

A Survey on RIS-Assisted Millimeter-Wave Channel Estimation

Chengji Li

*School of Communication and Information Engineering, Shanghai University, Shanghai, China
lichengji123@shu.edu.cn*

Abstract. Nowadays, with the rapid development of Reconfigurable Intelligent Surface (RIS), it has gradually become one of the most important research areas in 6G communication. By dynamically adjusting the phase shift of the reflection units, RIS can smartly arrange the wireless communication environment, so as to effectively solve the problems in millimeter-wave communication such as blockage by obstacles and the loss of the communication path. However, because of the passive nature of RIS and the large-scale antenna dimensions, channel estimation has become an important task in order to reduce the pilot expenses. Considering this, this article provides an overview of RIS-assisted channel estimation methods using in millimeter-wave systems. This article divides the estimation methods into three main methods: compressed sensing, deep learning, and beam training. Compressive sensing uses the sparse characteristic of millimeter-wave channels in the angular domain, thus turning the channel estimation problem into recovering the sparse channel. Deep learning uses the neural networks' fitting ability to learn and simulate the mapping relationship of the channel matrix from the received signals. Beam training avoids estimating the complete channel matrix. Instead, it directly tests preset phase codebooks to find the best beam direction. Then, this article provides a detailed comparison of these three methods in terms of pilot expenses, computational complexity, and hardware limitations. Finally, this article discusses the open challenges like the beam splitting in broadband millimeter-wave systems and the joint channel estimation problem in STAR-RIS systems, giving some direction for future study.

Keywords: RIS, millimeter wave, phase shift, channel estimation, pilot expenses

1. Introduction

Reconfigurable Intelligent Surface (RIS), acting as a key technology in 5G-Advanced and 6G wireless communications in the future, has been attracting a widespread attention in both academic and industry fields. RIS is made up of a large amount of cheap, passive scattering units, enabling it to proactively arrange the wireless communication environment by controlling the phase shift of electromagnetic waves. This capability is especially useful for millimeter-wave communication. Although the millimeter-wave communication provides abundant spectrum bandwidth, a fatal shortcoming is its severe path loss and the communication is susceptible to be blocked by obstacles like buildings and human bodies. But by properly deploying RIS, people can create virtual line-of-

sight links, which can effectively turn areas that were originally blind spots into reliable coverage areas.

Despite its great potential, the actual deployment of RIS still faces a fundamental challenge: channel estimation. Traditional RIS units are completely passive and lack the RF chain, which means that they can't actively send or receive pilot signals. So, the base station (BS) can't separate the individual channels and has to estimate the high-dimensional cascaded channel. The pilot expenses of the traditional least squares (LS) estimator increase linearly with the number of RIS units, which quickly becomes unbearable in large-scale deployments such as deploying hundreds or thousands of RIS units.

To overcome this challenge, the academic field has been doing a lot of comprehensive research over the past few years. Stamatia F. Drampalou and others analyzed and overviewed the different methods facilitated with RIS in MIMO and MISO systems [1]. Beixiong Zheng and others overviewed the different methods considering the limitations of various architecture, different environments, and non-ideal hardware conditions in RIS systems [2]. Yun Yu and others mainly analyzed and overviewed various algorithms of LS estimation, compressed sensing, and deep learning [3]. Qing Xue and others overviewed the cutting-edge beam training algorithms [4]. Unlike existing general reviews, this article chooses a narrow point, focusing on the millimeter-wave communication system. Simultaneously, unlike previous studies, besides compressed sensing and deep learning, this article also summarizes beam training as an implicit channel estimation method, making the analysis of RIS-assisted millimeter-wave communication system more comprehensive. This article divides communication methods in the RIS-assisted millimeter-wave system into three categories: compressed sensing, deep learning, and beam training. This paper aim to overview the cutting-edge channel estimation methods in RIS-assisted millimeter-wave systems, introduce the basic principles and typical algorithms of these three methods, and provide ideas for learning and improving these methods in the future.

2. Classification and comparison of RIS-assisted millimeter-wave channel estimation methods

2.1. Compressed sensing

Compressed sensing mainly uses the spatial sparsity of millimeter-wave channels in the angular domain to turn the channel estimation problem into sparse signal recovery.

Zhen Chen and others proposed an adaptive sparse channel estimation method based on a multi-objective algorithm, named Proposed Hybrid MOEA. Different from several existing single-objective compressed sensing methods, the Proposed Hybrid MOEA algorithm reinterprets the channel estimation problem as a dual-objective optimization problem which balances both the channel reconstruction accuracy and the sparsity constraints. This algorithm doesn't need to get statistical information like the standard deviation of system noise in advance, which lets it completely avoid the theoretical challenge in traditional compressed sensing algorithms where the regularization parameter λ is hard to adjust automatically. When choosing the solution, Proposed Hybrid MOEA gets an eclectic solution with the maximum slope change. Simulation results show that under the same signal-to-noise ratio conditions, the normalized mean square error of the Proposed Hybrid MOEA algorithm is significantly lower than that of the traditional OMP algorithm. Simultaneously, with the same accuracy and equal normalized mean square error, the pilot sequence P required by the Proposed Hybrid MOEA algorithm is much smaller than that of the traditional OMP algorithm, taking a great advantage in pilot expenses [5].

The Proposed Hybrid MOEA algorithm shows that by combining multi-objective algorithms with compressive sensing, it can completely get rid of the dependence on signal-to-noise ratio, which means a good adaptability. At the same time, it also shows that compared to traditional least squares estimation, compressed sensing algorithms can significantly reduce pilot expenses through the Kronecker operation.

2.2. Deep learning

This method's principle is to treat channel estimation as an image reconstruction task. Figure 1 shows how the CNN network works during the process [6].

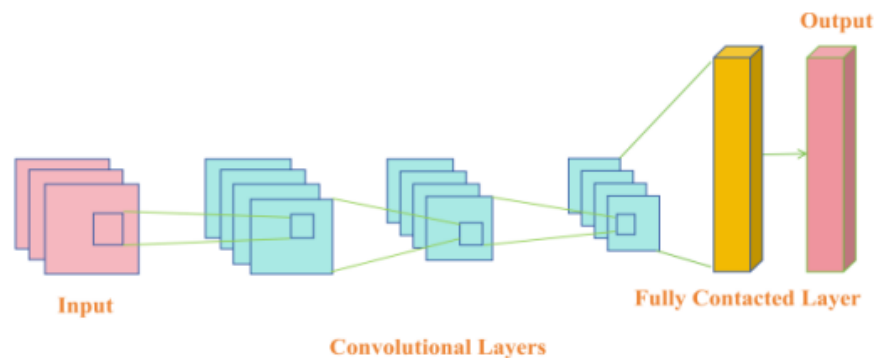


Figure 1. CNN working diagram (picture credit: original)

Thabang C. Rapudu and others proposed a DnCNN-GRU algorithm. The DnCNN-GRU algorithm divides the rows and columns of the channel matrix into respectively large and small time scales. On the large time scale, DnCNN noise-removing neural network is used to estimate the quasi-static channel from the first-level RIS to the base station, while on the small time scale, a GRU gated cycling unit is used to estimate the time-varying channel from the user to the first-level RIS, adaptively getting the correlation of the time-varying signals. In the normalized mean squared error test, the DnCNN-GRU algorithm performs significantly better than the common CNN neural network. At the same time, when facing different signal-to-noise ratio conditions, the normalized mean square error of the DnCNN-GRU algorithm is only about 1dB away from the theoretical lower bound. Additionally, because there are two time scales, the cascaded RIS channel doesn't need to be estimated repeatedly every time, which saves a lot of pilot expenses [7].

The DnCNN-GRU algorithm shows that with the introduction of multiple RIS units, deep learning networks like CNN has a strong dimension-reducing capability and can handle the estimation of high-dimensional channels with multiple RIS units. At the same time, deep learning algorithms have a good anti-interference ability to low signal-to-noise ratio conditions and can maintain stable reconstruction performance.

Regarding the discussion of the two kinds of algorithms above, compared to traditional compressed sensing and deep learning algorithms, Asmaa Abdallah and others proposed an improved DD-FS-CE algorithm that combines compressed sensing with deep learning. Based on the principles of compressed sensing, it changes the traditional model-driven approach to a data-driven one. It takes advantage of DnCNN's powerful feature extraction ability to predict all the non-zero columns at once, which effectively avoids error accumulation. And the results of this algorithm are also quite impressive. The text shows that in terms of computational complexity, the DD-FS-CE

algorithm reduces computational complexity by more than two orders of magnitude compared to the traditional DS-OMP algorithm. At the same time, the normalized mean square error of the DD-FS-CE algorithm is only about 0.5~1dB away from the theoretical optimal lower bound. In comparison, the traditional DS-OMP algorithm is off by as much as 5~6dB from the theoretical optimal lower bound. This also means that the DD-FS-CE algorithm has increased accuracy by nearly an order of magnitude. At the same time, when the number of users in the network drops from 16 to 4, the performance of the DD-FS-CE algorithm decreases much less than that of the DS-OMP algorithm, proving that it has better anti-interference ability. And during the online deployment, because the neural network skips complicated iterative calculations, the running complexity of DD-FS-CE algorithm is less than two orders of magnitude compared to the traditional DS-OMP algorithm [8].

By combining the traditional DS-OMP algorithm with the improved DD-FS-CE algorithm that uses DnCNN, this article summarizes the advantages of compressed sensing methods and deep learning methods respectively. First of all, compressed sensing algorithms have a natural theoretical advantage, allowing all the subcarrier channels to be recovered from observations of only a few subcarriers. Then, since the theory of compressed sensing algorithms is well-developed, the framework of compressed sensing algorithms has very high estimation accuracy. And deep learning methods can reduce the computational complexity of online deployment, with a good anti-interference ability to changes in the number of users.

2.3. Implicit estimation of beam training

To get rid of the complicated matrix solving in explicit channel estimation, the beam training method pre-designs some phase shift beams, named codebook, which represent several reflection directions. The base station and RIS take turns to scan these beams, so as to directly align with the physical space paths. As shown in Figure 2, both the base station and the user's devices have pre-designed different phase beams, which are expressed in the form of two-dimensional matrices [4, 9]. During the beam training process, researchers take the squared inner product of different main beams with the real channel's matrix, and pick the one with the highest squared inner product, which is the closest to the actual channel [10].

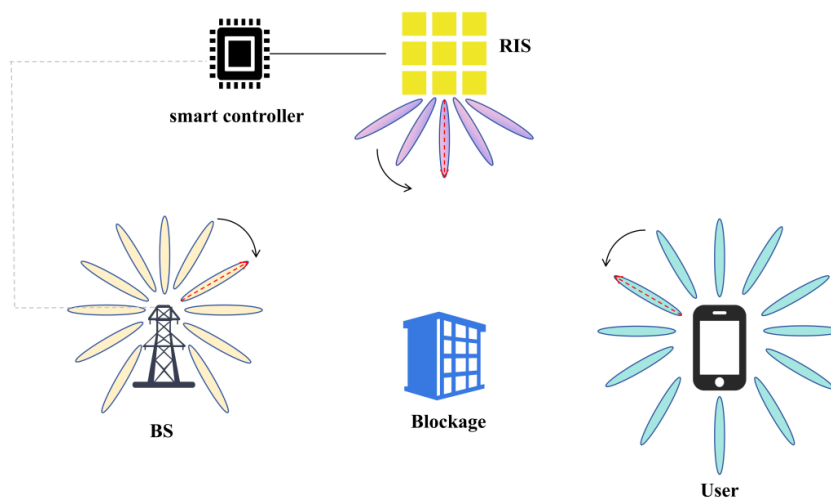


Figure 2. Beam training diagram (picture credit: original)

Jinghe Wang and others proposed a hierarchical codebook design and beam training algorithm based on passive RIS constraints [11]. In traditional systems, the main beam is usually widened by setting some of the RIS amplitudes to zero. However, because of hardware limitations, RIS units cannot be turned off while adjusting the phase. So, Jinghe Wang and others proposed a hierarchical multi-beam codebook and a high-fault-tolerance beam training algorithm which is suitable for passive RIS system. By setting up two complementary multi-main beams at each layer, this algorithm can achieve full-angle coverage without turning off any RIS units. At the same time, the algorithm also sets up a multi-layer joint decision agreement. The algorithm first scans each layer and finally makes a joint decision at the bottom layer, which effectively solves the decision problems caused by misjudgments in wide beam layers, having a strong self-correction ability. Simulation results show that in a 128-unit passive RIS system, the proposed algorithm has over 85% beam alignment success rate while only consuming about 17.2% of the exhaustive pilot expenses, effectively solving the core contradiction between hardware constraints and beam accuracy in passive RIS beam alignment.

This algorithm shows that beam training doesn't need massive online or offline computation, making it quite convenient for both algorithm implementation and practical deployment.

3. Comparison of the three main methods

Table 1 compares the three main methods from several perspectives, including basic principles, pilot expenses, offline and online computational complexity, hardware dependency, advantages and limitations, and suitable scenarios. From the perspective of pilot expenses, deep learning algorithms have the lowest, followed by compressed sensing algorithms, whose pilot expenses varies with sparsity. And the beam training algorithm requires a relatively high pilot expenses. In terms of offline and online computational complexity, compressive sensing algorithms mainly spend time on online computation, deep learning algorithms focus on offline neural network training, while beam training algorithms have a simple computation process. In terms of hardware dependence, compressed sensing and deep learning algorithms only need common hardware conditions, while beam training algorithms have higher requirements for hardware accuracy. Combining the advantages and limitations of three algorithms, this article concludes that compressed sensing algorithms are mainly suitable for analyzing sparse channel environments, deep learning algorithms are mainly suitable for environments with limited computing power, and beam training algorithms are mainly suitable for quick deployment scenarios.

Table 1. Comparison of the three methods

Comparison perspective	Compressed sensing	Deep learning	Beam training
Basic principle	Using sparsity in the angular domain to reconstruct as sparse signal recovery	Consider as image reconstruction using CNN/GNN	Preset phase shift codebook to do beam scanning for path alignment
Pilot expenses	Low, decreasing with sparsity	Extremely low, only a small amount of RIS units needed	Medium to high, depending on the codebook design

Table 1. (continued)

Offline computational complexity	None	High, requires a huge amount of data to train the neural network	Low, only need to design a phase shift codebook
Online computational complexity	Medium to high, iterative solving for sparse signal recovery	Low, with fast forward reasoning of neural network	Low, beam search and selection
Hardware dependence and feedback	Require the assumption of channel sparsity, standard hardware	Require training data and computing resources, standard hardware	Require phase shift codebook and is sensitive to hardware phase accuracy
Advantage	The theory is well-developed, easy to explain, and quite accurate	Low pilot expenses, fast online computation	Easy to implement, with no complex matrix operations
Limitation	The heaviness of iterative computation and it relies on the sparsity assumption.	Training has a high cost, and its ability to adapt to different environment is still unproven	Accuracy is limited, and codebook design affects performance
Applicable scenarios	Sparse Channel Environment for Theoretical Analysis	Environment with limited computing power	Scenario for quick deployment

4. Challenges and future research directions

From current research, this article summarizes a few open challenges. Firstly, one of the challenges is that in broadband millimeter-wave systems, the physical propagation angles of different subcarriers are different, resulting the simulated beam to split. Secondly, RIS material science is also developing rapidly. The emergence of STAR-RIS which can transmit and reflect signals at the same time has brought new channel estimation challenges. At the same time, in actual deployments, passive RIS has phase shift and amplitude-phase coupling errors caused by hardware limitation.

In response to the above-mentioned public challenges, future research on RIS-assisted millimeter-wave channel estimation has a few main directions. Firstly, one is to develop a channel estimation algorithm for broadband millimeter-wave systems that can work with very low expenses. In broadband millimeter-wave systems, because of the beam-split effect, the pilot expenses needed for channel estimation often grow exponentially with the number of antennas and subcarriers. So in the future, researchers need to combine compressed sensing principles with AI to develop a channel estimation algorithm that can achieve high accuracy with limited pilot resources, ensuring the transmission efficiency of broadband millimeter-wave systems. Secondly, a good direction is to develop an algorithm to jointly estimate the transmission and reflection cascaded channels in a STAR-RIS system with low pilot expenses. Since STAR-RIS can transmit and reflect signals at the same time, its cascaded channel dimension is doubled that of traditional RIS, so that channel estimation becomes much more challenging. In the future, researchers need to further explore the

relationship between transmission and reflection links and design efficient decoupling algorithms to achieve a fast and accurate estimation of the cascaded channels. Thirdly, to overcome errors caused by non-ideal hardware, people need to develop channel estimation algorithms with strong anti-interference in the future.

5. Conclusion

This article gives an overview of the latest channel estimation methods for RIS-assisted millimeter-wave communication systems. This article divides the existing algorithms for RIS-assisted millimeter-wave communication systems into three categories: compressed sensing, deep learning, and beam training. This article compares the advantage, limitation, and suitable using cases of each method through examples of typical algorithms, providing directions for future work on algorithm innovation and improvement in millimeter-wave system channel estimation. Then, this article compares three different methods in terms of pilot expenses, computational complexity, and hardware dependence. Finally, this article sorts out several cutting-edge research challenges, pointing out directions for further exploration of RIS-assisted 6G wireless communication systems in the future.

References

- [1] Drampalou, S. F., Miridakis, N. I., Leligou, H. C., & Karkazis, P. A. (2023). A survey on optimal channel estimation methods for RIS-aided communication systems. *Signals*, 4(1), 208–234. <https://www.mdpi.com/2624-6120/4/1/12>
- [2] Zheng, B., You, C., Mei, W., & Zhang, R. (2022). A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications. *IEEE Communications Surveys & Tutorials*, 24(2), 1035–1071. <https://ieeexplore.ieee.org/document/9722893/>
- [3] Yu, Y., Wang, J., Zhou, X., Wang, C., Bai, Z., & Ye, Z. (2023). Review on channel estimation for reconfigurable intelligent surface assisted wireless communication system. *Mathematics*, 11(14), 3235. <https://www.mdpi.com/2227-7390/11/14/3235>
- [4] Xue, Q., Ji, C., Ma, S., Guo, J., Xu, Y., Chen, Q., & Zhang, W. (2024). A survey of beam management for mmWave and THz communications towards 6G. *IEEE Communications Surveys & Tutorials*, 26(3), 1520–1559. <https://ieeexplore.ieee.org/document/10422712/>
- [5] Chen, Z., Tang, J., Zhang, X. Y., So, D. K. C., & Wong, K.-K. (2022). Hybrid evolutionary-based sparse channel estimation for IRS-assisted mmWave MIMO systems. *IEEE Transactions on Wireless Communications*, 21(3), 1586–1601. <https://ieeexplore.ieee.org/document/9521836/>
- [6] Gao, T., & He, M. (2023). Two-stage channel estimation using convolutional neural networks for IRS-assisted mmWave systems. *IEEE Systems Journal*, 17(2), 3183–3191. <https://ieeexplore.ieee.org/document/10025617/>
- [7] Rapudu, T. C., & Oyerinde, O. O. (2025). Machine learning-based channel estimation for multi-RIS-assisted mmWave massive-MIMO OFDM system in a dynamic environment. *IEEE Transactions on Wireless Communications*, 24(6), 5297–5309. <https://ieeexplore.ieee.org/document/10918600/>
- [8] Abdallah, A., Celik, A., Mansour, M. M., & Eltawil, A. M. (2023). RIS-aided mmWave MIMO channel estimation using deep learning and compressive sensing. *IEEE Transactions on Wireless Communications*, 22(5), 3503–3521. <https://ieeexplore.ieee.org/document/9944694/>
- [9] Wang, P., Fang, J., Zhang, W., Chen, Z., Li, H., & Zhang, W. (2022). Beam training and alignment for RIS-assisted millimeter-wave systems: State of the art and beyond. *IEEE Wireless Communications*, 29(6), 64–71. <https://ieeexplore.ieee.org/document/9771330/>
- [10] Wang, W., & Zhang, W. (2021). Joint beam training and positioning for intelligent reflecting surfaces assisted millimeter wave communications. *IEEE Transactions on Wireless Communications*, 20(10), 6282–6297. <https://ieeexplore.ieee.org/document/9410435/>
- [11] Wang, J., Tang, W., Jin, S., Wen, C.-K., Li, X., & Hou, X. (2023). Hierarchical codebook-based beam training for RIS-assisted mmWave communication systems. *IEEE Transactions on Communications*, 71(6), 3650–3662. <https://ieeexplore.ieee.org/document/10057262/>