An Indoor Fusion Fingerprint Localization Based on Channel State Information and Images

Wen Liu[1], Hong Chen[1], Zhongliang Deng[1], Changyan Qin[1] and Mingjie Jia[1]

[1]*School of Electronics Engineering, Beijing University of Posts and Telecommunications, Beijing 100089, China*

Abstract

With the intelligence of society, location-based service (LBS) plays an increasingly prominent role in daily life. In this paper, an indoor fusion fingerprint localization based on channel state information (CSI) and image is proposed to solve the limitation of single sensor in positioning. To reduce the dimension of image data, Shared Convolutional-Neural-Network based Add Fusion Network(SC-AFN) is proposed to process multi-directional images. In SC-AFN, initial features of images from different directions are extracted by shared Convolutional-Neural-Network (CNN), and then the features are fused by Add strategy and trained for vector representation of images. On this basis, the measure index based fusion representation model (MI-FRM) is proposed to fuse CSI features and SC-AFN image features. MI-FRM introduces the measurement index of fingerprint database into the fusion method. The parameters of MI-FRM are optimized by maximizing the discrimination of fusion fingerprint database to improve the matching accuracy for positioning. Experiments show that the fusion localization based on MI-FRM achieves better positioning performance, with the average error 0.62m in office scene.

Keywords

Indoor localization, channel state information, images, neural networks

1. Introduction

With the development of information and intelligence in society, location-based service (LBS) plays an increasingly prominent role in daily life. Indoor localization has been widely used in emergency rescue, logistics tracking, intelligent manufacturing and other aspects[1]. High quality LBS is based on high precision location information, so it is urgent to study on precise and reliable indoor localization technology to meet the needs of location service in complex indoor environment.

In recent years, a series of solutions for indoor localization have been proposed. According to the signal sensor used, the positioning technology can be divided into radio signals such as Wi-Fi, Bluetooth and Ultra-Wide-Band (UWB), as well as non-radio signals such as infrared, ultrasonic and vision [2, 3]. Although indoor localization with single sensor can obtain certain accuracy, its performance is limited in a complex indoor environment. The frequent shade, strong interference, weak light and non-line-of-sight in complex environment may lead to the feature losing, signal error and incomplete information when single sensor is used only. It

IPIN 2021 WiP Proceedings, November 29 – December 2, 2021, Lloret de Mar, Spain
liuwen@bupt.edu.cn (W. Liu); chenhong@bupt.edu.cn (H. Chen); dengzhl@bupt.edu.cn (Z. Deng); qin_changyan@163.com (C. Qin); jiamingjie@bupt.edu.cn (M. Jia)
0000-0002-6450-1969 (W. Liu); 0000-0002-4884-7433 (H. Chen)

would inevitably affect the accuracy and reliability of positioning system. Therefore, the fusion localization with multi-sensor has become an important trend in indoor positioning technology [4].

Among various signals for indoor positioning, Wi-Fi and visual signal have become the research hotspots currently due to their advantages of rich positioning information and low hardware cost. The regular indicators of Wi-Fi are the received signal strength (RSS) and channel state information (CSI), among which CSI provides more detailed subcarrier information with stronger time stability. Therefore, CSI-based indoor localization can achieve better positioning performance [5]. The accuracy of localization based on CSI has reached the meter-level in reports, but there are still some problems such as insufficient discrimination in data and difficulty in determining the unique location in the complex indoor environment. Visual signal-based positioning is widely used in navigation and positioning because it does not need to deploy equipment in advance and the cost of hardware is low [6]. The image match for indoor positioning has the advantages of stable features and low noise influence, but multi-directional images for positioning always have the problems of high feature dimension, complex work-flow and difficulty in achieving real-time performance. To sum up, the CSI features of Wi-Fi signal and the image features of visual signal have their own advantages and disadvantages, and the fusion localization with the two signal features can improve the completeness and accuracy of location information to achieve better performance.

At present, the multi-sensor fusion positioning systems usually include decision-level fusion and feature-level fusion [7]. Decision-level fusion is usually used to fuse the preliminary positioning results of different positioning sensors at a higher level to get the final result. Although the position accuracy can be improved to a certain extent in decision-level fusion, there are still problems such as complex process and insufficient fusion depth. In contrast, feature-level fusion can simplify the positioning process and increase the depth of fusion. In feature-level fusion, fingerprint database is usually constructed after fusion of heterogeneous features from different sensors, and then location estimation is realized by fingerprint matching [8]. Theoretically, the estimated location based on feature-level fusion can further integrate the rich location information in multi-sensor data. However, there is still a lack of effective unified representation of heterogeneous features for fusion fingerprint database with high discrimination, which leads to insufficient positioning accuracy.

The most important contribution of this paper is the design of an indoor fusion fingerprint localization based on channel state information and multi-directional images. In order to improve the positioning accuracy by combining the two heterogenous information from different sensors, Shared Convolutional-Neural-Network based Add Fusion Network (SC-AFN) is proposed to extract the key vector feature from multi-directional images and the measure index based fusion representation model (MI-FRM) of the two location features is proposed to realize the fusion representation with high discrimination for the fingerprint localization. In SC-AFN, the shared Convolutional-Neural-Network(CNN) model is used for the initial feature extraction of images in all directions, the Add strategy is adopted for the disordered fusion of multiple features and the full connection is used to extract the target related features for the final representation of the images. In MI-FRM, the perceptron is trained with the aim of optimizing the measure index of heterogeneous fusion features to obtain the best representation model, and the fusion fingerprint with high discrimination can be constructed for the final fingerprint localization.

The rest of this paper is organized as follows: The second section introduces the related work. In the third section, the proposed system is introduced, including the image feature extraction method SC-AFN, fusion fingerprint construction method MI-FRM and the localization method. The fourth section introduces the experimental results. In the fifth section, this paper is summarized and the future work is prospected.

2. Related works

Multi-carrier CSI fingerprint has rich location information, but is insufficient in data discrimination. In fingerprint positioning, data processing of original CSI data can enhance fingerprint discrimination and improve positioning accuracy [9]. In [10], with the analysis of CSI amplitude, we propose Local Connection based Deep Neural Network (LC-DNN) to realize high precision positioning. LC-DNN adopts multiple non-shared convolution kernels to extract the local feature in different frequency ranges of CSI amplitudes and adopts full connection to extract the target related global feature from the spliced local feature. Thus the features with high discrimination are obtained for positioning. Since this paper focuses on the dimensionality reduction for multidirectional images and the fusion for multi-sensor, LC-DNN is directly adopted for feature processing of CSI amplitude.

Image feature description is the key point in image fingerprint localization [11]. In [12], the authors propose a two-stage image fast search and matching localization method. Firstly, similar images are quickly selected through Histogram of Oriented Gradient(HOG) feature, and then the pose estimation and position estimation are further realized by the match of accurate image feature Affine-Scale Invariant Feature Transform(A-SIFT). However, HOG, SIFT and other traditional descriptors are still manually image feature extractors and models, which require a specific professional foundation. Deep learning technology provides the possibility of automatic extraction of image features and has become a research hotspot in recent years. In [13], the author introduces transfer learning to indoor positioning to improve the training speed and accuracy of model. The image sets collected by the robot are used to retrain the pre-trained VGG-16, and the last layer of the network is adopted as the image feature for matching and positioning. Nowadays, deep learning descriptors have been widely used in single image feature description. The images collected from different directions at the same location could provide more abundant position information for precise localization. However, indoor positioning based on multi-directional images still has problems such as high feature dimension and complex matching process. There is still a lack of research on unified vector representation of multi-directional images for positioning. In this paper, SC-AFN is designed to realize effective vector representation of multi-directional images through shared CNN and Add fusion, which is convenient for subsequent multi-sensor fusion.

For the feature-level fusion localization of multi-sensor signals, a variety of heterogeneous fusion fingerprint construction and matching algorithms have been proposed. In [14], based on the fusion of geomagnetism and image information, the author proposes a heterogeneous feature database to achieve a positioning accuracy of 0.85m in the laboratory environment. But this method is only applicable to the space with significant magnetic field features. In [15], the authors propose a feature level fusion of CSI amplitude information and geomagnetic intensity information to construct a heterogeneous fusion fingerprint for localization. And the multidimensional scaling K-Nearest-Neighbor matching (MDS-KNN) algorithm is proposed for matching location, with the mean error of 1.7 m. However, this method only realizes the simple combination of features, and fails to change the discrimination of features. To enhance the feature discrimination, many researchers try to introduce the measure index, which measures the degree of similarity between features, into feature construction and matching. In [16], the author proposes a new similarity measure method based on LP measure between fuzzy sets, which improves the accuracy of face recognition system. In [17], the authors divide the image into blocks to extract the contour and uses Jaccard coefficient as the similarity measure of the binary image, which effectively solves the recognition problems caused by the rotation and deformation of the image. In this paper, a fusion representation model based on measure index, MI-FRM, is proposed for distinctive fusion localization. The perceptron-based fusion representation model is trained with the discrimination as the optimization objective, and the heterogeneous fusion fingerprint database with high discrimination is constructed to improve the positioning accuracy.

3. Fusion localization based on MI-FRM

To realize multi-sensor localization, a fusion fingerprint localization based on MI-FRM is proposed in this paper. The fusion fingerprint database is the key point. In the fusion representation of multi-sensor signal features, CSI and multi-direction images need to be processed respectively first. In this paper, SC-AFN algorithm is proposed to construct the vector feature of multi-directional images. The MI-FRM algorithm is used to fuse CSI features and multi-directional images for precise indoor localization. In addition, the LC-DNN method [10] is adopted for data processing to ensure the validity of CSI data. The framework is shown as Figure 1.

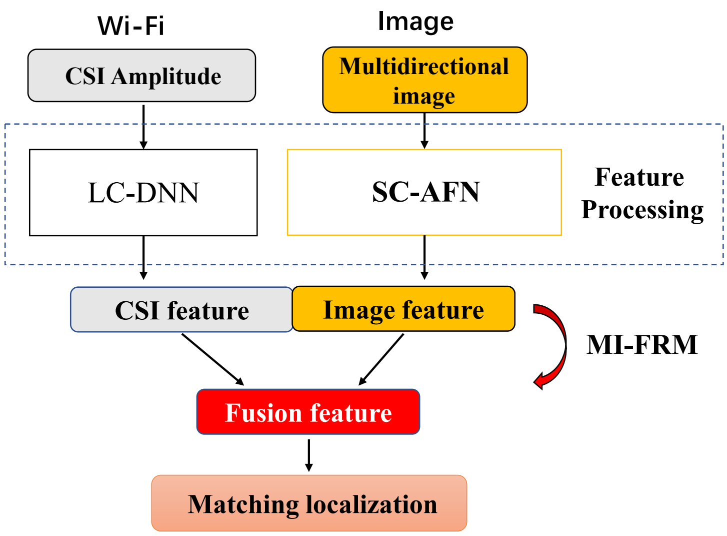


Figure 1: Framework of fusion indoor localization based on MI-FRM.

3.1. SC-AFN Algorithm

Different visual images of the environment can be collected at fixed position points from different directions, as shown in Figure 2. Multi-directional environment images could jointly record the rich position information about collected points and have better discrimination for position estimation. However, there are still some problems in the localization based on multidirectional images. First, it is usually difficult to obtain the direction label of the images in real time to ensure the sequence of the input images. Second, the images lead to higher dimensions in positioning. Therefore, the fusion representation of the disordered multi-directional images is still needed for fingerprint localization.

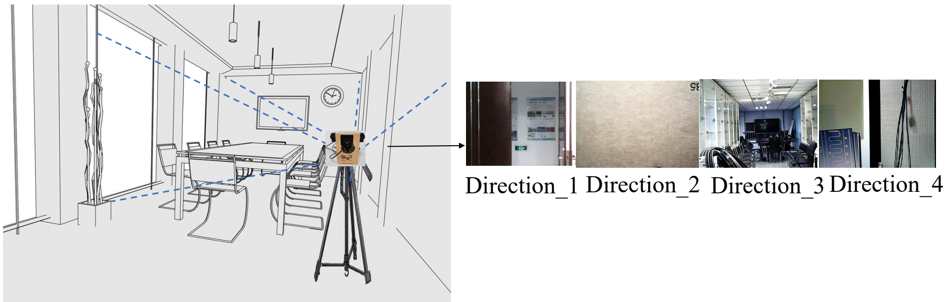


Figure 2: The schematic diagram of multi-directional image.

In this paper, SC-AFN is proposed to process and fuse the multi-directional images at each position, and construct the vector representation. This method innovatively realizes the fusion characterization of disordered multi-directional images, and reduces the dimension of images to complete the structural homogenization of heterogenous data, which lays the foundation for multi-sensor fusion. SC-AFN consists of three parts: shared CNN, Add fusion and full connection. The structure is shown as Figure 3.

In the feature extraction of image, the multi-directional images at each position belong to the same type environment, so the same model can be used for feature extraction and description. CNN is a common feature extraction and classification method for image, which has the advantages of strong generality, small number of parameters and high classification accuracy. Therefore, we adopt the shared CNN model to extract the initial features of images from each direction. The shared CNN includes convolution, pooling, dropout and flattening layers. Figure 4 shows the structure of shared CNN. Firstly, the rotation invariant and translation invariant features of $Image_n$(the input images of direction_n) are extracted by convolution method. And then, the max pooling simplifies the data and reduces the dimensions of the multi-dimensional feature images to enlarge the receptive field and prevent overfitting. ReLU function is adopted to realize the nonlinear transformation. The two-dimensional convolution and activation function are formulated as (1) and (2).

$$s(i,j) = (X * W)(i,j) + b = \sum_{k=1}^{n\_in} (X_k * W_k)(i,j) + b, \tag{1}$$

$$\mathrm{ReLU}(x) = \max(0, x), \tag{2}$$

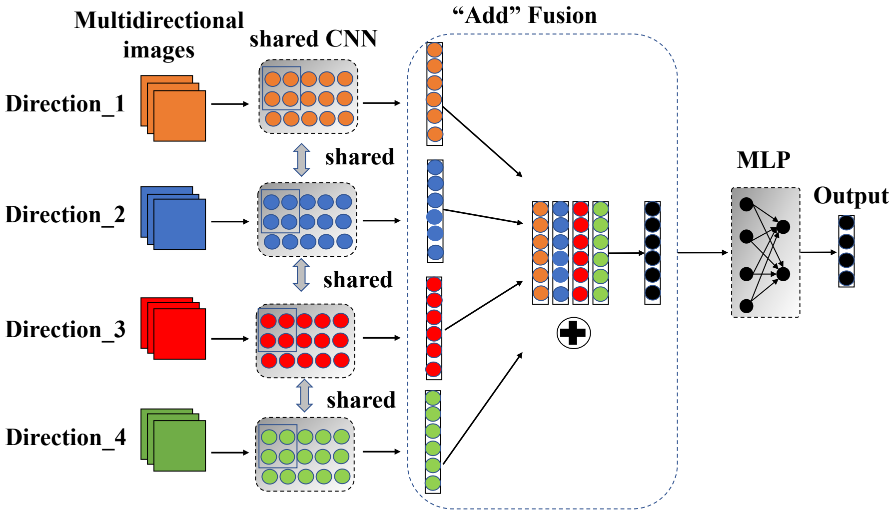


Figure 3: The structure of SC-AFN.

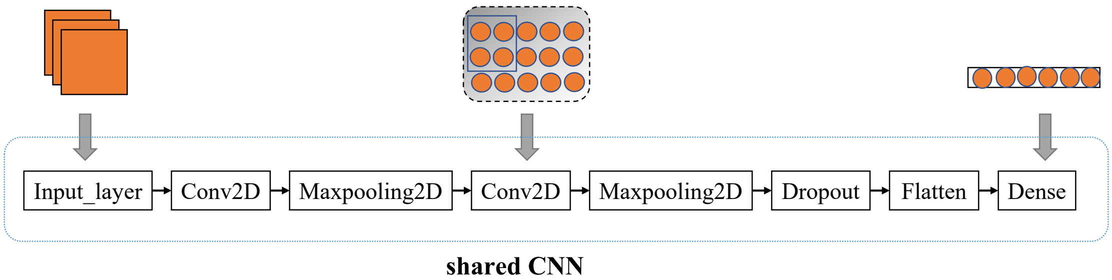


Figure 4: The structure of shared CNN.

where $s(i, j)$ is the output value at the corresponding position, n_in is the total number of channels for input data, X_k is the input matrix of channel k, W_k represents the kth channel's subconvolution kernel matrix and b is the bias. To enhance the robustness of the network, the dropout strategy is adopted to randomly discard part of the neuron nodes, so that the calculation results contain more random structures. Finally, the one-dimensional vector feature of the image h_n can be obtained. The feature extraction of the single-direction image could be expressed as:

$$h_n = \text{shared\_model}(Image_n), \tag{3}$$

where shared_model is the shared CNN.

After the extraction of single-direction image, the network needs to adopt the fusion strategy to integrate the multi-directional image features. The fusion strategy includes Contract and Add. Contract method is often used to slice features of different channels, which means that the number of feature dimensions increases but the amount of information in each dimension remains unchanged. The Add strategy mainly completes the superposition among features, and increases the information of each dimension while the feature dimension remains unchanged. For multi-directional images, Contract has the problems of high data dimension, much redundant information, and great dependence on the splicing order. Therefore, SC-AFN adopts Add strategy to fuse the features of images from different directions element by element to achieve disordered fusion. The feature map of the multi-directional image can be expressed as:

$$h = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n} \\ h_{2,1} & h_{2,2} & \dots & h_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{m,1} & h_{m,2} & \dots & h_{m,n} \end{bmatrix}, \tag{4}$$

where m is the number of directions and n is the dimension of image feature. The fusion feature based on Add strategy can be expressed as:

$$r = [r_1 \, r_2 \, \cdots \, r_n], \tag{5}$$

where $r_i = \sum_{k=1}^{m} h_{k,i}$.

Finally, the network adopts full connection layer to integrate the fusion features, and obtain the final fusion features related to position. And the Softmax function is used to classify the final features. The probability of class i is formulated as :

$$\mathrm{softmax}\,(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}, \tag{6}$$

where $(x)_i$ is the $i\text{-}th$ eigenvalue.

For multi-classification, the cross-entropy loss function is adopted for network optimization, which is shown as

$$\mathrm{H}_{y'}(y) = -\sum_i {y'}_i \log(y_i), \tag{7}$$

where y is the predicted probability distribution and y' is actual probability distribution. The network parameters are trained through Gradient descent method iteratively by minimizing the loss function.

3.2. MI-FRM Algorithm

In this paper, a fusion representation algorithm MI-FRM is proposed to transfer CSI and image feature domains to the fusion domain. This method mainly proposes the measure index to quantitative the discrimination of fingerprint database, and optimizes the parameters of the fusion model by maximizing the measure index. The structure is shown as Figure 5.

The MI-FRM adopts basic perceptron as the fusion representation model. The input is the splicing data of CSI amplitude feature and multi-directional image feature, and the output is the fusion feature. In MI-FRM, full connection and activation function are used to fit the process of fusion representation mapping. Linear transformation of the feature domain is realized by

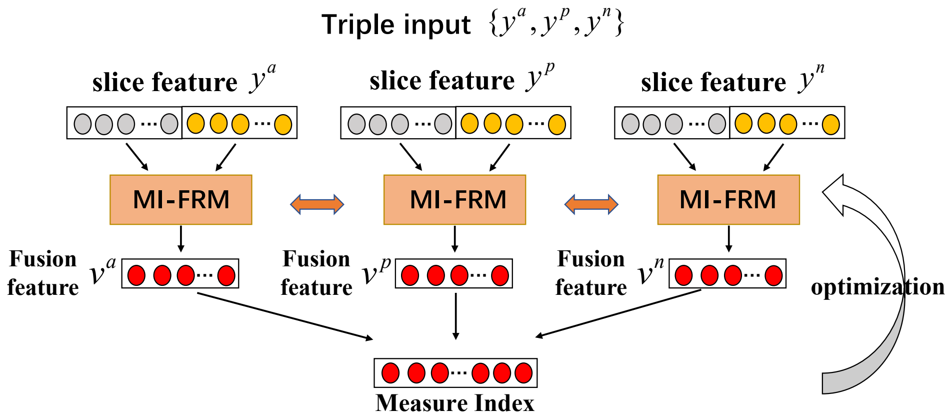


Figure 5: Fusion representation algorithm based on MI-FRM.

assigning weight parameters to different dimensions of the two kinds of features, and nonlinear transformation of the feature domain is realized by activation function. Then the feature fusion representation domain is constructed.

The expression of the initial fusion representation model is as follows:

$$v_k^m = \frac{1}{1 + \exp\{-c(\sum_{i=1}^{N_{CSI}} w_i^m y_{CSI,k}^i + \sum_{j=1}^{N_{IMA}} w_{j+N_{CSI}}^m y_{IMA,k}^j)\}} (m = 1, 2 \cdots, M), \tag{8}$$

where v_k^m is the m-th element in the fusion feature vectors at the kth fingerprint point, M is the dimension of the fusion feature, $y_{CSI,k}$ and $y_{IMA,k}$ are the CSI feature vectors and multi-directional image feature vectors at the kth fingerprint point, N_{CSI} and N_{IMA} are the dimensions of the two feature vectors, and $\{c, w_1, w_2, \cdots, w_{N_{CSI}+N_{IMA}}\}$ are the tunable parameters of the model.

High-quality fusion fingerprint should meet the following requirements: the fingerprints at the same location are closer to each other, while those at different locations are farther apart. According to this principle, the classical Euclidean distance is used to calculate the basic distance between two fusion representations, then the measurement index of the whole database can be defined. We divided all fusion fingerprints into triples. Each triplet contains anchor samples v^a, positive samples v^p with the same location as anchor samples, and negative samples v^n with different locations. To reduce the distance of the fingerprints at the same location and increase the distance of the fingerprints at different location points, the measurement index of the fusion database can be expressed as:

$$D = \sum_{i=1}^{N} \left[\left\| v_{(i)}^a - v_{(i)}^n \right\|_2^2 - \left\| v_{(i)}^a - v_{(i)}^p \right\|_2^2 - \alpha \right] (\alpha > 0), \tag{9}$$

where $v_{(i)}^a$, $v_{(i)}^p$and $v_{(i)}^n$ are the fingerprint of anchor position, positive position and negative position in the i-th triad respectively, α is the minimum threshold to distinguish the distance between negative pair and positive pair and N is the total number of triplet samples.

To construct discriminative fingerprints, this paper takes the measurement index of fusion feature space as the optimization objective, and adopts Adaptive Moment Estimation (Adam) training algorithm to optimize the parameters of MI-FRM. The parameters are as follows:

$$w = \{c, w_1, w_2, \cdots, w_{N_{CSI}+N_{IMA}}\}. \tag{10}$$

These parameters represent the contribution of each dimension in the feature vector to the fusion representation domain. In the optimization process, appropriate parameters are selected to make the information with high discriminative degree have relatively high contribution in the fusion representation.

The measure index of fusion database is proportional to the difference degree of fingerprints at different locations, and inversely proportional to the difference degree of fingerprints at the same location. The target of fingerprint database construction with high discrimination should maximize the measure index. Therefore, the negative number of the measure index is selected as the objective function of minimization:

$$L = -D = \sum_{i=1}^{N} \left[\left\| v_{(i)}^{a} - v_{(i)}^{p} \right\|_2^2 - \left\| v_{(i)}^{a} - v_{(i)}^{n} \right\|_2^2 + \alpha \right]. \tag{11}$$

Adam algorithm is adopted as the optimization algorithm, and the update formula is shown as follows:

$$w_t = w_{t-1} - \alpha \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t} + \varepsilon}, \tag{12}$$

where t is the number, $\widehat{m}_t$ is the correction of m_t, $\widehat{v}_t$ is the correction of v_t.

$$\widehat{m}_t = \frac{m_t}{1-\beta_1^t}, \widehat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{13}$$

where β_1 and β_2 are constants that control exponential attenuation, m_t is the exponential moving mean of the gradient (obtained by the first moment of the gradient), and v_t are square gradient (obtained by the second moment of the gradient).The update formula of, is as follows:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t, \tag{14}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2, \tag{15}$$

where g_t is the first derivative. The default setting for the parameters is : $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$.

3.3. Localization

The fusion localization based on MI-FRM is realized by the construction and matching of fusion fingerprints, including offline and online stages.

In the offline stage, original data of CSI and multi-directional images are collected at each reference point. For CSI, LC-DNN is trained with the CSI amplitudes and position labels to obtain

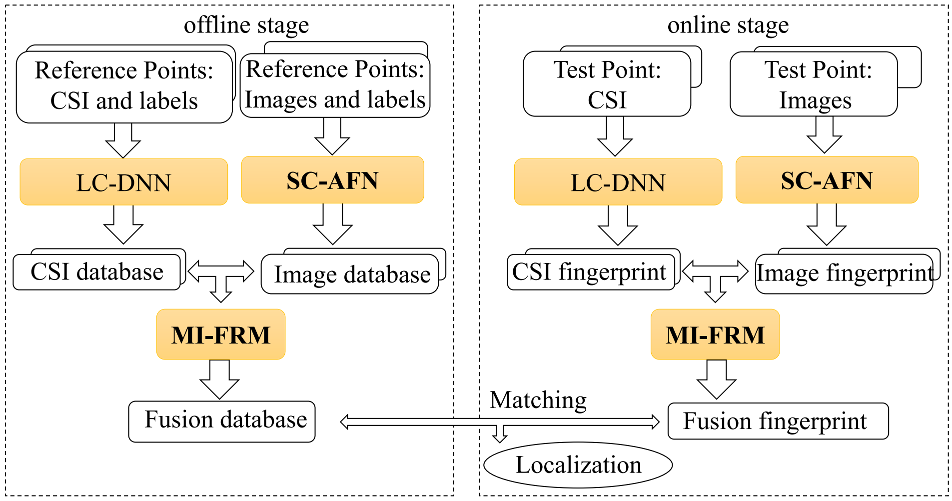


Figure 6: Fusion fingerprint matching localization.

the optimal weights of feature extraction and the effective feature database of CSI amplitudes. For images, SC-AFN is trained with multi-directional images and position labels to obtain the optimal weights and the vector representation database for multi-directional images. The data processing of the above two kinds of location data can realize the structural homogenization of features, which lays the foundation for the next feature-level fusion. To further fuse the CSI and image feature, the two features at the same location are spliced to construct the training set, and the parameters of the MI-FRM are optimized with the combination of the training set and labels to obtain the optimal weights and the fusion fingerprint database.

In the online stage, the real-time fusion fingerprint based on MI-FRM is constructed with the real-time CSI and image data and is matched with the fusion database to estimate the location. For CSI, the feature extraction is realized through the trained LC-DNN. For images, the trained SC-AFN is adopted for vector representation of multi-directional images. Then, the MI-FRM could transform the data of CSI amplitude feature and multi-directional image into the fusion representation domain. Finally, the location is estimated by the match of fusion fingerprint. Since the matching algorithm is not focused in this paper, the Softmax classifier is used directly here.

4. Experiments

4.1. Experiment environments

This part mainly verifies the performance of fusion localization based on MI-FRM. Data collection in experiments includes two sources, CSI and image. The equipment is shown in Figure 7 and Figure 8 respectively. The equipment for CSI collection includes transmitter and receiver, both of which are mobile terminals with built-in Intel 5300 wireless network card. The transmitter sends packets at 5ms intervals under the bandwidth of 20MHz, and the receiver collects and stores the corresponding packets. The CSI information can be solved by modifying the underlying driver.

The equipment for image collection is jointly built by the stereo box and four low-cost cameras. Four cameras are fixed in the four directions of stereo box and connected with the same laptop computer through the data line respectively. Therefore, the multi-directional image data can be collected by single computer operation.

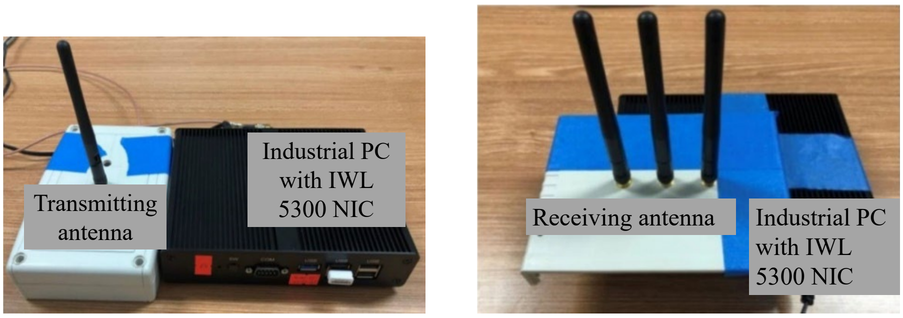


Figure 7: The transmitter (left) and receiver (right) of CSI.

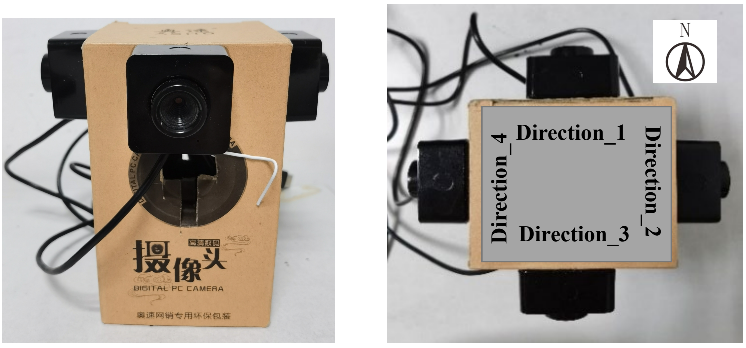


Figure 8: The collect equipment of multidirectional image (Left: front view, Right: top view).

Datasets: The data collection included 52 reference position points and 52 test points, 1000 CSI packets and 40 multi-directional images collected at each position. The experiment environment is the comprehensive office scene, which is composed of laboratory, meeting room and corridor. It is characterized by many obstacles, independent rooms and frequent flow of people. In the office scene, obstacles in room mainly include tables, chairs and computers, and three rooms are separated by walls. The total area in experiment is 152.9m^2, of which the laboratory area is 16.4m*4.4m and the meeting room area is 8.4m*1.8m. 52 reference points and 52 test points are set, and the interval of reference points is 1.2m. In CSI collection, four fixed transmitter and one mobile receiver are used. In image collection, four cameras are controlled by a laptop computer at the mobile end. Data collection is mainly arranged in non-work hours. Figure 9 shows the layout of the office scene and the distribution of points.

(a) laboratory (b) meeting room (c) corridor

(d) Layout of the office scene and reference/test points

Figure 9: The diagram of the experiment environments.

4.2. The evaluation of SC-AFN

To verify the positioning performance of SC-AFN, this part mainly compares SC-AFN with the positioning method based on VGG-16 proposed by Wozniak P et al in 2018 from positioning accuracy and stability[13]. Both methods adopt the above datasets in office scene to ensure the objectivity of experiments.

Figure 10(a) shows the comparison of CDFs in two image localization. Overall, the trend of CDF curves in the two methods are basically similar because the data sets and feature matching methods are similar. However, the CDF of SC-AFN is always higher than the image positioning method based on VGG-16 and more inclined to the upper left, which indicates that SC-AFN achieves superior positioning performance. At the same time, the maximum positioning errors of SC-AFN and VGG-16 methods are 13.5m and 17.7m respectively. In contrast, the unified vector representation based on multi-directional images (SC-AFN) effectively reduces the probability of large errors and improves the positioning accuracy.

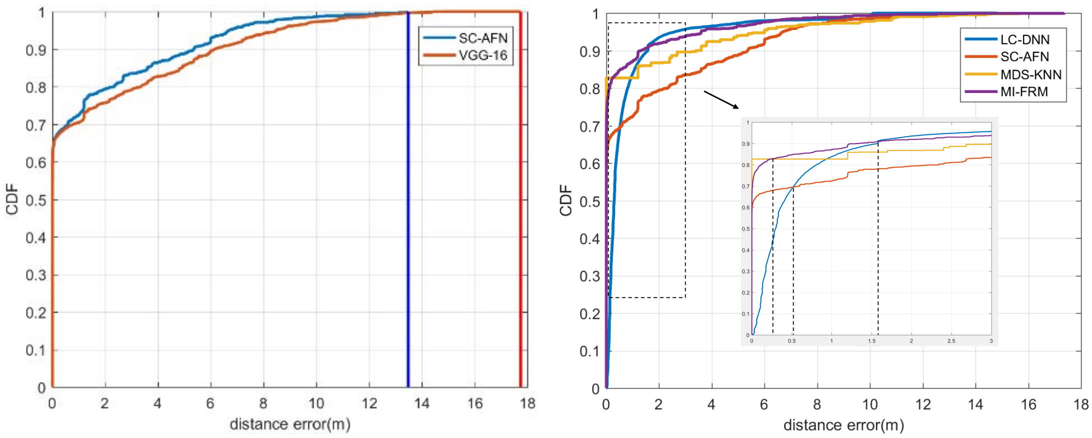
Table 1
Image Localization Errors of Different Methods

	SC-AFN	VGG-16
mean/m	1.25	1.96
std/m	2.47	2.96

Table 1 and Figure 11(a) show the comparison of the positioning mean error and standard deviation in two image localization methods. The mean error of SC-AFN is 1.25m, which is 44.3% lower than that of the 1.96m of VGG-16 method. In terms of stability, the standard deviation of SC-AFN is 2.47m, which is 16.5% less compared with 2.96m based on VGG-16 method. Combined with CDFs, the vector representation network of multi-directional image features (SC-AFN) further improves the positioning accuracy and stability by reducing the probability of large errors.

4.3. The evaluation of MI-FRM

To further evaluate the indoor fusion localization based on MI-FRM, this part mainly compares the proposed algorithm with LC-DNN (CSI database)[10], SC-AFN (image database), and MDS-KNN (fusion database)[15] to verify the effectiveness of positioning performance improvement by introducing measure index into fingerprint construction. The experiments focus on accuracy and stability of positioning.



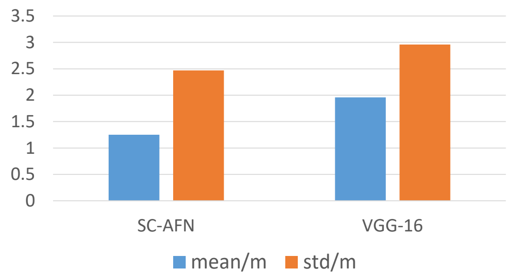
(a) (b)

Figure 10: The CDFs of different methods.

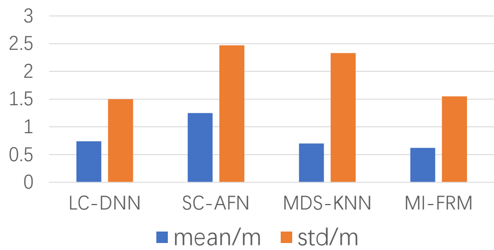
Figure 10(b) shows the CDF of the four methods for indoor localization. As can be seen, in the comparison of single fingerprint database, the SC-AFN (image localization) and LC-DNN (CSI localization) have an intersection point at the error of 0.5m. It indicates that the image localization achieves better effect when the error is less than 0.5m and the CSI localization achieves better effect when the error is larger than 0.5m. Both fingerprint databases have their own advantages. Compared with single fingerprint database, MI-FRM have significant improvement within the error of 1.5m, and the overall performance is more superior. It also shows that fusion fingerprint could effectively combines the information of two positioning features and has higher discrimination than single fingerprint. Compared with MDS-KNN, MI-FRM performs slightly worse within the error of 0.25m, which may be caused by the reduction of feature dimension and the direct loss of information. However, MI-FRM has obvious advantages when the error is larger than 0.25m, with more stable curve and better positioning performance. Therefore, it is further verified that the introduction of the measure index in the database construction could improve the feature discrimination and positioning accuracy.

Table 2
Localization Errors of Different Methods

	LC-DNN	SC-AFN	MDS-KNN	MI-FRM
mean/m	0.74	1.25	0.70	0.62
std/m	1.55	2.47	2.33	1.55



(a) Image localization errors in different methods.



(b) Localization errors in different methods.

Figure 11: The comparison of localization errors in different methods.

TABLE 2 and Figure 11(b) show the comparison of mean error and standard deviation in four positioning methods. The mean error of MI-FRM is 0.62m, which is 16.2%, 50.4% and 11.4% lower than the 0.74m of LC-DNN (CSI database), 1.25m of SC-AFN (image database) and 0.70m of MDS-KNN (fusion database) respectively. The standard deviation of MI-FRM is 1.55m, which is basically unchanged compared with LC-DNN (CSI database), 37.3% higher than that of SC-AFN (image database) and 33.4% higher than that of MDS-KNN (fusion database). The results show that MI-FRM could combine the advantages of two localization features effectively to construct a high-precision positioning fingerprint database with more abundant information, and achieve better stability than single fingerprint or other fusion method.

5. Conclusion

In this paper, we present a fusion fingerprint localization based on MI-FRM to improve the positioning accuracy and robustness. In this method, SC-AFN is proposed to construct vector features for images from different directions and MI-FRM is designed to fuse image features and CSI features to a new distinctive representation for precise fingerprint positioning. Experiments in office scene show that SC-AFN realizes vector representation and reduces the probability in large errors, with an average error of 1.25m. The MI-FRM effectively combined the two positioning signals to further improve the feature discrimination and positioning accuracy. The average error of MI-FRM is 0.62m, which is 16.2%, 50.4% and 11.4% higher than LC-DNN, SC-AFN and MDS-KNN, respectively.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China under Grant No.61871054.

References

[1] H. Ht, T. Le Vu, Location-Based Services, ABC Journal of Advanced Research 8 (2019) 89–94. doi:10.18034/abcjar.v8i2.91.

[2] B. Lu, J. Niu, J. Juny, L. Cheng, Y. Guy, WiFi Fingerprint Localization in Open Space, in: Proc. IEEE LCN, 2013, pp. 1–4.

[3] J. Aparicio, F. J. Álvarez, Á. Hernández, S. Holm, A review of techniques for ultrasonic indoor localization systems, The Journal of the Acoustical Society of America 145 (2019) 1884–1884. doi:10.1121/1.5101825.

[4] L. Antsfeld, B. Chidlovskii, E. Sansano-Sansano, Deep Smartphone Sensors-WiFi Fusion for Indoor Positioning and Tracking, arXiv preprint arXiv:2011.10799 (2020).

[5] W. Liu, Q. Cheng, Z. Deng, H. Chen, X. Fu, X. Zheng, S. Zheng, C. Chen, S. Wang, Survey on CSI-based Indoor Positioning Systems and Recent Advances, in: 2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2019, pp. 1–8. doi:10.1109/IPIN.2019.8911774.

[6] N. Ravi, P. Shankar, A. Frankel, A. Elgammal, L. Iftode, Indoor Localization Using Camera Phones, in: Seventh IEEE Workshop on Mobile Computing Systems & Applications (WMCSA'06 Supplement), IEEE, 2007, pp. 1–7. doi:10.1109/WMCSA.2006.4625206.

[7] J. Bleiholder, F. Naumann, Data fusion, ACM computing surveys (CSUR) 41 (2009) 1–41. doi:10.1145/1456650.1456651.

[8] X. Guo, N. Ansari, L. Li, H. Li, Indoor Localization by Fusing a Group of Fingerprints Based on Random Forests, IEEE Internet of Things Journal 5 (2018) 4686–4698. doi:10.1109/JIOT.2018.2810601.

[9] V. Honkavirta, T. Perala, S. Ali-Loytty, R. Piché, A comparative survey of WLAN location fingerprinting methods, in: 2009 6th workshop on positioning, navigation and communication, IEEE, 2009, pp. 243–251. doi:10.1109/WPNC.2009.4907834.

[10] W. Liu, H. Chen, Z. Deng, X. Zheng, X. Fu, Q. Cheng, LC-DNN: Local Connection Based Deep Neural Network for Indoor Localization with CSI, IEEE Access 8 (2020) 108720–108730. doi:10.1109/ACCESS.2020.3000927.

[11] J. Zhang, A. Hallquist, E. Liang, A. Zakhor, Location-based image retrieval for urban environments, in: 2011 18th IEEE International Conference on Image Processing, IEEE, 2011, pp. 3677–3680. doi:10.1109/ICIP.2011.6116517.

[12] Y. Huang, H. Wang, K. Zhan, J. Zhao, P. Gui, T. Feng, IMAGE-BASED LOCALIZATION FOR INDOOR ENVIRONMENT USING MOBILE PHONE, ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-4/W5 (2015) 211–215. doi:10.5194/isprsarchives-XL-4-W5-211-2015.

[13] P. Wozniak, H. Afrisal, R. G. Esparza, B. Kwolek, Scene Recognition for Indoor Localization of Mobile Robots Using Deep CNN, in: International Conference on Computer Vision and Graphics, Springer, Springer, Cham, 2018, pp. 137–147. doi:10.1007/978-3-030-00692-1_13.

[14] Z. Liu, L. Zhang, Q. Liu, Y. Yin, L. Cheng, R. Zimmermann, Fusion of Magnetic and Visual Sensors for Indoor Localization: Infrastructure-Free and More Effective, IEEE Transactions on Multimedia 19 (2017) 874–888. doi:10.1109/TMM.2016.2636750.

[15] X. Huang, S. Guo, Y. Wu, Y. Yang, A fine-grained indoor fingerprinting localization based on magnetic field strength and channel state information, Pervasive and Mobile Computing 41 (2017) 150–165. doi:10.1016/j.pmcj.2017.08.003.

[16] M. A. El-Sayed, K. Hamed, et al., Study of Similarity Measures with Linear Discriminant Analysis for Face Recognition, Journal of Software Engineering and Applications 8 (2015) 478–488. doi:10.4236/jsea.2015.89046.

[17] N. Kane, K. Aznag, A. El Oirrak, M. N. Kaddioui, Binary Data Comparison using Similarity Indices and Principal Components Analysis, The International Arab Journal of Information Technology 13 (2016) 232–237.