, which is given by

$$m = -\frac{\ln(2)}{\ln(\cos(\varphi_{semi}))}, \quad (4)$$

where φ_{semi} represents the LED transmitter semi-angle at a half power level.

Motivated by the work in [11], after obtaining the channel matrix \mathbf{H} of the m-MIMO VLC system, the channel matrix can be treated as a noise-free channel image x with a size $w \times l$. And it becomes a noisy channel image y of size $w \times l$ after it is corrupted by AWGN with real noise level σ_o . Then the noisy channel image y is input to a trained FFDNet network to estimate the channel.

B. The FFDNet Architecture

As shown in Fig. 3 at the top of the next page, the first layer of FFDNet is a reversible downsampling process which reshapes the noisy channel image y into four downsampled sub-images with size $\frac{w}{2} \times \frac{l}{2} \times 4$. The downsampling process can significantly improve the training speed without reducing the modeling ability. Moreover, different from the denoising convolutional neural network (DnCNN), the denoising on downsampled sub-images can effectively increase the receptive field without employing the dilated convolution and lead to a moderate network depth [12]. After the operation of downsampling, a tunable noise level map \mathbf{M} with the input noise level¹ σ is combined with the downsampled sub-images to establish a tensor \tilde{y} with a size $\frac{w}{2} \times \frac{l}{2} \times (4+1)$ as the input of the CNN.

With the tensor \tilde{y} as input, the following layer is a CNN with the depth D_e . As shown in Fig. 3, in order to effectively speed up the training process and improve the denoising performance of FFDNet, each layer in the CNN is composed of a specific combination of three types of operations: convolution (Conv), rectified linear units (ReLU) and batch normalization (BN) [12]. The first convolution layer ‘‘Conv+ReLU’’ uses 64 filters with a size 3×3 to generate 64 feature maps, and ReLU is adopted for nonlinearity. The middle layers ‘‘Conv+BN+ReLU’’ adopt 64 filters with a size $3 \times 3 \times 64$ and BN is added between the operations of convolution and the ReLU. And the last layer ‘‘Conv’’ uses a filter with the size $3 \times 3 \times 64$ to reconstruct the output. Furthermore, after each convolution, zero-padding is employed to guarantee that the size of the feature maps is not changed. After the CNN, an upscaling operation is exploited as the reverse process of the downsampling process applied in the input stage to produce the denoised channel image \tilde{x} of size $w \times l$.

¹ σ is the noise level of the noise level map \mathbf{M} which is used to control the trade-off between noise reduction and detail preservation. σ_o is the noise level of the AWGN that added to the channel matrix of the m-MIMO VLC system.

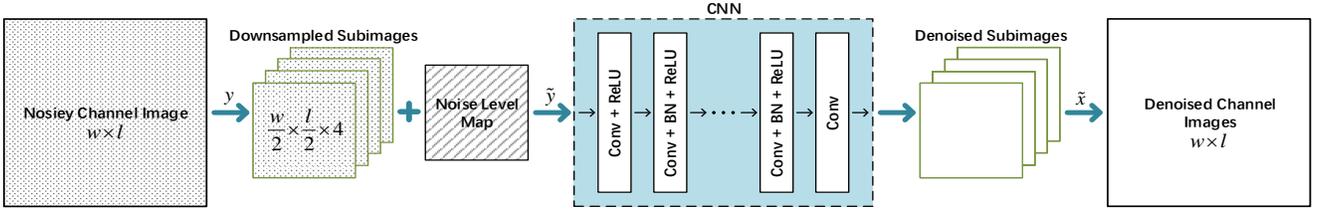


Fig. 3: The system architecture of FFDNet.

C. The Noise Level Map

Most of the model-based denoising methods aim to solve the problem, given as [12]

$$\tilde{x} = \arg \min_x \frac{1}{2\sigma^2} \|y - x\|^2 + \lambda \Phi(x), \quad (5)$$

where $\frac{1}{2\sigma^2} \|y - x\|^2$ is the data fidelity term with the input noise level σ , and $\Phi(x)$ represents the regularization term associated with image prior. λ is used to control the balance between the data fidelity term and the regularization term. In order to solve this problem, an implicit function can be defined as [12]

$$\tilde{x} = F(y, \sigma, \lambda; \Theta), \quad (6)$$

where Θ is the trainable parameters in CNN. Since λ can be equal to σ , eq. (6) can be rewritten as

$$\tilde{x} = F(y, \sigma; \Theta). \quad (7)$$

In this case, the input noise level σ can also control the balance between the data fidelity term and the regularization term. And these model-based methods are flexible in image denoising by setting various input noise levels. Thus it is promising to exploit CNN to learn the explicit mapping of eq. (7) which takes y and σ as inputs. However, y and σ cannot be fed into CNN directly since the dimensions of them are different. Therefore, in the FFDNet, a tunable noise level map \mathbf{M} which has the same dimension with y is utilized. And all the elements of \mathbf{M} are σ . Then the implicit function can be further written as

$$\tilde{x} = F(y, \mathbf{M}; \Theta). \quad (8)$$

D. Dataset Generation and Network Training

In order to train the FFDNet network, it is necessary to obtain N_p training data $\{(y_k, \mathbf{M}_k; \mathbf{H}_k)\}_{k=1}^{N_p}$. Firstly, we obtain the channel matrix \mathbf{H}_k by implementing the m-MIMO VLC channel and treat it as a noise-free channel image. Secondly, the noise-free channel image \mathbf{H}_k is added with AWGN that has zero mean and variance $(\sigma_o/255)^2$ to form the noisy channel image y_k . \mathbf{M}_k is the noise level map. In order to keep the balance between complexity and performance, similar to the parameter settings in [12], the depth of the network is set as 15 and it has a receptive field of 62×62 . The patch size is set to 70×70 since it should be larger than the receptive field of FFDNet. Then the noisy patches are obtained by adding AWGN to the noise-free patches. The noisy patches and \mathbf{M}_k are input into the CNN. Finally, during the network

training, the adaptive moment estimation algorithm is adopted to optimize the FFDNet by minimizing the loss function given as

$$L(\Theta) = \frac{1}{2N_p} \sum_{k=1}^{N_p} \|F(y_k, \mathbf{M}_k; \Theta) - x_k\|^2. \quad (9)$$

III. SIMULATION RESULTS

In this section, in order to effectively reflect the accuracy of channel estimation by FFDNet, the simulation results are given to evaluate the peak signal to noise ratio (PSNR) performance of channel estimation achieved with FFDNet in the m-MIMO VLC systems. Moreover, the achieved performance is compared with those obtained with the traditional channel estimation scheme based on MMSE.

Table 1: BASIC SIMULATION PARAMETERS

Description	Notation	Value
The size of the room	$l_r \times w_r \times h_r$	$8\text{m} \times 8\text{m} \times 4\text{m}$
Transmitter semi-angle	φ_{semi}	50°
LED power	P	0.02 W
FOV of the PD	ϕ_{oc}	45°
Physical area of PD	A_r	1 cm^2
Gain of optical filter	$T_s(\phi_{ij})$	1
Refractive index of OC	n	1.5

Based on the basic parameters of the m-MIMO VLC system shown in Table 1 [15], a number of channel images are obtained by implementing the m-MIMO VLC systems. Some of these images are used as the training data of the FFDNet network and the remaining images are exploited as the testing data. After the training and testing of the network, the performance of FFDNet for the channel estimation of the m-MIMO VLC system is presented and evaluated. Without loss the generality, the size of the transceiver array $N_t \times N_r$ is set to 128×128 or 256×256 . And Fig. 4 is given to illustrate the noise level sensitivity curves of FFDNet under both transceiver array sizes.

As shown in Fig. 4, the PSNR is demonstrated by varying different real noise levels $\sigma_o \in [0, 50]$ for the specific noise level maps with $\sigma = 5, 15, 25, 50$. It can be seen that the PSNR value begins to decrease obviously when the real noise level σ_o is larger than the input noise level σ . And the PSNR is decreased with the increase of σ_o . Therefore, in practical applications when FFDNet is exploited to estimate the m-MIMO VLC channel, it is better to choose an appropriate σ which is larger than the current σ_o to achieve a satisfactory noise reduction. Moreover, it can also be seen that the PSNR

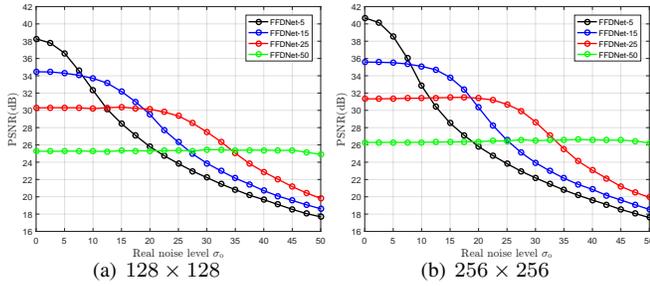


Fig. 4: The PSNR performance of channel estimation achieved by FFDNet with different sizes of the transceiver array.

values for the specific σ in Fig. 4 (b) are larger than those for the same specific σ in Fig. 4 (a). This is because the number of LEDs and PDs in the transceiver array of the m-MIMO VLC system increases and the array becomes more dense. In this case, the channel matrix has a more obvious sparsity, and its feature information is closer to the 2D natural image. Thus, the channel estimation performance with the FFDNet network based on the image denoising is better.

In order to evaluate the performance of channel estimation with FFDNet in the m-MIMO VLC system, based on the same settings as in Fig. 4 (b), Figs. 5 and 6 are given to compare the performance achieved by FFDNet with those obtained by the traditional channel estimation scheme based on MMSE.

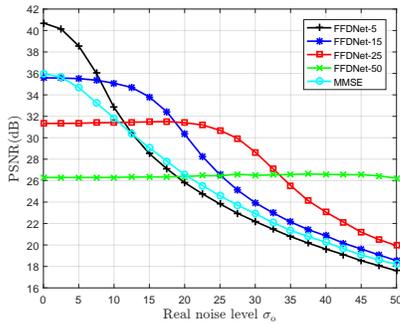


Fig. 5: The PSNR performance of channel estimation achieved by FFDNet with a fixed σ and that obtained by MMSE.

As shown in Fig. 5, the PSNR performance of the curve named FFDNet-15 is basically better than that of MMSE under different σ_0 . Although the PSNR values of MMSE are larger than those obtained by FFDNet in the case of the low σ_0 region, the performance of FFDNet is still better than that of MMSE at high σ_0 . Moreover, the biggest advantage of FFDNet is to exploit a variable σ to cope with the interference of different σ_0 . Therefore, we set an appropriate σ in FFDNet for the current σ_0 and compare with the results of MMSE. As shown in Fig. 6, a more satisfactory denoising effect than that of MMSE can be achieved by FFDNet. And it can be seen that the advantages of FFDNet become more obvious as the σ_0 increases.

IV. CONCLUSION

A FFDNet-based channel estimation scheme was proposed to estimate the channel of a m-MIMO VLC system. In order to

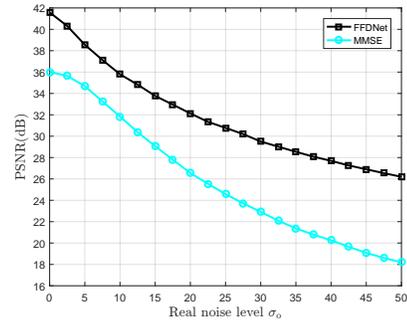


Fig. 6: The PSNR performance of channel estimation achieved by FFDNet with a tunable σ and that obtained by MMSE.

realize it, the channel matrix was established as a 2D natural image. The training method for the FFDNet was provided. Simulation results have shown the superiority of our proposed FFDNet-based channel estimation scheme compared with the existing MMSE scheme.

REFERENCES

- [1] R. Wang, Q. Gao, J. You, E. Liu, P. Wang, Z. Xu, and Y. Hua, "Linear transceiver designs for MIMO indoor visible light communications under lighting constraints," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2494-2508, Jun. 2017.
- [2] T. Fath and H. Haas, "Performance comparison of MIMO techniques for optical wireless communications in indoor environments," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 733-742, Feb. 2013.
- [3] J. Zhang, X. Xue, E. Bjrnson, B. Ai, and S. Jin, "Spectral efficiency of multipair massive MIMO two-way relaying with hardware impairments," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 14-17, Feb. 2018.
- [4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.
- [5] J. Yang, C. K. Wen, S. Jin, and F. Gao, "Beamspace channel estimation in mmWave systems via cosparsity reconstruction technique," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4767-4782, Feb. 2018.
- [6] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmWave MIMO Systems," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 245-259, Jan. 2019.
- [7] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065-3073, Dec. 2001.
- [8] X. Chen and M. Jiang, "Adaptive statistical bayesian MMSE channel estimation for visible light communication," *IEEE Trans. Signal Process.*, vol. 65, no. 5, pp. 1287-1299, Mar. 2017.
- [9] J. Gao, X. Yi, C. Zhong, X. Chen, and Z. Zhang, "Deep learning for Spectrum Sensing," *IEEE Wireless Commun. Lett.*, DOI:10.1109/LWC.2019.2939314.
- [10] H. Huang, J. Yang, H. Huang, Y. Song, and Guan Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549-8560, Apr. 2019.
- [11] H. He, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 852-855, Oct. 2018.
- [12] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608-4622, Sep. 2018.
- [13] K. Xu, H. Yu, and Y. J. Zhu, "Channel-adapted spatial modulation for massive MIMO visible light communications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 23, pp. 2693-2696, Dec. 2016.
- [14] H. Marshoud, P. C. Sofotasios, S. Muhaidat, G. K. Karagiannidis, and B. S. Sharif, "On the performance of visible light communication systems with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6350-6364, Oct. 2017.
- [15] L. Yin, W. O. Popoola, X. Wu, and H. Haas, "Performance evaluation of non-orthogonal multiple access in visible light communication," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5162-5175, Dec. 2016.