NLOS identification for UWB based on channel impulse response

Zhuoqi Zeng[1], Steven Liu[2], Lei Wang[3]
1 CR/RTC5-AP, Bosch (China) Investment Ltd., Shanghai, China, Email: Zhuoqi.ZENG@cn.bosch.com
2 Institute of Control Systems, University of Kaiserslautern, Germany, Email: sliu@eit.uni-kl.de
3 Sino-German School for Postgraduate Studies of Tongji University, Shanghai, China, Email: leiwang@tongji.edu.cn

***Abstract*—The localization accuracy of ultra-wide band (UWB) system could be dramatically degraded, if the signal is propagated under non-line-of-sight (NLOS) condition. The detection of the NLOS propagation is very important to guarantee the accuracy of the UWB system. Based on the channel impulse response (CIR) sample, the NLOS condition could be identified. However, for the decawave chips, each CIR sample contains 1015 points. Thus the real-time realization of the NLOS detection with CIR is very hard, since the import and calculation of such a large amount of data cause to huge delay. In order to reduce the delay, the minimal needed size of the points in CIR for accurate NLOS identification is discussed in this paper. The support vector machine (SVM) is used for the classification based on the original CIR points or the eight different features extracted from each CIR. Furthermore, a new method is proposed for the identification based on the convolution algorithm. Compared to the existing approach with CIR, the needed CIR points for the detection are dramatically reduced, which makes the on-line identification realization possible. The accuracy of the NLOS identification with less CIR points is even better. The new proposed method using convolution algorithm also shows very promising results compared the other approaches.**

Keywords—localization; UWB; NLOS identification; CIR; SVM; convolution algorithm

I. INTRODUCTION

Indoor position awareness plays an important role in many applications such as service, indoor localization and tracking systems. Among all the existing indoor localization systems, ultra-wide band (UWB) is one of the most promising technologies due to its low-power consumption, low complexity, and robust operation in harsh indoor environments. However, if the direct signal propagation path between the mobile station (MS) and base stations (BSs) is obstructed, a non-negligible positive bias is added to the measurements, which is defined as non-line-of-sight (NLOS) propagation. The localization accuracy can be dramatically degraded under NLOS propagation.

The range estimation [1] together with CIRs feature based methods [2], [3], [4] are the most widely used approaches to identify the NLOS propagation. The main drawback of the range estimation based detection method is the unavoidable additional latency due to the collection of the ranges. Furthermore, the detection accuracy is also not so attractive. Two different features (skewness and root-mean-square delay spread) extracted from CIR are used in [8] to identify the NLOS condition. In [9], the NLOS identification is realized with six features based on import vector machine. These used features are received signal energy, maximum amplitude, rise time, mean excess delay, RMS delay spread and kurtosis. A survey of the most widely used NLOS identification methods for UWB could be found in [10]. For the existing CIR feature based methods, there is no discussion about the needed points in each CIR sample for accurate identification. The more points used for detection, the more data need to be saved, thus huge delay might be caused. Furthermore, with the increase of the BSs, the data amounts enhance dramatically. The NLOS identification could be realized with the help of an additional sensor. The NLOS identification and accurate UWB measurements selection approach with the help of the inertial measurement unit (IMU) is proposed in [5], [11]. Another UWB/IMU sensor fusion approach is presented in [6]. In this method the line of sight condition can be guaranteed for the foot mounted hardware set-up, which makes the NLOS detection unnecessary. These approaches require a second sensor source. In this paper, the NLOS detection for the stand-alone UWB system based on CIR stays in the focus.

The full CIR received by the BS with a Decawave chip is shown in Fig. 1. Only a small part of the points (marked in red) provide the useful information about the received impulse signal, the rest (marked in blue) are just noise. Furthermore, even for the actual useful information, not all the points have to be used for the identifications of the NLOS condition. The latency of the system could be reduced if less points are saved for the detection purpose.

In this paper, the minimal needed points in CIR for the features extracted method based on SVM is discussed. The

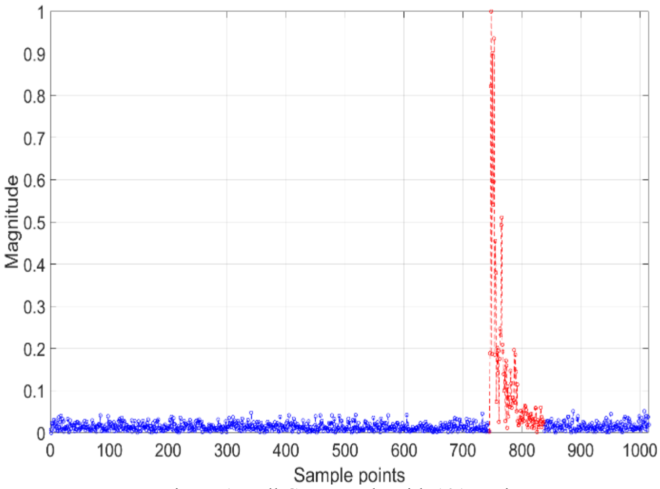

Figure 1. Full CIR sample with 1015 points

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

performance of the direct use of the original CIR points for the detection is evaluated. By using the original data, the calculation complexity can be reduced, since no calculation of the features is needed. Furthermore, a new approach is proposed based on the convolution. Based on the rise time to the maximal value after the convolution of the standard LOS CIR and the CIR measurement, the NLOS condition can be detected. The experimental results show that higher detection accuracy can be achieved even with less CIR points for both feature and original CIR points based SVM approaches. The advantage of choosing less CIR points could be easily explained: if six BSs are used for localization, six range measurements and six CIR samples have to be used for the calculation. There are 1015 times more data contained in each CIR sample than in each range measurements in the used hardware for this paper. If all the data in CIR are used, on-line calculation is almost impossible. Thus it is very important to select only the minimal necessary data in each CIR for the further computation. Additionally, the new approach with convolution has also good NLOS identification accuracy. The rest of the paper is organized as follows: section II describes the used hardware, the CIR samples and the testing environment. In section III the feature and original CIR points based dataset are presented. The basic principle of the SVM algorithm is shown. Furthermore, a new convolution based approach is proposed in this section. Section IV presents the field test results with the mentioned three different approaches, followed by the conclusions in section V.

II. HARDWARE DESCRIPTION AND TESTING ENVIRONMENT

One BS and one MS, which contain the decawave chips, are used for the testing. The distance between BS and MS is measured by the two way ranging method. During the test, the full CIR sample (as shown in Fig.1) is collected. Besides the CIR data, the BS also provide the first path CIR start index, which indicates the real received signal starts. Based on our observation, the 10 points before the index and 110 points after the index describe the true character of the received impulse signal. Thus, these 120 points are taken out for further calculation.

In order to build the training data for SVM, the true distance is also recorded. It is measured by the Bosch GLM 100 C Professional laser measure. Metal, water, people, concrete column etc. are used to block the signal propagation between BS and MS. During the testing, several outliers are found even under LOS. These outliers have to be selected out. The training data is divided into two classes: the LOS class and the NLOS class. In the LOS class, all the data are measured under LOS and the difference between the range measurements from UWB and the true distance is smaller than 25 cm. In the NLOS class, the data are collected with different blocks and the calculated range error is bigger than 25 cm. The testing data is collected in the same way.

During the testing in the office building, the measurements under LOS or with blocks (metal, people, water, woods, concrete etc.) are recorded. The map of the office is shown in Fig.2. The typical CIR samples under LOS and with people or water block with 120 CIR points are presented in Fig.3. As seen in Fig.3, the CIR under LOS and NLOS has different characters, for instance the rise time to the maximal amplitude, the energy for the first path etc. Fig. 4 shows the CIR sample with much less points (30 points). Different characters could still be observed with less points, which proves NLOS identification with less points is also possible. Overall, more than thirty thousand CIR samples are collected at different places within the yellow area (Fig.2) in the office. These collected data are divided into two groups: LOS and NLOS group. They are used as training dataset for SVM. Another thirty thousand CIR samples are recorded within the blue area for testing. The testing results are shown in section IV.

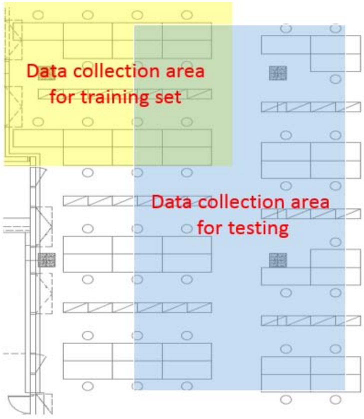

Figure 2. Map of the office with the data collection area

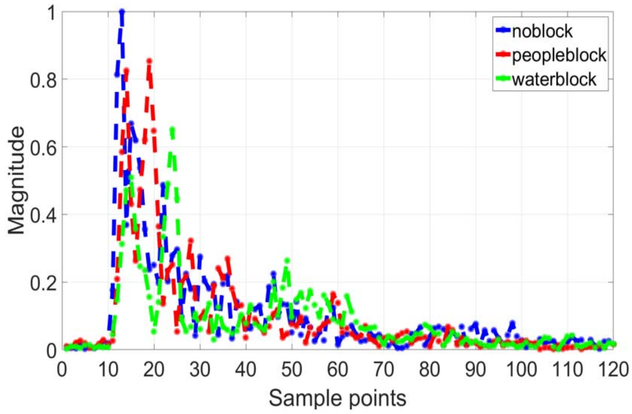

Figure 3. Typical CIR sample under LOS, people and water block with 120 CIR points

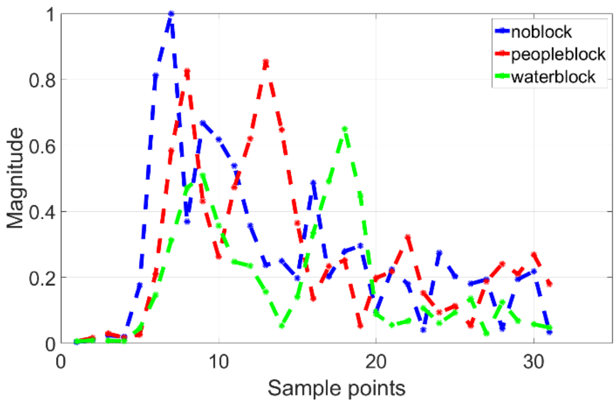

Figure 4. Typical CIR sample under LOS, people and water block with 30 CIR points

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

III. NLOS Identification Approach

The sum of all the received pulses can be used to describe the channel impulse response $c(t)$:

$$c(t) = \sum_{k=1}^{K} a_k \delta(t - \tau_k) \tag{1}$$

Where K is the total number of the multipath components, a_k and τ_k represent the amplitude and time delay of the k^{th} arrived path. τ_1 is the arrival time of the first arrived path. In LOS condition, the first path in CIR is the strongest, while in NLOS condition the first arrived path is attenuated and might not be the strongest. The delay spread of LOS/NLOS condition is also different. Based on those differences, the NLOS condition could be identified. Three different methods are discussed here. Firstly, 8 different features could be extracted from the CIR and used as dataset for SVM. In the second approach, all CIR points are directly used as feature for SVM. These datasets used for detection, together with the SVM algorithm are explained in this section. Lastly, the new proposed approach based on the convolution of the standard CIR under LOS and the measurements for NLOS condition detection is explained. For the first two approaches, in order to decide the needed minimal amount of the CIR points for accurate detection, different sizes of the CIR points are used. For the third approach, however, the minimal needed amount of the data in convolution could be easily determined.

A. *Feature Based Dataset*

Based on the observation of the difference between the LOS and NLOS CIR samples, 8 features can be extracted:

1) Received signal energy

$$E = \int_{-\infty}^{+\infty} |c(t)|^2 dt \tag{2}$$

2) Maximal amplitude

$$c_{max} = \max c(t) \tag{3}$$

3) Rise time to the maximal amplitude: T_{rise}

The rise time can be represented as the index of the maximal amplitude in the CIR.

4) Standard deviation: σ

5) Mean excess delay

$$\tau_m = \frac{\int_{-\infty}^{+\infty} t|c(t)|^2 dt}{E} \tag{4}$$

6) Root-mean-square (RMS) delay spread

$$\tau_{RMS} = \sqrt{\frac{\int_{-\infty}^{+\infty} (t - \tau_m)^2 |c(t)|^2 dt}{E}} \tag{5}$$

7) Kurtosis

$$\kappa = \frac{E[(|c(t)| - \mu)^4]}{\sigma^4} \tag{6}$$

Where μ denotes the mean of $c(t)$.

8) Skewness

$$\gamma = \frac{E[(|c(t)| - \mu)^3]}{\sigma^3} \tag{7}$$

The reasons for selecting these features can be explained as following:

i) The signals are more attenuated or even blocked under NLOS condition. Thus the energy and maximal amplitude of the received impulses are smaller. Due to the attenuation, the first path might not be the strongest path, so the rise time to the maximal amplitude takes longer.

ii) Because the first path is attenuated due to the block, the amplitude of the first path becomes smaller. Because of the block, more reflected, diffracted, scattered later received impulses are added together after the first path, which could enhance the amplitude after the first path. Thus the data points should be closer to the mean under NLOS, which leads to a lower standard deviation. The rest features are selected due to the different energy dispersion in CIR under NLOS and under LOS.

The combination of different features or even one single feature could be used for the NLOS identification. The cumulative distribution function (CDF) could be used to check how good a single feature for the identification of the NLOS/LOS condition is. The CDF of the features under LOS, metal, people, water block are illustrated in Fig 5. These

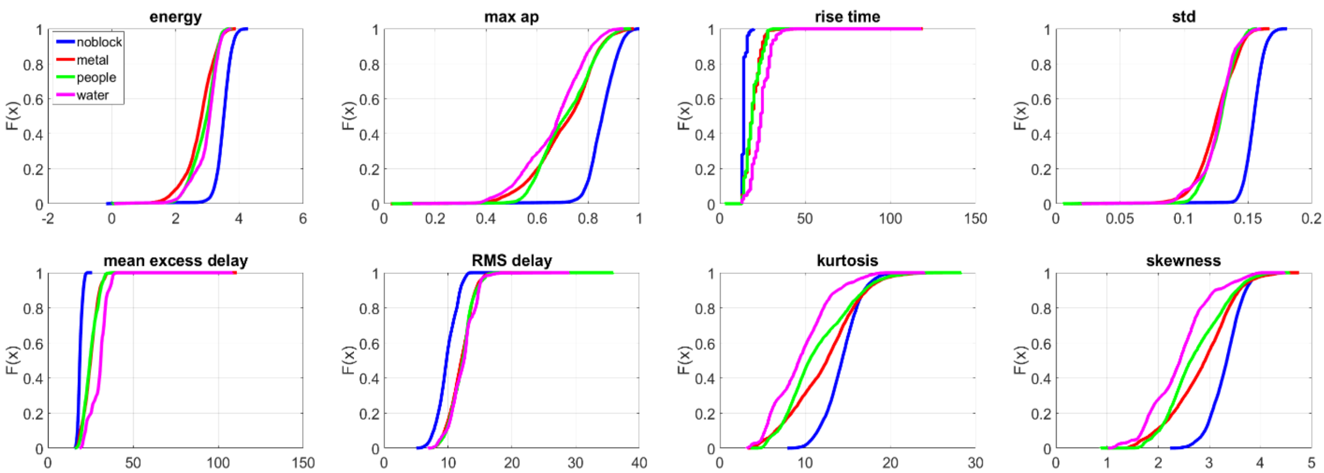

Figure 5. Cumulative distribution function of the features with 120 CIR points

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

features are calculated with the selected 120 CIR points. As expected, more under LOS CIR samples have larger energy, kurtosis and skewness, bigger maximal amplitude and standard deviation, shorter rise time, smaller mean excess delay and RMS delay compared to these under NLOS. 80% of the under LOS CIR samples has an energy value around 3.5, while only 8% of the under NLOS CIR samples has the same energy level, which means based on the energy alone, 80% NLOS CIR samples could be identified with only 8% LOS CIR sample detected wrongly as NLOS. Similarly to the energy, most of the mean excess delay and standard deviation calculated by the LOS samples have a different value than these computed by the NLOS samples. However, the kurtosis and skewness do not show so much significant difference as the energy.

If the size of the CIR points is changed, the value of the corresponding features is also different. As mentioned before, the size of the used CIR points has significant influence on the delay. These features calculated with 30, 40…100 and 110 points in the 120 CIR points are analyzed. Fig. 6 shows the CDF of these features computed with the first 30 points. Based on our observation, as the size becomes smaller, the difference between some of the features under LOS and NLOS has the trend to become smaller. The maximal amplitude remains more or less the same. The difference of some features under LOS and NLOS changes irregularly with the change of the used size of the CIR points.

The feature with the most significant difference is used first for the identification. Then this feature is combined with the second most significant one, till the least significant one. Based on the difference of the feature under LOS and NLOS with different used CIR points, for instance, as shown in Fig. 5 and Fig. 6, at first only the energy is used to build the dataset. Then the mean excess delay, rise time, standard deviation, maximal amplitude, RMS delay, kurtosis and skewness are added to the data set sequentially.

B. *Original CIR Points Based Dataset*

In the feature based dataset, the features have to be calculated first. An alternative option is to build the dataset with the original CIR points without any calculation. The dimension of the dataset is the number of used CIR points.

C. *Support Vector Machine*

As one of the most popular supervised learning models, support vector machine (SVM) is widely used for classification. SVM can be used for linear and non-linear classification. If ideal separation of two classes is not possible, the slack variables are introduced to tolerate the misclassifications, so that the rest data could still be optimally classified. Principally, the linear classification is realized by solving the following optimization equations:

$$\min_{\omega,b,\xi} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\xi_i \tag{8}$$
$$s.t.\ y_i(\omega^T x_i + b) \geq 1 - \xi_i,$$
$$\xi_i \geq 0, i = 1,2,\cdots,n$$

Where the slack variables ξ_i is used as a measure of the misclassification errors. C is the regularization parameter to control the toleration level of the misclassification in the training example.

To solve a non-linear classification, the kernel function has to be used. More details about the SVM algorithm can be found in [7].

After the collection of the training data, the above mentioned two different datasets are built up. The SVM algorithm is used to train the datasets.

D. *Convolution Based Approach*

Convolution is a mathematical operation to combine two functions to produce a third function. The produced function could be understood as the overlap area of one function as it moves over the second function. In our case, if the under LOS CIR is used as one function, after convolution of a second under LOS CIR, the rise time to the maximal amplitude should be

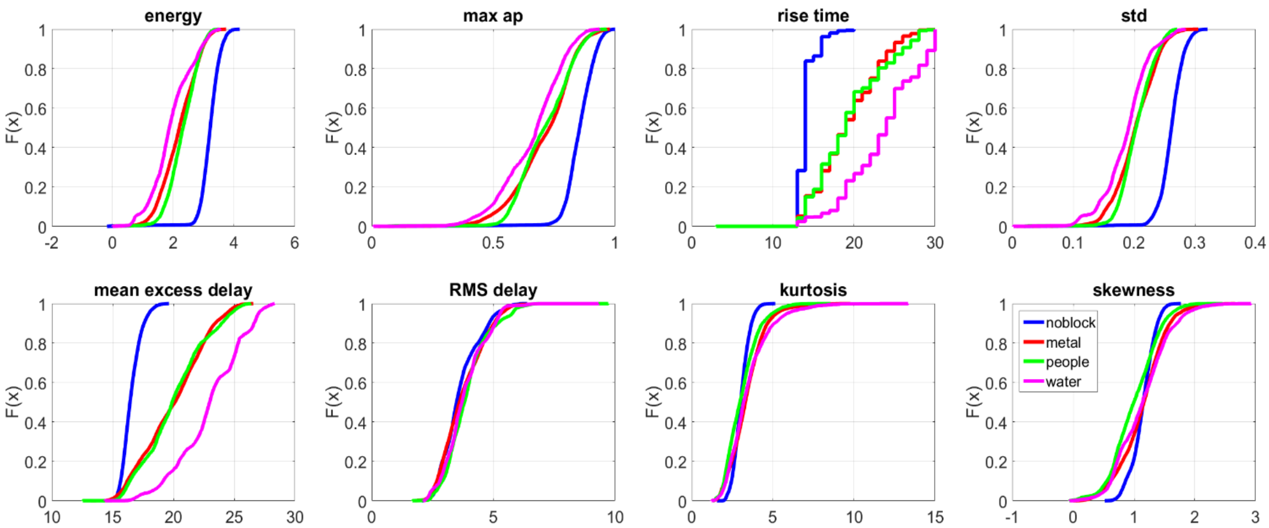

Figure 6. Cumulative distribution function of the features with only the first 30 CIR points

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

shorter than after the convolution of a under NLOS CIR. This is because the energy of the LOS CIR is mainly in the beginning. During convolution, the overlap area reaches the biggest value earlier compared to the NLOS CIR. Fig. 7 shows the signal after the combination of the LOS CIR and the LOS/waterblock/peopleblock CIR with convolution. The rise time of the convolution with two LOS CIR is shorter. To determine the minimal necessary number of the CIR points, we make sure that the maximal amplitude of the signal after convolution can be properly calculated. Based on our observation, only 30 points in CIR are needed. The advantage of this approach is that no big data collection is needed and the calculation is simple. Only a few LOS CIR samples have to be collected at first. Then the average of these signals is calculated and used as the first function for the convolution. A proper threshold needs to be determined. After the convolution with the measured CIR, the rise time is compared to the threshold. If the rise time is shorter than the threshold, the measured CIR is under LOS, otherwise it is under NLOS.

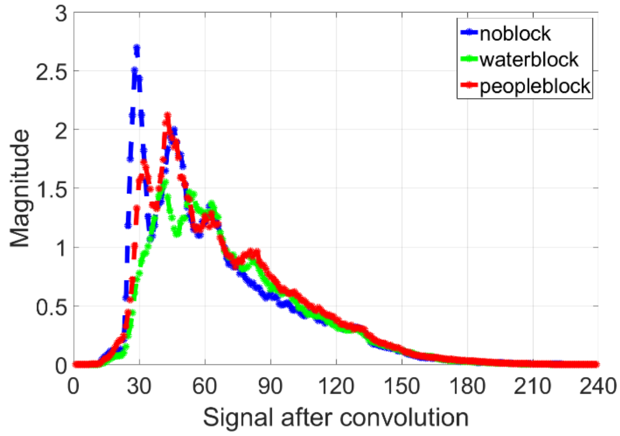

Figure 7. Convolution signal the combination of the LOS CIR and the LOS/waterblock/peopleblock CIR

IV. Field Test Results

The data are collected in the office building as discussed in section II. The SVM approach is used to train the feature based dataset and the original CIR points based dataset. Different datasets are built up with different numbers of the CIR points, starting from 30 points, 40 points to 120 points. To evaluate the performance of the SVM algorithm with different testing datasets, two parameters are used here. The first one is the true accurate LOS detection rate P_t and the second one is the true accurate NLOS detection rate P_f.

$$P_t = \frac{TL}{TL + FL} \tag{9}$$

$$P_f = \frac{TN}{TN + FN} \tag{10}$$

Where TL is the number of correctly detected LOS CIR, FL is the false detected LOS CIR, which means these measurements are actually under LOS, but they are identified as NLOS. TN is the number of correctly detected NLOS CIR, FN is the false detected NLOS CIR. Table 1. shows the final testing results. The first row indicates how many points are used in each CIR sample. The first column indicates which feature is used in the dataset. 1 represents energy, which means only the energy feature is used for building the dataset. 2,3… 8 represent separately the mean excess delay, rise time, standard deviation, maximal amplitude, RMS delay, kurtosis and skewness. CIR data SVM indicates the original CIR points. The dataset built by the energy, mean excess delay and rise time with 50 points shows the best performance among the feature based approaches. Although the NLOS identification with energy has over 90% accuracy, average over 50% LOS measurements are identified as NLOS. As the used CIR points becomes more, the accuracy does not always get better. Overall, the original CIR points with 30 points based dataset shows the best results. There is no specific relationship between the detection accuracy and the used features could be observed from the Table 1, neither is the used CIR points. However, it is obvious that as the increase of the used features or used points in CIR sample, the results

Table 1. Testing results with three different approaches

P_t%/ P_f%	30 points	40 points	50 points	60 points	70 points	80 points	90 points	100 points	110 points	120 points
1	43.14/ 91.35	50/ 91.57	55.14/ 91.46	54.86/ 91.68	53.14/ 91.68	53.71/ 91.68	54.86/ 91.68	54.29/ 91.68	55.43/ 91.68	54.57/ 91.68
1,2	80/ 89.3	83.14/ 89.73	78.86/ 90.38	62.86/ 91.03	78/ 90.27	83.71/ 89.63	75.14/ 90.49	78.29/ 90.27	77.71/ 90.38	84.29/ 89.84
1,2,3	79.14/ 90.17	82/ 90.38	91/ 90.38	86/ 90.27	83.43/ 90.38	83.43/ 90.38	86.57/ 90.38	86/ 90.38	85.14/ 90.38	86.57/ 90.38
…	71.71/ 91.03	70.86/ 91.24	82.29/ 90.6	74.57/ 90.81	86/ 90.38	72.57/ 91.57	74.29/ 91.57	72.86/ 91.24	78.29/ 91.68	70.57/ 91.46
…	79.14/ 89.95	86.86/ 89.95	86.29/ 90.06	88.86/ 89.95	71.43/ 91.14	73.14/ 91.03	90/ 89.84	86.86/ 90.06	86.86/ 90.17	88.86/ 90.17
…	86.86/ 90.17	80.57/ 89.63	77.43/ 90.17	87.71/ 89.84	86.57/ 89.73	82.57/ 89.95	87.14/ 89.63	84.29/ 90.17	78.57/ 90.38	88.57/ 90.06
1,2…7	80.86/ 90.38	89.71/ 89.63	87.71/ 89.95	87.43/ 89.84	87.43/ 89.73	84/ 89.84	88/ 89.3	88.86/ 89.2	88.29/ 89.84	88.57/ 89.84
1,2…8	79.43/ 90.27	88.57/ 89.3	85.14/ 89.52	78.57/ 90.17	87.71/ 88.87	84.29/ 89.3	81.71/ 90.06	88.86/ 88.98	89.71/ 89.3	86.29/ 89.41
CIR data SVM	91.43/ 91.03	89.57/ 90.49	85.71/ 91.03	88/ 89.95	82.86/ 89.73	87.43/ 89.09	85.71/ 88.33	83.14/ 88.33	83.14/ 88.33	83.14/ 88.33
Convolution based	88/ 89.8	—	—	—	—	—	—	—	—	—

978-1-5386-5602-0/18/$31.00 ©2018 IEEE

could even become worse. Choosing the proper size of the used CIR points could not only reduce the delay but also improve the identification accuracy. Compared to the feature based approach, the CIR data SVM requires no extra calculation. Although the accuracy of the convolution based approach is not as attractive as the first two approach. This approach does not require lots of data collection to build the training data. As discussed before, the best accuracy is achieved if the maximal amplitude of the signal after convolution can be properly calculated. In our case, 30 points in CIR are needed here. The accuracy of this approach could be further improved if more features are extracted from the signal after convolution.

V. Conclusion

In this paper, the necessary size of the used CIR points for NLOS identification is discussed. Many CIR samples are collected and used as training data and testing data for three different approaches. The final results show that with proper size, the accuracy could even be improved. For the convolution based approach, only a few LOS CIR need to be recorded. Although the approach is relatively simple, the detection accuracy has comparable results to the other two approaches.

References

[1] J. Schroeder, S. Galler, K. Kyamakya, and K. Jobmann, “NLOS detection algorithms for Ultra-Wideband localization,” in *2007 4th Workshop on Positioning, Navigation and Communication*, Hannover, Germany, 2007, pp. 159–166.

[2] S. Marano, W. Gifford, H. Wymeersch, and M. Win, “NLOS identification and mitigation for localization based on UWB experimental data,” *IEEE J. Select. Areas Commun.*, vol. 28, no. 7, pp. 1026–1035, 2010.

[3] K. Gururaj, A. K. Rajendra, Y. Song, C. L. Law, and G. Cai, “Real-time identification of NLOS range measurements for enhanced UWB localization,” in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sapporo, 2017, pp. 1–7.

[4] K. Wen, K. Yu, and Y. Li, “NLOS identification and compensation for UWB ranging based on obstruction classification,” in *2017 25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 2017, pp. 2704–2708.

[5] Z. Zeng, S. Liu, and L. Wang, “UWB/IMU integration approach with NLOS identification and mitigation,” in *2018 52nd Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, 2018, pp. 1–6.

[6] Z. Zeng, S. Liu, W. Wang, and L. Wang, “Infrastructure-free indoor pedestrian tracking based on foot mounted UWB/IMU sensor fusion,” in *2017 11th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Surfers Paradise, QLD, 2017, pp. 1–7.

[7] Asa Ben-Hur and Jason Weston. A user’s guide to support vector machines. Methods in Molecular Biology, 609:223–239, 2010.

[8] Landolsi, Mohamed Adnan and Ali F. Almutairi. “Reliable Line-of-Sight and Non-Line-of-Sight Propagation Channel Identification in Ultra-Wideband Wireless Networks.” (2016).

[9] Yang Xiaofeng, Zhao Feng, Chen Tiejun. (2018). NLOS Identification for UWB Localization Based on Import Vector Machine. AEU - International Journal of Electronics and Communications. 87. 10.1016/j.aeue.2018.02.003.

[10] J. Khodjaev, Y. Park, and A. S. Malik, “Survey of NLOS identification and error mitigation problems in UWB-based positioning algorithms for dense environments,” Annals Telecommun., 2009.

[11] Z. Zeng, S. Liu, and L. Wang, “A novel NLOS mitigation approach for TDOA based on IMU measurements,” in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, Barcelona, Spain, 2018, pp. 1–6.