NLOS Satellite Detection Using Fish-Eye Camera and Semantic Segmentation for Improving GNSS Positioning Accuracy in Urban Area

Kenta Horide\textsuperscript{1}, Akihiro Yoshida\textsuperscript{1}, Reo Hirata\textsuperscript{1}, Yukhiro Kubo\textsuperscript{1} and Yoshiharu Koya\textsuperscript{2}

\textsuperscript{1} Department of Electrical and Electronic Engineering, Ritsumeikan University  
E-mail: ykubo@se.ritsumei.ac.jp  
\textsuperscript{2} Kobe City College of Technology

Abstract

In this paper, for the GNSS (Global Navigation Satellite System) positioning, we propose a method to distinguish non-line-of-sight (NLOS) signals from line-of-sight (LOS) signals by utilizing the satellite geometry and the image data of the fish-eye view of the zenith direction at the antenna position. The NLOS signal is the diffracted signal by buildings or structures, and it greatly degrades the positioning accuracy. By applying the proposed method, we can exclude the NLOS signal and can improve the positioning results in urban areas. In recent years, image recognition using deep learning has developed rapidly. We use the semantic segmentation method using deep learning for segmentation distinguishing sky area from obstacle area.

1 Introduction

GNSS (Global Navigation Satellite System) is widely used because it can perform highly accurate positioning. However, the accuracy can be easily degraded in places where radio waves are difficult to reach such as urban areas and mountainous areas and so on.

Fig. 1: Multipath Signal

Fig. 1 shows the multipath signals in urban areas. GNSS signals tend to reflect on glass, metals, wet surfaces. In urban areas, in addition to signals received directly from satellites, the GNSS receiver may also receive signals reflected on buildings, walls, grounds, etc. This signal is called “multipath signal”. Especially multipath signals received from NLOS (Non Line Of Sight) satellites, which are obstructed by obstacles and can not be directly received, are not perfectly corrected even if multipath mitigation techniques are applied [1]. In such case, the positioning accuracy in urban areas is greatly degraded. In this research, we aim to improve the positioning accuracy by estimating the NLOS satellites and excluding them from the positioning calculation.

2 Methods

Fig. 2 shows a flow chart of the proposed method.

Fig. 2: Flow Chart

First, we set omnidirectional fish-eye camera with optical axis aligned to the zenith and, divide the sky and obstacles into regions by using image processing for the images taken by the fish-eye camera. Next, we calculate the elevation angle and the azimuth angle using the satellite position and the approximated receiver...
position calculated from the data obtained from the satellite and plot the satellite position on the image after region division while taking the camera projection method into consideration. Finally, the satellite in the obstacle region is estimated as the NLOS satellite and excluded from the positioning calculation. Fig. 3 shows an example of the detected NLOS Satellites. The left figure in Fig. 3 shows the satellites’ positions with satellite identifier such as “G32”. In the right figure, satellites G12, G10, R7 and G29 are detected as the NLOS satellites.

![Satellites Positions](image)

Fig. 3: NLOS Satellite Detection

### 3 Image Processing

#### 3.1 FCN

In this research, FCN (Fully Convolution Networks) [2] which is a method of semantic segmentation is used for discrimination of sky area. FCN is an extension of the classification method CNN (Convolutional Neural Network) to the pixel level. This method can be efficiently applied when there exist disturbances such as cloud movement and car body shaking, and it is possible to discriminate the sky region even in such situations, so we consider that the proposed method can discriminate the sky region more accurately in various circumstances. The model applied in this research is fcn8s with the highest accuracy among the three methods described in the paper [2].

![fcn8s](image)

Fig. 4: fcn8s

FCN has a structure in which convolution layer and pooling layer are alternately arranged. Formula of convolution layer [3] is as follows:

$$\hat{z}_{m,n,k}^{l+1} = \sum_{p,q,k} \omega_{p,q,k,k'} z_{m+p, n+q, k}^{l} + b_{k'}^{l+1}$$

where \((m, n, k)\) is the feature map \((z_l)\) size at \(l\) layer, \((p, q, k)\) is the kernel \((w)\) size, \(b\) is the bias and \(k'\) is the number of kernel and bias channel. Formula of pooling layer [3] is as follows:

$$z_{m,n,k}^{l+1} = \max_{(i,j) \in R_{m,n}} z_{i,j,k}^{l}$$

where \((m, n, k)\) is the output \((z_{l+1})\) size and \(R_{m,n}\) is the rectangular area corresponding to the output destination \((m, n)\). We update the kernel and bias values with learning so that the FCN’s output matches the teacher data. However, it is necessary to set the size and number of kernels arbitrarily. Table 1 shows the model and the parameters used in this research. We used the fcn8s, and the parameters in the model are those of VGG16 [4] which is a representative model of the classification in image recognition.

| layer     | kernel | bias  |
|-----------|--------|-------|
| input     | (3, 3, 3)*64 | -     |
| conv1-1 + ReLU | (3, 3, 64)*64 | (1, 1, 64) |
| conv1-2 + ReLU | (3, 3, 64)*64 | (1, 1, 64) |
| pooling1  | -      | -     |
| conv2-1 + ReLU | (3, 3, 128)*64 | (1, 1, 128) |
| conv2-2 + ReLU | (3, 3, 128)*128 | (1, 1, 128) |
| pooling2  | -      | -     |
| conv3-1 + ReLU | (3, 3, 256)*256 | (1, 1, 256) |
| conv3-2 + ReLU | (3, 3, 256)*256 | (1, 1, 256) |
| conv3-3 + ReLU | (3, 3, 256)*256 | (1, 1, 256) |
| pooling3  | -      | (1)   |
| conv4-1 + ReLU | (3, 3, 512)*512 | (1, 1, 512) |
| conv4-2 + ReLU | (3, 3, 512)*512 | (1, 1, 512) |
| conv4-3 + ReLU | (3, 3, 512)*512 | (1, 1, 512) |
| pooling4  | -      | (2)   |
| conv5-1 + ReLU | (3, 3, 512)*512 | (1, 1, 512) |
| conv5-2 + ReLU | (3, 3, 512)*512 | (1, 1, 512) |
| conv5-3 + ReLU | (3, 3, 512)*512 | (1, 1, 512) |
| pooling5  | -      | -     |
| conv6 + ReLU | (3, 3, 512)*4096 | (1, 1, 4096) |
| conv7 + ReLU | (3, 3, 4096)*4096 | (1, 1, 4096) |

(1) \(\rightarrow\) Score1 | (1, 1, 256)*class | (1, 1, class) |
(2) \(\rightarrow\) Score2 | (1, 1, 512)*class | (1, 1, class) |
(3) \(\rightarrow\) Score3 | (1, 1, 4096)*class | (1, 1, class) |

We create teacher data using index color with reference to Visual Object Classes Challenge 2012 [5]. Fig. 5 shows an example of teacher data. The right figure in Fig. 5 shows that the red area is the sky area, and the black area is the obstacle area.
3.2 Projective Transformation

In this research, it is necessary to consider the projection method of the camera when we plot the satellite on the image. The lens of the fisheye camera is an orthographic projection; therefore, we need to transform from equidistant projection to orthogonal projection for plotting. Equidistant projection has a characteristic that the distance from the center and the incident angle are proportional. Fig. 6 shows each projection method.

\[ x_{equ} = 1 - \frac{2}{\pi} \theta \quad (0 \leq x_{equ} \leq 1) \]  
\[ x_{ort} = \cos(\theta) \quad (0 \leq x_{ort} \leq 1) \]  
\[ x_{ort} = \cos \left( \frac{\pi}{2} (1 - x_{equ}) \right) \]  

where \( x_{ort} \) and \( x_{equ} \) are expressions of orthographic projection and equidistance projection of the fisheye camera, and \( \theta \) [rad] is the elevation angle from the receiver (0 \( \leq \theta \leq \frac{\pi}{2} \)). Eq. (5) gives the ratio of the distance to the location after projective transformation from the center scaled with a minimum value of 0 and a maximum value of 1. Fig. 7 shows how to calculate the coordinates. We change the scale so that the maximum value becomes the radius of the shootable range of the fisheye camera, and we obtain the coordinates from the azimuth. Coordinate transformation is given by using Eq. (5) as follows:

\[ u = r \cdot x_{ort} \cos(\phi) \]  
\[ v = r \cdot x_{ort} \sin(\phi) \]  

where \( u \) and \( v \) are uv coordinates with the center of the image as the origin, \( r \) is radius of the circle in the shootable range of the fisheye camera, and \( \phi \) [rad] is azimuth of the satellite from the receiver (0 \( \leq \phi < 2\pi \)).

4 Experimental Results

4.1 Semantic Segmentation

In this experiment, the model is learned by using the sky image and teacher data, and we examine the output accuracy of the inference result. Table 2 shows the experimental conditions.

| Location            | BKC, Ritsumeikan Univ. |
|---------------------|------------------------|
| Model               | fcn8s                  |
| Model Parameter     | VGG16                  |
| Epoch               | 100                    |
| Camera              | α5100 (SONY)           |
| Angle of view       | 180                    |
| Projection method   | Orthographic projection|

BKC (Biwako Kusatsu Campus) is one of the campuses of Ritsumeikan Univ. We prepared 4 cloudy images taken on August 07, 2018 and 6 sunny images taken on August 22, 2018, and expanded these data. We rotated and reversed the images at random 25 times for data augmentation. We set the number of epoch to 100 so that it can be learned sufficiently. Fig. 8, 9, and 10 show the inference result. The left figures in Fig. 8, 9, and 10 show the unlearned images, and the right figures show
the inference results. The white area is the sky area, and the black area is the obstacle area.

From Fig. 8 and 9, it can be seen that we can infer the sky area with high accuracy. Fig. 10 is an example of failure due to insufficient teacher data of windows reflecting the sky and the sky that pass through the windows of building. It is considered that the accuracy can be improved by increasing teacher data.

4.2 Positioning

In this experiment, we estimate NLOS satellites using inference images obtained by semantic segmentation and exclude them from the positioning calculation. Since the purpose of this experiment is to evaluate the accuracy improvement when excluding NLOS satellites, we use single point positioning and least squares estimation which is the simplest positioning method. Single point positioning is a positioning technique that receives a signal from a satellite with a single receiver and estimates the coordinates of the receiver from the geometric distance between each satellite and the receiver. The geometric distance can be measured from pseudo distance data called C/A code pseudo range and the broadcast ephemeris [6]. Table 3 shows the experimental condition.

Table 3: Experimental Condition (Positioning)

| Date             | October.09, 2018                  |
|------------------|-----------------------------------|
| GPS-Time         | 08:49:30 ~ 08:57:00               |
| Location         | BKC, Ritsumeikan Univ.            |
| Receiver         | FLEX6-TAQ-B0G-TTN (NovAtel)      |
| Antenna          | GPS-703-GGG (NovAtel)             |
| Epoch interval   | 1 [s]                            |
| Used Satellites  | GPS                              |
| Positioning method | Single Point Positioning         |
| Estimation method | Least squares method             |

We acquired the data of 450 epochs at 1 second interval in a stationary state with the FLEX6 receiver made by NovAtel and GPS-703 antenna made by NovAtel. Fig. 11 shows the experiment location and antenna position. In order to reproduce the environment in the urban area, the experiment was conducted on the road nearby BIO LINK which is a high-rise building in Ritsumeikan University BKC.

Fig. 8: Inference Result

Fig. 9: Inference Result (Buildings and Trees)

Fig. 10: Inference Result (Lack of Teacher Data)

Fig. 11: Experiment Location
Fig. 12 shows the inference result of the image taken at antenna position. It can be seen that we can infer the sky area with high accuracy.

![Fig. 12: Inference Result (Antena-Position)](image1)

Fig. 13 shows the plots of satellite constellation before and after the projective transformation. It can be seen that G5 and G12 satellites are plotted in the obstacle area. Since the satellite constellation calculated from the elevation angle and the azimuth angle have characteristic of an equidistant projection, it is converted into an orthogonal projection by Eqs. (6) and (7) which is a characteristic of the fisheye lens. We inverted the image horizontally in order to match east and west because we took the image in the zenith direction. As a result, the satellite plotted in the obstacle area is estimated as an NLOS satellite. In this experiment, when the NLOS satellite is excluded and the number of satellites used becomes less than 5, we excluded NLOS satellites with low elevation so that there are 5 used satellites instead of excluding all NLOS satellites, considering that the positioning accuracy sharply decreases when there are few satellites used.

![Fig. 13: Projective Transformation](image2)

Fig. 14 shows the positioning results plotted on Google Maps.

![Fig. 14: Positioning Result](image3)

The yellow pin is the position obtained by the relative positioning method [6] and it is assumed to be true position in evaluating the positioning results in this experiment. The relative positioning is a method that uses two or more receivers and estimates a baseline vector, which is a relative position between each receiver. By using the true value for one of the receiver coordinates, we can perform very accurate positioning. For this reason, we used the reference station data and daily coordinate values (F3 solution) provided by the Geospatial Information Authority of Japan for one receiver [7] [8].

From this result, it can be observed that the coordinates which were scattered in northwest area by the conventional method are gathered around the provisional true value by the proposed method. Fig. 15, 16 and 17 show the ENU error and satellites visibility with each method.

![Fig. 15: ENU Error](image4)
Table 4: Positioning Error

| Direction | Conventional Method (All Satellites) | Proposed Method (Excluded NLOS) |
|-----------|-------------------------------------|---------------------------------|
|           | East                                | East                            |
|           | North                               | North                           |
|           | Upper                               | Upper                           |
| STD [m]   | 1.50                                | 0.83                            |
| RMS [m]   | 1.50                                | 1.19                            |

Table 4 shows the statistics of positioning error. From these results, it can be considered that the positioning accuracy in North and Upper direction is improved by excluding NLOS satellites PRN 5 and 12, whereas in the East direction, there is no great improvement in positioning accuracy. It is the influence that the NLOS satellites exist in the south direction. Therefore, we consider that the pseudo-range data from the NLOS satellites contain large multipath errors, consequently they caused the large position errors in the North-South direction. From the above, we confirmed improvement of positioning accuracy by excluding NLOS satellite in SPP (Single Point Positioning).

5 Conclusion

In this paper, we proposed the NLOS satellite estimation method using a fisheye camera. The FCN used in the sky area segmentation made it possible to infer highly accurate despite learning data a little. It is expected that more accurate inference can be performed by increasing learning data. In addition, we could improve the positioning accuracy by excluding NLOS satellite in SPP (Single Point Positioning). Since there is a problem that positioning can not be performed when the used satellite has not reached the required minimum number of satellites when excluding NLOS satellites, we will also consider not only excluding NLOS satellites but also methods to correct them in the future work.

References

[1] Mark Petovello, “GNSS Solutions: Multipath vs. NLOS signals”, Inside GNSS, November / December, pp. 40-44, 2013.
[2] Jonathan Long, Evan Shelhamer and Trevor Darrell, “Fully Convolution Networks for Semantic Segmentation”, Computer Vision and Pattern Recognition (CVPR) 2015 paper, pp. 3431-3440.
[3] Tatsuya Harada, “Image Recognition”, Koudan-sya, 2017 (in Japanese).
[4] Karen Simonyan and Andrew Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”, CoRR, abs/1409.1556, 2014 (from : http://arxiv.org/abs/1409.1556).
[5] Visual Object Classes Challenge 2012 (VOC2012) (from : http://host.robots.ox.ac.uk/pascal/VOC/voc2012/).
[6] Sueo Sugimoto and Ryosuke Sibasaki(eds.), “GPS Handbook”, Asakura, 2010 (in Japanese).
[7] H. Nakagawa, et al, “Development and Validation of GEONET New Analysis Strategy (Version 4)”, Journal of Geographic Survey Institute, Vol. 118, pp. 1-8, 2009
[8] Geodetic Observation Center, “Establishment of the nationwide observation system of 1,200 GPS-based control stations”, Journal of Geographic Survey Institute, Vol. 103, pp. 2-8, 2004