Peer-reviewed version available at *Sensors* **2020**, *20*, 4728; doi:10.3390/s20174728

Article

Skymask Matching Aided Positioning Using Sky-Pointing Fisheye Camera and 3D City Models in Urban Canyons

Max Jwo Lem Lee, Shang Lee, Hoi-Fung Ng, Li-Ta Hsu

Interdisciplinary Division of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China; 16044209d@connect.polyu.hk (M.J.L.L.), 16044131d@connect.polyu.hk (S.L.)

* Correspondence: lt.hsu@polyu.edu.hk

Abstract: 3D-mapping-aided (3DMA) global navigation satellite system (GNSS) positioning that improves positioning performance in dense urban areas has been under development in the recent years, but it still faces many challenges. This paper details a new algorithm that explores the potential of using building boundary for positioning. Rather than applying complex simulations to analyze and correct signal reflections by buildings, the approach utilizes a convolutional neural network to differentiate between the sky and building in a sky-pointing fisheye image. A new skymask matching algorithm is then proposed to match the segmented fisheye images with skymasks generated from a 3D building model. Each matched skymask holds a latitude and longitude coordinate to determine the precise location of the fisheye image. The results are then compared with the conventional GNSS and advanced 3DMA GNSS positioning methods. The aims of the proposed algorithm are to increase positioning and heading accuracy in a rich urban environment.

Keywords: GPS; GNSS; Localization; Navigation; Autonomous Driving; Urban Canyon; Land Application; Cameras; Image segmentation

1. Introduction

Global navigation satellite systems (GNSS) provide geographical longitude and latitude positioning with meter-level accuracy in open areas [1]. This accuracy, however, suffers in dense urban areas because buildings block, reflect, and diffract the signals. These cause errors in satellite positioning and reduces accuracy-in severe cases, the position error could exceed 50 meters [2, 3]. An improvement in the real time-positioning accuracy of low-cost GNSS systems in dense urban areas to within 5m would benefit many different potential applications [4], such as cloud-sourced mobile mapping and object tracking. As such, there is a need for a low-cost positioning device that has great solution availability and accuracy.

To improve the positioning performance in urban environments, researchers have designed different methods to identify then correct or exclude the unhealthy measurements by receiver based GNSS or with extra equipment. The increased availability of open-sourced 3D building models allows utilization of 3D mapping aided (3DMA) GNSS positioning to improve urban positioning [4, 5]. This includes shadow matching, ray tracing and likelihood-based ranging methods. Shadow matching utilizes 3D building models with building geometry to match satellite visibility [1, 6], allowing the exclusion of NLOS measurements [7]. Ray tracing GNSS [8, 9] and Skymask based 3DMA GNSS [10] predicts the signal transmission path based on building geometry to calculate the reflection delay distance and provide correction on the pseudorange of the NLOS signals [8-10]. The likelihood-based ranging method is another pseudorange correction method, in which the conventional GNSS position is corrected via a statistical model [11]. In addition, the satellite visibility can be used to estimate the sky-visibility of the environment the receiver is located at [12]. Another

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

approach towards enhancing positioning accuracy in the urban environment is using extra equipment to collect additional data for positioning. This approach is most suitable for vehicular applications due to weight, space and power usage concerns. Two popular approach includes the use of 3D light detection and ranging (LiDAR) and the usage of fisheye cameras. LiDAR is used to retrieve surrounding buildings or obstacles information, which in turn can be used to perform NLOS classification [13]. With the ability to get building distance, pseudorange can be corrected by the NLOS propagation model [14]. Sky-pointing fisheye cameras are capable of detecting obstacles and buildings in the local environment. When used in conjunction with image processing algorithms [2, 4, 14-17], they allow the exclusion of NLOS satellites from position calculations, improving positioning accuracy [15]. Furthermore, research indicates that if both LiDAR and fisheye cameras are used in conjunction, positioning accuracy in urban areas can be further enhanced [3].

Fisheye cameras are used in autonomous driving, where a front-facing camera is commonly used for lane detection and road sign identification. To facilitate its use in improving GNSS positioning performance, the fisheye camera is pointed towards the zenith, which reduces the probability of capturing uncertainties while allowing the use of the positions of immovable objects such as high-rise buildings. As such, this paper makes use of images captured from a zenith sky-pointing fisheye camera, which are matched to computer-generated boundary skyplot (skymask) to obtain a position and heading. This is similar to the approach employed in the research paper [18]. However, this study differs in several ways. Firstly, in this research, the roll and pitch are fixed with a fisheye camera pointing upward, hence the 3D model can be used to generate skymask ahead of time, while the aforementioned paper utilizes real-time generation of computer-generated images from a 3D model. This study's method uses less computing power, while the [18] study's method allow generation of images better suited to the camera's condition (pitch, roll, etc.). This study also utilizes a Convolutional Neural Network (CNN) to segment images into sky and building classes, while the [18] study uses the Otsu method to differentiate between sky and building. Semantic segmentation was chosen for this study due to its potential to segment additional classes in the future. The proposed matching algorithm can not only provide the position of the user but also the heading angle. These are the main novelties of this research paper.

The proposed skymask matching method integrates GNSS, supervised deep learning, and a matching algorithm. This approach takes advantage of sky-pointing fisheye images and matching it with candidate skymask. The results are then compared with other enhanced methods of positioning. The rest of the paper is organized as follows. The proposed algorithm will be explained in detail in section 2. Section 3 describes the experimentation results. Section 4 contains the concluding remarks.

2. Proposed Skymask Matching Method

An overview of the skymask matching method is shown in Figure 1. This study proposes the calculation of the position solution and heading solution with sky-pointing fisheye image and 3D building model generated skymask. The method is divided into 2 main categories-offline processes and online processes. The offline process is done before the real-time processing. It consist of training the CNN and generating skymasks, which are generated outside of buildings in square grid intersections, with a length and width separation of 2m. In the online process, the user captures a sky pointing fisheye image, with initial position and heading estimated by a conventional GNSS receiver. Then, candidate positions are distributed around the initial positioning in a 50m radius to generate candidate skymasks. Meanwhile, the image is simultaneously segmented using the convolutional neural network to distinguish between building and sky, then converted to a binary format. Then, the binary image is converted from pixel format to azimuth and elevation format for comparison. The image converted skymask is compared to candidate skymasks using GMSD, FMSDD, FMBPD and HD techniques. The techniques were weighted and combined into a score to calculate the likelihood of the candidates. The chosen position and heading solution is determined by the highest combined score.

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

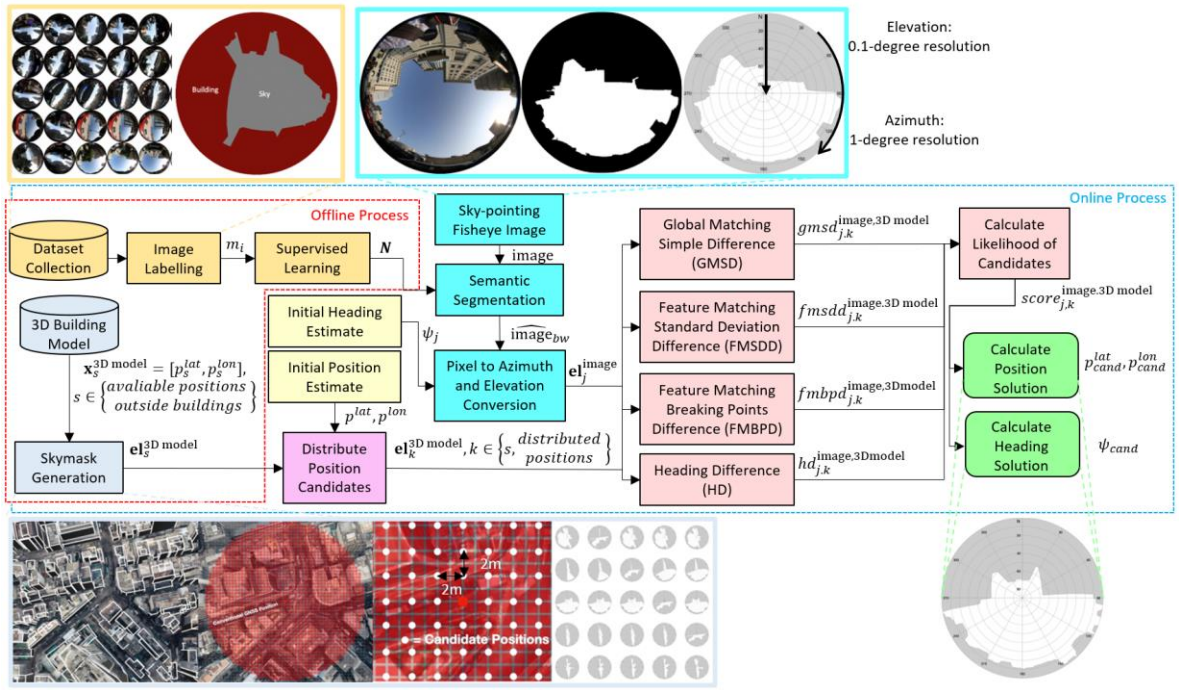


Figure 1. Flowchart of the Proposed Skymask Matching based on Images taken by Sky-pointing Fisheye Camera and Skymask Generated by 3D Building Models.

2.1 Skymask Generation

In this paper, the sky-pointing fisheye image is compared to precompute skymasks that are identical to those utilized in shadow matching and skymask 3DMA GNSS. Skymasks store the highest elevation angle of the surrounding building boundaries (where buildings meet the sky) at each azimuth angle. When used in Skymask Matching, these skymasks are generated in the offline stage and stored in a database, an example is shown above in Figure 1. As the offline stage is done ahead of time, skymask matching requires less computing power and reduces processing times by a factor of 10 when compared to the real-time use of the 3D building model [19]. The precomputed skymasks are then used in the online phase for matching. The details of the generation processes are listed in [15]. Each skymask has a resolution of 360 azimuth angles with 1-degree resolution, and 90 elevation angles with 0.1-degree resolution. They also contain the WGS84 latitude and longitude of the location in decimal degrees format.

2.2 Semantic Segmentation of the Image Taken by Sky-pointing Camera

Pre-processing is needed before matching the fisheye camera image with the skymask in the online stage. There are two main steps for the image processing used: 1) image segmentation by a neural network; 2) image coordinate conversion from pixel to elevation and azimuth angles. These will be discussed in sub-section 2.2.1 and 2.2.2, respectively.

2.2.1 *Convolutional Neural Network (CNN) Training*

Training datasets were collected at deep urban canyons in Hong Kong. The locations were chosen due to its challenging urban topography. The high density of skyscrapers and other tall buildings create steep urban canyons, reducing the accuracy of conventional GNSS positioning. The high concentration of skyscrapers also leads to an abundance of features in a sky pointing image, providing an ideal testing ground for skymask matching.

Over 1200 daytime images were taken using a Canon DSLR camera. Of these, 570+ images were manually labelled and used to train the CNN semantic segmentation network. The dataset is split

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

into three parts: 60% are used as a training set, 20% as the cross-validation set, and the remaining 20% as the test set. The training dataset was labelled manually with Image Labeler application in MATLAB, which is part of the Computer Vision Toolbox [20]. MATLAB's Deep Learning toolbox was also utilized to retrain Resnet50, a 50-layer CNN [21]. The trained CNN is then used for semantic segmentation. In this paper, semantic segmentation is used to differentiate between the sky and buildings only. In the future, however, more classes could be added. This will be discussed in further detail in Section 4.

To further improve the segmentation accuracy, Active Contours is also applied. Active contouring utilizes the Chan-Vese segmentation algorithm [22]. The Chan-Vese method is based on the approach to curve evolution to separate foreground from background based on the means of two regions.

2.2.2 *Image Coordination Transformation: Pixel to GNSS Skyplot*

To match the segmented fisheye image with the pre-computed skymask, the segmented image is first converted into a format that is identical to that of the precomputed skymask, known as the GNSS skyplot format (elevation and azimuth angles). The resolution of elevation and azimuth angle information are identical to the pre-computed skymask in Section 2.1. For a given position in pixel, the calculation to convert into azimuth and elevation angle is the same approach described in [3]. Assuming the optical center of the camera is zenith pointing, each pixel inside the sky view image will be converted to a corresponding azimuth and elevation angle.

$$\theta = \frac{\pi}{2} - el_z^{\text{image}}$$
$$d_{z,\text{pix}} = 2 \cdot f_c \tan\left(\frac{\theta}{2}\right), z \in \{index\ of\ pixels\} \quad (1)$$

To determine the elevation angle for a pixel $(x_{z,\text{pix}}^{\text{image}}, y_{z,\text{pix}}^{\text{image}})$ the focal length (f_c) of the fish-eye camera is needed. Where $d_{z,pix}$, is the pixel distance from the center of the sky-pointing image and correlated with the elevation angle el_z^{image} of the pixel. Given the center of the sky-pointing image in pixel position $(x_{\text{c,pix}}^{\text{image}}, y_{\text{c,pix}}^{\text{image}})$, the azimuth and elevation angle can be expressed as follows:

$$x_{z,\text{pix}}^{\text{image}} = x_{c,\text{pix}}^{\text{image}} + d_{\text{pix}} \cdot \cos\left(az_{\boldsymbol{z}}^{\text{image}}\right)$$
$$y_{z,\text{pix}}^{\text{image}} = y_{c,\text{pix}}^{\text{image}} - d_{\text{pix}} \cdot sin\ \left(az_{\boldsymbol{z}}^{\text{image}}\right) \quad (2)$$

The expression is a quadratic equation that solves the azimuth and elevation angles simultaneously. For practicality, the relationships between the angles and corresponding pixel coordinates is precomputed offline and stored in a database. This means that during the online matching process, the angles information can be retrieved by mapping given pixel coordinates in the pre-compute lookup table, reducing computational load and time.

$$\mathbf{M}_{z,\text{pix}}^{\text{image}} = \left[az_0^{\text{image}}, el_0^{\text{image}} \ldots az_z^{\text{image}}, el_z^{\text{image}}\right] \quad (3)$$

2.3 *Skymask Matching Positioning and Heading Resolution*

In the online stage, the sky-pointing fisheye image converted skymask is rotated 359 times with an increment of 1-degree clockwise to generate 360 skymasks, each with a different heading angle. The image converted skymask is then compared to the skymask generated from each candidate. The matching algorithm will provide two pieces of information: 1) the location of the fisheye camera, and 2) the heading angle of the camera. The target function is to find the candidate skymask with the smallest differences with respect to elevation angle, standard deviation, feature and heading angle.

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

$$\begin{aligned} \mathbf{x}_k &= [p_k^{lat}, p_k^{lon}], k \in \{distributed\ candidates\} \\ \mathbf{el} &= [el_0, \dots, el_{az}, \dots, el_{359}], az = [0°, 359°] \\ \mathbf{el}_k^{\text{3D model}} &= \text{skymask}(\mathbf{x}_k) \\ \mathbf{el}_j^{\text{image}} &= \text{image}(\psi_j), j = [0°, 359°] \end{aligned} \tag{4}$$

There are four methods to compare two skymasks, which are 1) Global Matching Simple Difference (GMSD), 2) Feature Matching Standard Deviation Difference (FMSDD), 3) Feature Matching Breaking Points Difference (FMBPD) and 4) Heading Difference (HD). In total, four set of differences can be obtained to use for the estimation of the likelihood of the distributed candidates.

2.3.1 Global Matching Simple Difference (GMSD)

At each azimuth angle, the difference of elevation angles between the image converted skymask, and the skymask at different candidates based on 3D model is calculated and averaged. The averaged elevation angle difference will be known as the $gmsd_{j.k}^{\text{image,3D model}}$ value.

$$gmsd_{j.k}^{\text{image,3D model}} = \frac{\left|\mathbf{el}_j^{\text{image}} - \mathbf{el}_k^{\text{3D model}}\right|}{360} \tag{5}$$

The average elevation angle difference is obtained by summing the absolute elevation angle difference at each azimuth angle. This difference is summed up and divided by 360 (the number of azimuths) to acquire the average elevation angle difference at a specific heading angle. A large average difference means the candidate skymask has a low probability to be the image converted skymask, whereas a smaller value represents a similar overall average elevation difference, and thus a higher probability of the candidate skymask being similar to the image converted skymask.

2.3.2 Feature Matching Standard Deviation Difference (FMSDD)

The second technique is FMSDD. At each azimuth, the elevation angle is compared to the average ($\overline{\mathbf{el}_j^{\text{image}}}$ and $\overline{\mathbf{el}_k^{\text{3D model}}}$) to calculate its deviation from the mean. Features are defined as a change in elevation angle between two adjacent azimuth points. The standard deviation algorithm is used to measure the features in specified skymask. A small standard deviation indicates little to no features, whereas a large standard deviation indicates an abundance of them. The standard deviation of the candidate skymask is then compared to the standard deviation of the image converted skymask. A smaller difference means higher similarity.

$$fmsdd_{j.k}^{\text{image,3D model}} = \sqrt{\frac{(\mathbf{el}_j^{\text{image}} - \overline{\mathbf{el}_j^{\text{image}}})^2}{360}} - \sqrt{\frac{(\mathbf{el}_k^{\text{3D model}} - \overline{\mathbf{el}_k^{\text{3D model}}})^2}{360}} \tag{6}$$

2.3.3 Feature Matching Breaking Points Difference (FMBPD)

The third one calculates the feature breaking points only. Breaking points are defined as points where the elevation angle suddenly changes 10-degree or more between two adjacent azimuth points, an example of which can be seen in Figure 2. In this secenario, the rest of the data points are treated as being at 0 elevation angle. Algorithm 1 shows the feature identifying process in a skymask.

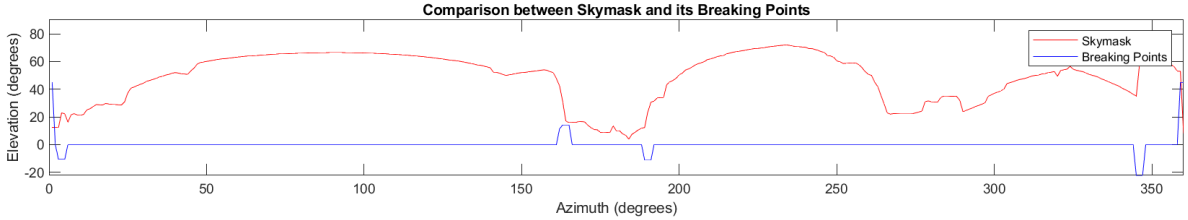


Figure 2. Comparison between Skymask and its Breaking Points

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

Algorithm 1: Determining the Feature Matching Breaking Points in a Skymask

Input: Skymask, **el**
Output: Skymask Breaking Points, **el**′

```
for each azimuth, az in el
  calculate az and az + 1 elevation angle difference, β
  if any(|β| < 10°)
    update el′(az) = 0
  end if
  if any(|β| ≥ 10°)
    update el′(az) and adjacent azimuths to β
  end if
end for each azimuth, az
```

(7)

The breaking points are calculated for both image converted skymask ($\mathbf{el}_j^{\text{image}}$) and candidate skymask ($\mathbf{el}_k^{\text{3D model}}$). By using the extremly distinctive features on a skymask only, a third FMBPD score can be obtained by both skymask with breaking point estimation.

$$fmbpd_{j.k}^{\text{image,3D model}} = \begin{cases} \text{No score} & \left|\mathbf{el}_j^{\text{image}'}\right| + \left|\mathbf{el}_k^{\text{3D model}'}\right| = 0° \\ \mathbf{el}_j^{\text{image}'} - \mathbf{el}_k^{\text{3D model}'} & \left|\mathbf{el}_j^{\text{image}'}\right| + \left|\mathbf{el}_k^{\text{3D model}'}\right| > 0° \end{cases} \tag{8}$$

2.3.4 Heading Difference (HD)

The forth one calculates the heading difference between the rotated image converted skymask and the conventional heading, which is the heading recorded by the Broadcom BCM47755 used in this study. A smaller heading difference from the conventional heading of the image converted skymask when matched with a candidate skymask means they are more similar to one another.

$$\varphi = \psi_{j,k}^{\text{image,3D model}} - \psi_j^{\text{conventional}}$$
$$hd_{j,k}^{\text{image,3D model}} = \begin{cases} \varphi + 360^o & \varphi < -180^o \\ \varphi - 360^o & \varphi > 180^o \\ \varphi & -180^o \le \varphi \le 180^o \end{cases} \tag{9}$$

2.4 *Candidate Scoring*

A higher score is given to the candidate position with a higher similarity between the image converted skymask and the candidate skymask. Gaussian distributions are assumed and used to model the the similarity of the candidate skymasks. In theory, the sky pointing fisheye image taken at the corresponding computer-generated GT skymask should have the smallest difference. Ten sky pointing fisheye images are taken at the corresponding known GTs to calibrate the Gaussian probability distribution function (PDF). A smaller elevation angle difference will obtain a bigger probability value and therefore a higher similarity to the image converted skymask and vice versa. The four differences are used to calculate the corresponding probability value in their respective distributions. The combined likelihood becomes the weightings of each candicate.

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

$$
\begin{aligned}
s_{gmsd,j,k}^{\text{image,3D model}} &= \frac{1}{8*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{gmsd_{j.k}^{\text{image,3D model}}}{8}\right)^2} \\
s_{fmsdd,j,k}^{\text{image,3D model}} &= \frac{1}{15*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{fmsdd_{j.k}^{\text{image,3D model}}}{15}\right)^2} \\
s_{fmbpd,j,k}^{\text{image,3D model}} &= \frac{1}{20*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{fmbpd_{j.k}^{\text{image,3D model}}}{20}\right)^2} \\
s_{hd,j,k}^{\text{image,3D model}} &= \frac{1}{40*\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{hd_{j.k}^{\text{image,3D model}}}{40}\right)^2}
\end{aligned} \tag{10}
$$

$$
s_{combined,j,k}^{\text{image,3D model}} = s_{gmsd,j,k}^{\text{image,3D model}} + s_{fmsdd,j,k}^{\text{image,3D model}} + s_{fmbpd,j,k}^{\text{image,3D model}} + s_{hd,j,k}^{\text{image,3D model}}
$$

Finally, the candidate skymask with the largest combined score will be selected as the chosen candidate skymask. The combined score will then be normalized and rescaled between 0 to 100%, which represents the total score of the candidate. Heading angle offset corresponds to the chosen candidate skymask.

3. Experiments Results and Analysis

3.1 *Experiment Setup*

In this study, the experiment locations were selected within the Tsim Sha Tsui area of Hong Kong as shown in Figure 3. These locations were selected using the following factors, proximity to obstacles, the height and features of nearby buildings, and ability to determine location relative to landmarks both on the ground and by satellite image. Images were then taken at each of the selected ground truths using a digital single-lens reflex (DSLR) camera. The DSLR (Canon 5D Mk III DSLR) with the fisheye-len (8-15 mm f/4L EF Fisheye USM Lens) was used to capture the image and the Galaxy Note 9 Broadcom BCM47755 was used to record the low-cost GNSS solutions and heading. The images taken were manually categoried into 4 distinctive environments, with distinctions specified in Table 1 below. Categorization was based on the frequency of different obstacles, buildings and its features. Four images were chosen, one from each category, to demonstrate the proposed algorithm.

Figure 3. Experiment Locations

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

Table 1 Fisheye Image Categories, and Experiment Locations

Pos #	Category	Environment	Mean of Skymask μ (degree) [25]	Std of Skymask σ (degree) [25]
1	Clean	Few buildings and obstacles visible	18.18	8.44
2	Urban – Distinctive	High rise buildings, distinctive feature visible	52.44	18.04
3	Urban – Complex	High rise buildings, mixed features	56.14	18.84
4	Multiplex	High frequency of trees and other obstacles	36.65	18.29

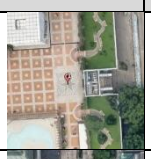
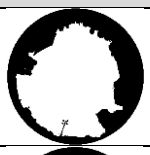
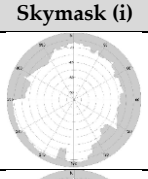
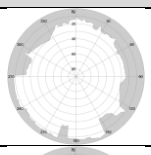
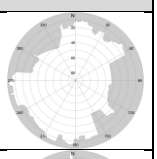
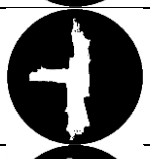
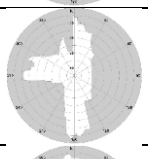
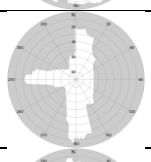
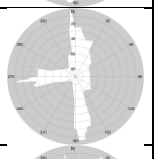
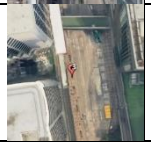
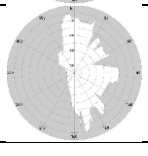
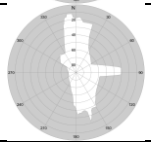
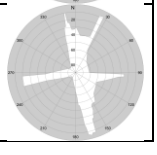
The experimental results are then post-processed and compared to the ground truth and different positioning algorithms, including:

1) Allystar: Commercial GNSS solution by Allystar TAU1302+AGR6303 active antenna.
2) Broadcom: Low-cost convetional GNSS solution by Galaxy note 9 Broadcom BCM47755
3) WLS: Weighted-least-square. (WLS) [23]
4) 3DMA: Integrated solution by 3DMA GNSS algorithm on shadow matching, skymask 3DMA and likelihood-based ranging GNSS. [24]
5) SM: The skymask matching algorithm proposed in this paper.
6) Ground Truth: Data was collected at the landmark location on Google Earth. The accurarcy is within 1-2 meters based on our experience.

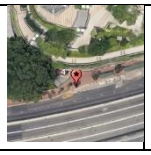
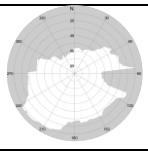
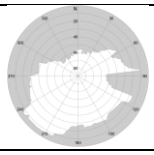
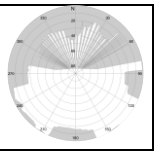
The images used for testing were not used in training the CNN. In the images, the heading of the camera (and thus the fisheye image) faces north at 0 degree in azimuth angle. During the image collection process, the camera was aligned to true north. The north was determined by estimation using *Google Earth* and observations of nearby objects. Slight manual adjustments $\pm 2^{\circ}$ to the heading was made after the images were taken to ensure heading of the image faces north for verification purposes. Manual adjustments were also determined based on Google Earth. It is important to note the conventional heading was also adjusted the same magnitude based to the manual adjustments.

To evaluate the accuracy of the segmentation processing on the images, they were compared to their corresponding hand-labelled counterparts. As seen in Table 2, four images were tested to see if the proposed method provides a consistent segmentation to differentiate between building and sky.

Table 2. Close view of Experiment Locations 1 to 4 and comparison between Neural Network Segmented Images, Labelled Images, their respective skymasks and GT skymasks

Pos #	Location	Fisheye Image	Neural Network Segmented Image	Hand Labelled Image	Neural Network Segmented Image Converted Skymask (i)	Hand Labelled Image Converted Skymask (ii)	3D Model Generated GT Skymask (iii)
1							
2							
3							

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

4							

3.2 Evaluation of the Skymask generated based on Image and 3D building models

Inaccuracies were largely the result of two sources; 1) overexposure of the image due to sunlight, and 2) failure to recognize reflective surfaces of glass buildings or buildings of a similar color to the sky. This poses a significant challenge for the segmentation process, as all these error sources are quite commonly encountered in the dense urban areas where the skymask matching is designed to be most helpful. Position 3 reflects this inaccuracy. These error sources can be mitigated in several ways. Improving the convolutional neural network could help refine pixel classification and prevent the mislabeling of pixels. To prevent the overexposure, narrowing the aperture setting could also help.

The skymasks included:

(i) The image converted skymask, segmented by the CNN.

(ii) The hand-labelled image converted skymask.

(iii) The ground truth (GT) skymask generated from 3D model.

The comparison calculation of the skymasks can be found below. The first comparison between (i) and (ii) measures the accuracy of the Neural Network. Ideally, there should be no mean difference (MD) and standard deviation difference (SDD) in elevation angles. The second comparison between (i) and (iii) measures the discrepancy of the image converted skymask to the GT skymask generated by 3D building model at the same location. In theory, the elevation angles MD and SDD should also be the same. Similarity results are shown in Table 3.

$$\mathbf{el}_* \in \{\mathbf{el}^{\text{hand labelled image}}, \mathbf{el}_{GT}^{\text{3D model}}\}$$

$$\text{MD} = \overline{\mathbf{el}^{\mathbf{image}}} - \overline{\mathbf{el}_*} \tag{11}$$

$$\text{SDD} = \sqrt{\frac{(\mathbf{el}^{\text{image}} - \overline{\mathbf{el}^{\text{image}}})^2}{360}} - \sqrt{\frac{(\mathbf{el}_* - \overline{\mathbf{el}_*})^2}{360}}$$

Table 3. Comparison between Image-Converted Skymask and Ground Truth Skymask.

	Image Converted Skymask vs Hand Labelled Image Skymask (degree)		Image Converted Skymask vs GT Skymask generated from 3D Model (degree)	
Ex#	Mean Difference	S.D Difference	Mean Difference	S.D Difference
1	1.42	0.12	-4.02	-5.05
2	-3.90	-0.22	-10.52	-1.93
3	-6.32	2.38	-9.69	0.30
4	-0.77	-0.69	8.28	-9.3

The mean difference and S.D difference between the fisheye image converted skymask compared to the hand labelled fisheye image converted skymask ranges from -6.32 to 1.42 degrees and -0.69 to 2.38 degrees respectively. The difference is due to segmentation inaccuracy, which may incorrectly identify building boundaries in the image. The mean difference between the fisheye image skymask compared to the GT skymask ranges from -10.52 to 8.28 degrees. Mentioned earlier, the segmentation inaccuracy from the neural network contributes to the difference. The second main reason suspected is the assumption that the images are taken at mean sea level, however during the experiments, the images might not exactly be at sea level due to handling difficulties. The change in position will decrease the size of buildings in the image and therefore reduce the elevation angle. Thirdly, as shown in the experiment 4, the generated skymask suggests some inaccuracies in the 3D

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

building model. The level of detail in the 3D building model directly affects the accuracy of the generated skymasks.

3.3 *Positioning Results*

Dots on the heatmap (score map) represent the similarity of the candidates skymasks to the image converted skymask, ranging from dark blue (0% similarity) to dark red (100% similarity). The heat maps are displayed below in Figure 4, with each diamond represents a different method. The positioning error of each method is also recorded in Table 4.

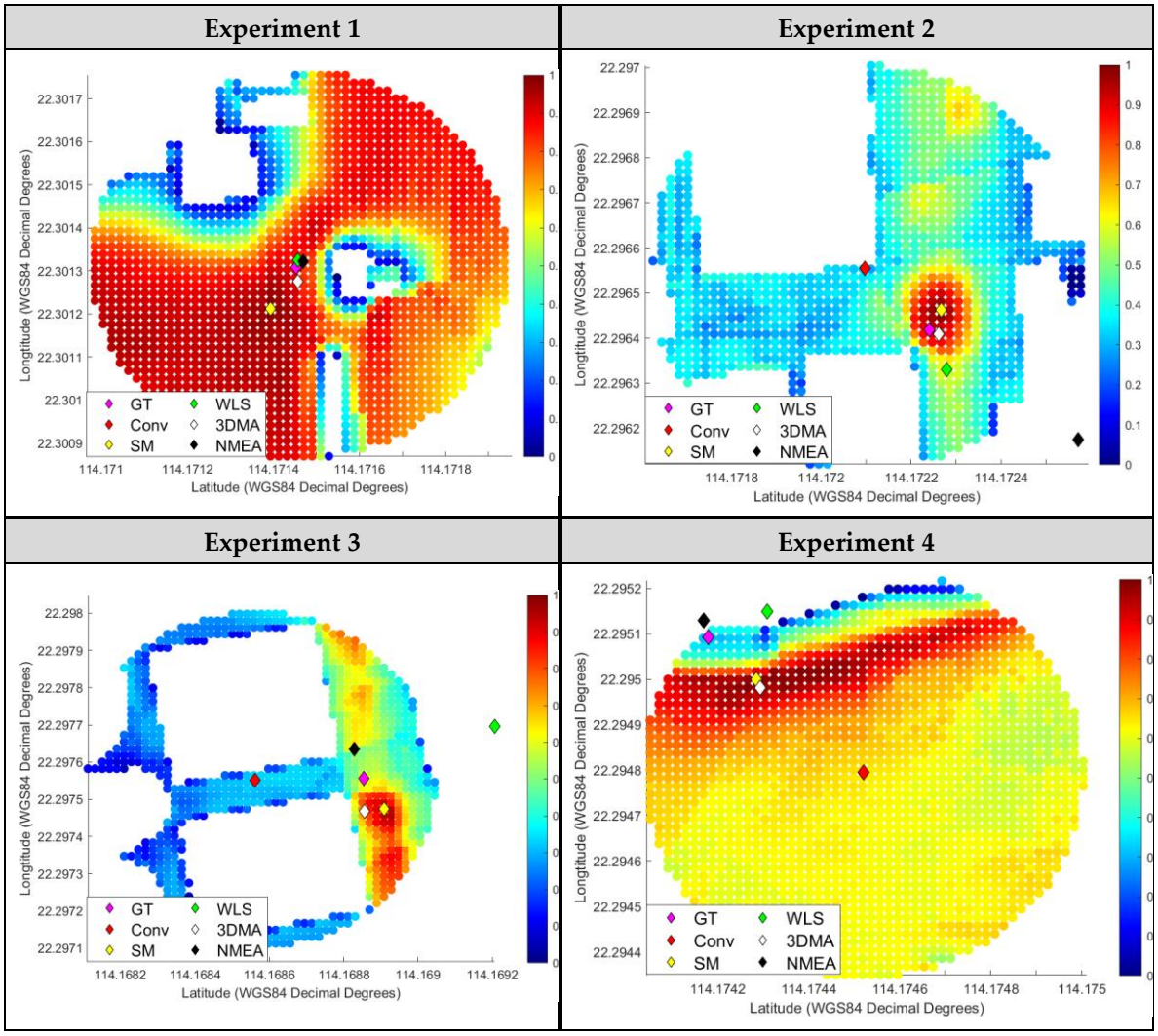


Figure 4. Heatmap on the similarity between the skymasks generated based on the fisheye images and 3D models based on the proposed skymask matching algorithm

Table 4. Performance Comparison of Positioning Results for Different Methods. Unit: meter.

Method	Position 1			Position 2			Position 3			Position 4		
	2D Error	Along-Street Error	Across-Street Error	2D Error	Along-Street Error	Across-Street Error	2D Error	Along-Street Error	Across-Street Error	2D Error	Along-Street Error	Across-Street Error
BroadCom	1.21	0.77	0.90	21.24	15.28	12.13	30.03	4.30	28.16	48.23	26.92	41.33
WLS	2.76	1.96	1.81	10.41	9.87	2.52	40.00	10.59	37.95	14.67	13.91	6.79
Allystar	2.31	1.26	1.70	43.72	24.73	31.66	9.39	9.33	1.06	4.54	1.19	4.46

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

3DMA	3.83	3.83	0.00	2.56	1.33	1.82	10.37	9.97	2.86	17.74	7.74	14.87
SM	12.30	10.39	5.82	7.03	6.93	0.85	10.69	9.30	5.45	15.41	9.92	10.33

In Experiment 1, the results show the accuracy of the proposed Skymask Matching lags behind other methods in open-sky areas. A 2D error increase of 12m from the low-cost GNSS position suggests that SM should not be used in open areas. Other methods makes the positioning more inaccurate (1-2m) as well. This result was expected, as the lack of nearby structures meant that there were few building features to match. This is represented by the high similarity on most candidate skymask in the open areas, as seen in the heat map, most candidate skymask are deep red in color. There is a risk of increasing the along/across street error if the image converted skymask is matched with a wrong skymask, demonstrated in this experiment. A simple workaround to address this problem is to disable skymask matching when the sky takes up more than 50% of the area in a sky pointing image. In such situations, relying on the conventional GNSS coordinates would yield better results as satellite measurements are likely in LOS to the receiver.

Experiment 2 is located in an urban environment, an environment with multiple distinctive high-rise buildings, which provided the features for image matching. When within these feature-rich environments, the skymask matching method improved upon the low-cost GNSS accuracy. The 2D error is 7m for the SM method. Overall, SM improved positioning accuracy to an acceptable degree, and performed second best out of the post-processing methods, coming behind only 3DMA GNSS processing, which had a position error of approximately 2 meters. The inability of NMEA and WLS to establish an accurate position was likely due to the nature of the highly urbanized environment. This environment, however, proved advantageous to SM, which had a bounty of distinctive features to match with.

Experiment 3 was categorized as urban with complex features. This meant that while buildings occupied a larger portion of these locations, resulting in poorer GNSS reception, SM also had more complex features to match with. The SM method yielded good improvements compared to conventional GNSS solutions and WLS method. The 2D error is about 11m for the SM method, which is 19m improvement from the conventional result. While a noticeable improvement, the positioning accuracy still leaves something to be desired. It should be noted that SM had similar accuracy to 3DMA GNSS. Overall, these results suggest that SM can and should be considered being implemented in a complex feature environment and used in conjunction with other methods.

The 4th experiment belongs to the multiplex category, an environment with a balanced mix of trees, buildings and other urban clutter. The fourth experiment showed improvements in the accuracy of the positioning. Skymask matching reduced the positioning error from 48 to 15 meters, with both along and across street error being substantially reduced. In this case, SM performed well, but lacks behind other methods. This suggests SM can be used in positioning in a complex environment and would be enhanced if SM can identify other obstacles such as trees and other urban clutter.

3.5 Heading Resolution Results

Before any images were taken, the camera (and thus the fisheye image) was aligned to north to as described in 3.1. Skymask matching heading error and conventional heading error for each image relative to the true north is shown in Table 5 (positive is clockwise).

Table 5. Skymask Matching Heading Error and Conventional Heading Error

Experiment	Skymask Matching Heading Error (°)	Conventional Heading Error (°)
1	-1	-8
2	0	8
3	0	-1

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

4	1	-45

The predicted bearing offsets for images 1, 2, 3 and 4 are all within acceptable parameters. The results show that in an urban environment with features, the boundary of buildings can be used to accurately estimate the heading offset. Figure 5, below, compares the boundaries of the image converted skymask and predicted candidate skymask. The boundaries of image converted skymask is rotated/adjusted for the predicted bearing offset.

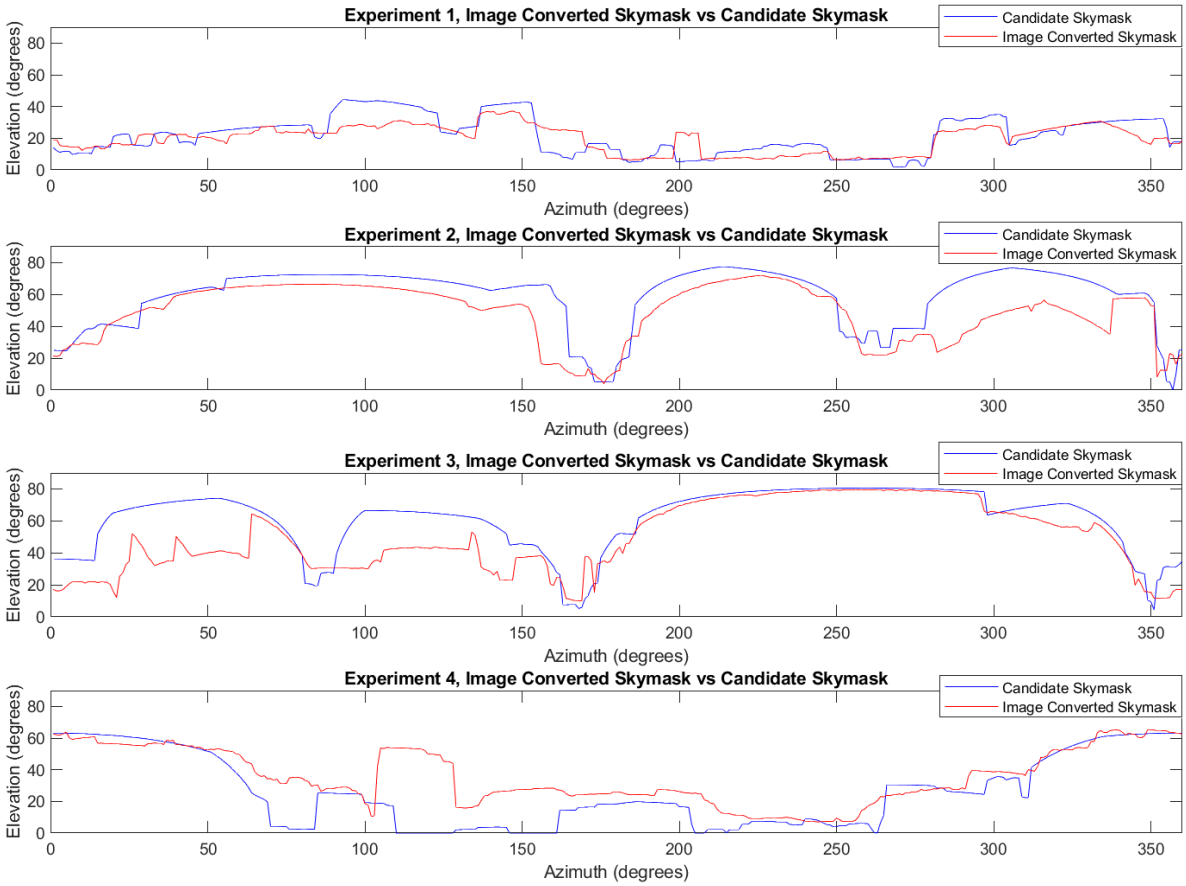


Figure 5. E, Azimuth & Elevation comparison between Image Converted Skymask and Candidate Skymask

All experiments show elevation angle discrepancies between the rotated image converted skymasks and predicted candidate skymasks. In experiment 1, features were lacking, nonetheless the few features were enough for SM to successfully gauge the orientation of the image converted skymask by comparing to the candidate skymask. Experiment 2 shows high accuracy, likely due to the distinctive building boundaries in the image. The accurate prediction of the bearing offset in turn increased position accuracy. Experiment 3's high accuracy suggests that skymask matching can perform well in an urban environment with complex building features. Experiment 4 shows the large discrepancies to the candidate skymask elevation. Despite this challenge, the predicted bearing offset was only 1°. Hence, SM matching can be considered an accurate approach to estimate the heading of the user in a rich urban environment.

3.6 *Discussions*

Table 6 displays the limitations and assumptions made during this study. These factors will now be explained in further detail. The proposed skymask matching concept is similar to GNSS shadow matching by matching satellite visibilities and building boundaries, but the Skymask

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

Matching provides more features to be matched with. For example, the building edges can be matched between image and skymask generated from 3D building model.

Table 6. Summary on the Limitations of the Proposed Skymask Matching

Process	Assumptions/Limitations in this Experiment
Sky-pointing Fisheye Image	Assumes images taken from mean sea level
	Assumes center of images are zenith pointing
	Images were only taken during the day with sunlight
Training Datasets	Requires fully accurate training datasets
	Hand-labelled inaccuracy
Semantic Segmentation	Risk of overfitting
	Limited number of identifying classes
Skymask Database	3D model might be outdated and/or imprecise
Skymask Matching	Limited search radius small than the error of the initial guess
	Only one elevation angle at each azimuth angle

There are several limitations stemming from the usage of sky-pointing fisheye images. The first drawback comes from the location at which the fisheye images are taken. In this study, images are assumed to be taken at mean sea level. Due to human error in this experiment, the fisheye image may be slightly slanted in one direction or otherwise not taken correctly, negatively affecting accuracy because the building heights visible on the image will change. This problem can be mitigated when used for vehicular purposes, where the mounted sky-pointing fisheye camera will stay level.

Precautions must also be taken with the image datasets used to train the convolutional neural network. Because the network labels each individual pixel, the images must be highly accurate. This is a significant problem due to the hand-labelled nature of these images, human errors may result in inconsistent labelling, especially around objects like trees. This can largely be mitigated by setting strict guidelines on how to segment images. For example, in this project, trees were assumed to be solid objects, any patches of sky visible between leaves and branches were ignored.

The semantic segmentation process also had some limitations. First, it requires many unique images in the dataset to increase variation and validation accuracy, there is the risk of overfitting on the images in the dataset. Another flaw is the limited number of classes, the current iteration of the code can identify only 2 classes, the Sky, Building classes. This means that objects that are not buildings, such as trees, signposts, and vehicles, are also classified as buildings. This also limits the usefulness of semantic segmentation, as different building materials are not identified. The former problem can be solved by labelling more unique images to increase the accuracy of the network. The latter problem can be solved by implementing more classes for objects, including building materials.

Another limitation comes from the inaccurate precomputed skymasks. This can occur in several ways, the 3D building model used to generate the skymask could be inaccurate, or the 3D model was out of date, new buildings could have been constructed, or old ones were torn down, leading to discrepancies between the fisheye image and skymask. The limitation can be solved by ensuring that the utilized 3D maps are highly accurate and constantly updated.

The primary limitation in skymask matching comes from the search radius of 50m. If the convolutional GNSS position error is larger than 50m, then the ability of the image matching code to return an accurate result is limited by the search radius, reducing the ability of the code to return an accurate result. This could be offset by increasing the size of the search radius but would also increase computing time and power required.

4. Conclusion and Future Works

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

This paper proposes a new method by introducing a new source of fisheye image data. First, a fisheye camera is used to capture the sky view fisheye image. Then, a convolutional neural network works in conjunction with active contouring to segment the fisheye image. The segmented image is then converted into a skymask and matched with its pre-computed counterparts. The similarity between the image skymask and pre-computed skymask is then regarded as the score of the position candidate. Compared with the ray-tracing based 3DMA GNSS, the proposed algorithm can provide similar correction in an environment with distinctive building features. In addition, the heading angle estimated by the proposed skymask matching algorithm is very accurate. As such, skymask matching is a promising candidate for use in vehicular navigation.

However, the proposed skymask matching method still has limitations. The image capture process may be affected by a number of factors. The CNN is time-consuming to train due to the hand-labelled nature of the datasets. Even after training, semantic segmentation may sometimes mislabel building surfaces due to lens flare or reflections. The nature of algorithm used in this research also means it only compares one elevation angle at each azimuth angle. Additionally, the precomputed skymasks may not be accurate due to poor model quality or being outdated, while too large of a conventional GNSS position error and lack of building features may limit the accuracy of image matching.

In the future, the algorithm is to distinguish between not only sky and buildings but also classes like trees and materials. Subclasses of building material will be added as well, as larger amounts of details provide the possibility for higher position accuracy. Allowing distinction between these semantic classes opens further avenues to increase segmentation accuracy, as areas with vehicles and tree labels can be weighted less in scoring. While building materials can be factored into similarity matching to allow improve results in Skymask Matching and other 3DMA GNSS processing techniques, because different materials reflect GNSS signals differently, the ability to identify textures could benefit raytracing GNSS. The skymask matching could also be further improved by extending the functionality in different weather, e.g. only sunny visible weathers, for safety concern.

Author Contributions: Conceptualization, M.J.L.L.; methodology, M.J.L.L., and S.L.; validation, M.J.L.L., S.L.; formal analysis, M.J.L.L., S.L.; data curation, M.J.L.L., S.L.; writing—original draft preparation, M.J.L.L., S.L.; writing—review and editing, M.J.L.L., S.L., H.-F.N., and L.-T.H.; visualization, M.J.L.L., S.L.; supervision, L.-T.H., H.-F.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Final Year Project Committee of Interdisciplinary Division of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University.

Acknowledgments: The authors would like to acknowledge the financial support of the Hong Kong Polytechnic University. Finally, the authors would like to thank the journal editor and peer reviewers for their constructive comments on this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

References

[1] X. Li *et al.*, "Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo," *Journal of Geodesy*, vol. 89, no. 6, pp. 607-635, March 2015 2015, doi: 10.1007/s00190-015-0802-8.

[2] L.-T. Hsu, "Analysis and modeling GPS NLOS effect in highly urbanized area," *GPS Solutions*, vol. 22, no. 1, p. 7, 2017, doi: 10.1007/s10291-017-0667-9.

[3] P. D. Groves and M. Adjrad, "Likelihood-based GNSS positioning using LOS/NLOS predictions from 3D mapping and pseudoranges," *GPS Solutions*, vol. 21, no. 4, pp. 1805-1816, 27th July 2017 2017, doi: 10.1007/s10291-017-0654-1.

[4] P. D. Groves and M. Adjrad, "Performance assessment of 3D-mapping–aided GNSS part 1: Algorithms, user equipment, and review," *Navigation*, vol. 66, no. 2, pp. 341-362, 26 March 2019 2019, doi: 10.1002/navi.288.

[5] M. Adjrad, P. D. Groves, J. C. Quick, and C. Ellul, "Performance assessment of 3D-mapping-aided GNSS part 2: Environment and mapping," *Navigation*, vol. 66, no. 2, pp. 363-383, 18th March 2019 2019, doi: 10.1002/navi.289.

[6] T. Suzuki and N. Kubo, "2A2-N04 Localization Based on GNSS Visibility and Positioning Error using 3D City Model," in *JSME annual Conference on Robotics and Mechatronics (Robomec)*, 17th May, 2015 2015, doi: 10.1299/jsmermd.2015._2A2-N04_1 [Online]. Available: https://www.jstage.jst.go.jp/article/jsmermd/2015/0/2015__2A2-N04_1/_article/-char/ja/

[7] L. Wang, P. D. Groves, and M. K. Ziebart, "GNSS Shadow Matching: Improving Urban Positioning Accuracy Using a 3D City Model with Optimized Visibility Scoring Scheme," *Navigation*, vol. 60, no. 3, pp. 195-207, 28th August 2013 2013, doi: 10.1002/navi.38.

[8] T. Suzuki and N. Kubo, "N-LOS GNSS signal detection using fish-eye camera for vehicle navigation in urban environments," in *27th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2014*, Tampa, United States of America, 8-12 September 2014 2014: Institute of Navigation, pp. 1897-1906.

[9] M. Smyrnaios, S. Schn, and M. Liso, "Multipath Propagation, Characterization and Modeling in GNSS," in *Geodetic Sciences - Observations, Modeling and Applications*, S. Jin Ed. London, United Kingdom of Great Britian and Northern Ireland: Intech, 2013, ch. Chapter 2, pp. 99-125.

[10] E. Realini and M. Reguzzoni, "goGPS: open source software for enhancing the accuracy of low-cost receivers by single-frequency relative kinematic positioning," *Measurement Science and Technology*, vol. 24, no. 11, pp. 595-603, 16th October, 2013 2013, doi: 10.1088/0957-0233/24/11/115010.

[11] M. Obst, S. Bauer, and G. Wanielik, "Urban multipath detection and mitigation with dynamic 3D maps for reliable land vehicle localization," in *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, Myrtle Beach, South Carolina, United States of America, 23-26 April, 2012 2012: IEEE, pp. 685-691, doi: 10.1109/PLANS.2012.6236944. [Online]. Available: https://ieeexplore.ieee.org/document/6236944/

[12] H. Xu, L.-T. Hsu, D. Lu, and B. Cai, "Sky visibility estimation based on GNSS satellite visibility: an approach of GNSS-based context awareness," *GPS Solutions*, vol. 24, no. 2, 16th March, 2020 2020, doi: 10.1007/s10291-020-0973-5.

[13] W. Wen, G. Zhang, and L.-T. Hsu, "GNSS NLOS Exclusion Based on Dynamic Object Detection Using LiDAR Point Cloud," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-10, 31st December, 2019 2019, doi: 10.1109/TITS.2019.2961128.

[14] J. Moreau, S. Ambellouis, and Y. Ruichek, "Fisheye-Based Method for GPS Localization Improvement in Unknown Semi-Obstructed Areas," *Sensors (Basel)*, vol. 17, no. 1, p. 119, Jan 17 2017, doi: 10.3390/s17010119.

[15] L. Lau and P. Cross, "Development and testing of a new ray-tracing approach to GNSS carrier-phase multipath modelling," *Journal of Geodesy*, vol. 81, no. 11, pp. 713-732, 1st November 2007 2007, doi: 10.1007/s00190-007-0139-z.

[16] L.-T. Hsu, Y. Gu, and S. Kamijo, "3D building model-based pedestrian positioning method using GPS/GLONASS/QZSS and its reliability calculation," *GPS Solutions*, vol. 20, no. 3, pp. 413-428, 28th March, 2015 2015, doi: 10.1007/s10291-015-0451-7.

[17] P. D. Groves and Z. Jiang, "Height Aiding, C/N0 Weighting and Consistency Checking for GNSS NLOS and Multipath Mitigation in Urban Areas," *Journal of Navigation*, vol. 66, no. 5, pp. 653-669, 2nd July 2013 2013, doi: 10.1017/s0373463313000350.

[18] T. Suzuki and N. Kubo, "GNSS Photo Matching: Positioning using GNSS and Camera in Urban Canyon," in *28th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Tampa, Florida, United States of America, 14-18 September 2015 2015, vol. 4: Institute of Navigation, pp. 2470-2480.

[19] H.-F. Ng, G. Zhang, and L.-T. Hsu, "A Computation Effective Range-Based 3D Mapping Aided GNSS with NLOS Correction Method," *Journal of Navigation*, vol. 73, no. 3, pp. 1-21, 30th June 2020 2020, doi: 10.1017/s037346332000003x.

[20] "Get Started with the Image Labeler." MATLAB. https://www.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html (accessed.

[21] "ResNet-50 Convolutional Neural Network." MATLAB. https://www.mathworks.com/help/deeplearning/ref/resnet50.html (accessed.

[22] T. F. Chan and L. Vese, "An Active Contour Model Without Edges," in *Scale-Space Theories in Computer Vision*, no. 1682). Heidelberg, Germany: Springer Verlag GmbH, 1999, pp. 141-151.

[23] N. S. Gowdayyanadoddi, J. T. Curran, A. Broumandan, and G. Lachapelle, "A Ray-Tracing Technique to Characterize GPS Multipath in the Frequency Domain," *International Journal of Navigation and Observation*, vol. 2015, pp. 1-16, 28th September, 2015 2015, doi: 10.1155/2015/983124.

[24] A. Cohen, C. Meurie, Y. Ruichek, and J. Marais, "Characterization of the reception environment of GNSS signals using a texture and color based adaptive segmentation technique," in *2010 IEEE Intelligent Vehicles Symposium*, San Diego, California, United States of America, 21-24 June 2010 2010: IEEE, pp. 275-280, doi: 10.1109/IVS.2010.5548018. [Online]. Available: https://ieeexplore.ieee.org/document/5548018/

[25] W.Wen, Y. Zhou, G.Zhang, S. Fahandezh-Saadi, X.Bai, W. Zhan, M.Tomizuka, L. Hsu, "UrbanLoco: A Full Sensor Suite Dataset for Mapping and Localization in Urban Scenes", *IEEE ICRA 2020*, Paris, France.

Peer-reviewed version available at Sensors 2020, 20, 4728; doi:10.3390/s20174728

Max Jwo Lem Lee is currently a student in the Bachelor of Engineering (Honours) in Aviation Engineering from the Hong Kong Polytechnic University, Hong Kong. His research interests include precise positioning in urban environments, navigation and unmanned aerial vehicle.

Shang Lee received a Bachelor of Engineering (Honours) in Aviation Engineering from the Hong Kong Polytechnic University, Hong Kong, in 2020. He is currently studying as a M.Sc. student in the Hong Kong University of Science and Technology's M.Sc. in Aeronautical Engineering Program. His research interests include aviation and meteorology.

Hoi-Fung Ng received a Bachelor of Engineering (Honours) in Air Transport Engineering from The Hong Kong Polytechnic University, Hong Kong, in 2018.
He is currently an M.Sc. student at the Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong. His research interests including GNSS localization, navigation.

LI-TA HSU (S'09-M'15) received the B.S. and Ph.D. degrees in aeronautics and astronautics from National Cheng Kung University, Taiwan, in 2007 and 2013, respectively. He is currently an assistant professor with the Division of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University, before he served as post-doctoral researcher in Institute of Industrial Science at University of Tokyo, Japan. In 2012, he was a visiting scholar in University College London, U.K. He is an Associate Fellow of RIN. His research interests include GNSS positioning in challenging environments and localization for pedestrian, autonomous driving vehicle and unmanned aerial vehicle.