ORIGINAL ARTICLE

Machine learning based LOS/NLOS classifier and robust estimator for GNSS shadow matching

Haosheng Xu[1], Antonio Angrisano[2], Salvatore Gaglione[3] and Li-Ta Hsu[1*]

Abstract

Global Navigation Satellites Systems (GNSS) is frequently used for positioning services in various applications, e.g., pedestrian and vehicular navigation. However, it is well-known that GNSS positioning performs unreliably in urban environments. GNSS shadow matching is a method of improving accuracy in the cross-street direction. Initial position and classification of observed satellite visibility between line-of-sight (LOS) and non-line-of-sight (NLOS) are essential for its performance. For the conventional LOS/NLOS classification, the classifiers are based on a single feature, extracted from raw GNSS measurements, such as signal noise ratio, pseudorange, elevation angle, etc. Especially in urban canyons, these measurements are unstable and unreliable due to the signal reflection and refraction from the surrounding buildings. Besides, the conventional least square approach for positioning is insufficient to provide accurate initialization for shadow matching in urban areas. In our study, shadow matching is improved using the initial position from robust estimator and the satellite visibility determined by support vector machine (SVM). The robust estimator has an improved positioning accuracy and the classification rate of SVM classification can reach 91.5% in urban scenarios. An important issue is related to satellites with ultra-high or low elevation angles and satellites near the building boundary that are very likely to be misclassified. By solving this problem, the SVM classification shows the potential of about 90% classification accuracy for various urban cases. With the help of these approaches, the shadow matching has a mean error of 10.27 m with 1.44 m in the cross-street direction; these performances are suitable for urban positioning.

Keywords: GNSS, Shadow matching, NLOS and machine learning, Robust estimation

Introduction

Positioning has become a part of our everyday life. People heavily rely on GNSS-enabled applications to navigate himself or herself to a destination. However, GNSS positioning is greatly affected by the notorious multipath and non-line-of-sight (NLOS) reception phenomena (Ji et al. 2010; Tsakiri et al. 1998). These effects are due to the signal blockage and reflection by and on the buildings. In other words, the more urbanized the city is, the more

challenge on the GNSS positioning is. This is one of the current problem of smartphone service providers; consequently, a solution for multipath and NLOS is needed.

To solve these problems, many technologies have been developed such as: pseudorange error modeling (Viandier et al. 2008), consistency checking (Hsuet al 2017), robust estimation (Gaglione et al. 2017), inertial sensors (El-Sheimy and Youssef 2020) and more available satellites (Yang et al. 2020). Beyond conventional approaches, one of the innovative solutions is related to the use of 3D building models. Since the rise of smart cities, the 3D city models become wid ely available, especially for highly urbanized cities, including Hong Kong, New York, Tokyo, and London. These models can be used to effectively

*Correspondence: lt.hsu@polyu.edu.hk
[1] The Hong Kong Polytechnic University, Kowloon, Hong Kong, People's Republic of China
Full list of author information is available at the end of the article

Xu *et al. Satell Navig*   (2020) 1:15

Page 2 of 12

simulate the GNSS signal transmission in urban areas. The methods that used the 3D mapping database to facilitate the GNSS positioning are called 3D mapping aided (3DMA) GNSS. One of the most effective 3DMA GNSS method is the GNSS shadow matching (Groves 2011). The shadow matching technique compares the visible GNSS satellites, from the hypothesized locations on 3D map, with the measurements, classified by received signal strength. This approach improves the GNSS positioning in urban canyons, especially reducing the positioning error in the across-street direction. This improvement could be extremely meaningful for applications like car sharing. The smartphone users are sometimes fooled by incorrect GNSS positioning, causing unpleasant user experiences. Theoretically, the GNSS shadow matching provides a good solution if the following assumptions are met: the GNSS measurements are correctly classified; the initial position accuracy of the shadow matching algorithm is within 40 m (Groves et al. 2015). The paper aims to provide an algorithms able to validate these two assumptions.

The proposed LOS/NLOS classification for shadow matching, is tested on measurements provided by a smartphone GNSS receiver firstly. The classification is based on the signal strength of satellites, expressed by the signal to noise ratio (SNR) measurement. The satellite with high SNR value is more likely to be considered as the LOS satellite. However, surrounding buildings, with facades made of materials likes steel and glass, increase the reflected signal strength, making the classification based on SNR insufficient (Wang et al. 2015). After that, a robust classification, considering the SNR, pseudorange, and elevation angle of the received signals in a decision tree method, is raised. Comparing with the SNR classification, the robust classifier takes more satellite features into consideration and have a better classification rate in LOS satellite detection (Yozevitch et al. 2016). Furthermore, the robust classifier has been combined with the shadow matching over a particle filter (Yozevitch and Moshe 2015). The support vector machine (SVM) is also applied to the LOS/NLOS classification (Hsu 2017). From the paper, the classification is compared in single feature and multiple features, and the difference between delta pseudorange and pseudorange rate (pseudorange rate consistency) is proved to have a positive impact on the classification. In the paper, various machine learning methods are compared, including k-nearest neighbors (KNN), neural network (NN), SVM and decision tree (TREE) (Xu et al. 2018). Among these approaches, the SVM method using the features of commercial GNSS receiver has a good performance in different urban scenarios and decent generalization ability. In this paper, we extended the SVM classifier based on the features

available in smartphone level GNSS chip. In addition, a robust estimator on single point positioning (SPP) is implemented for the initialization (Gaglione et al. 2017). The experiment results showed that the improved GNSS shadow matching with SVM classification could achieve 10.27 m of mean error and 1.44 m of error in across street direction in the urban canyons of Hong Kong.

This paper is organized as follows. Section 2 gives an overview of the improved GNSS shadow matching based on the classifier and robust SPP. In Sect. 3, the proposed machine learning LOS/NLOS classifier based on several GNSS measurement features is introduced. Section 4 provides a brief description of weighted robust estimator. Section 5 presents the results of the classification, different SPP solutions, and the integrated shadow matching. Section 6 gives the conclusion and future work based on the findings of this study.

Shadow matching based on LOS/NLOS classifier and weighted robust estimator

The shadow matching (SDM) is integrated with a robust estimator and the SVM classification as shown in Fig. 1. The modules highlighted are the contribution of this paper, compared with the shadow matching algorithm proposed in Wang (2014). The main modules of innovating shadow matching are introduced as follows:
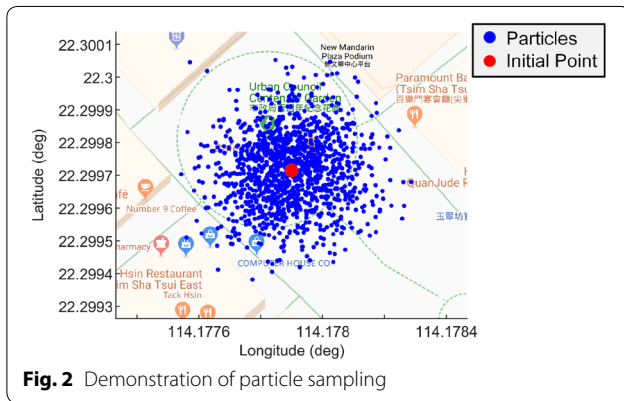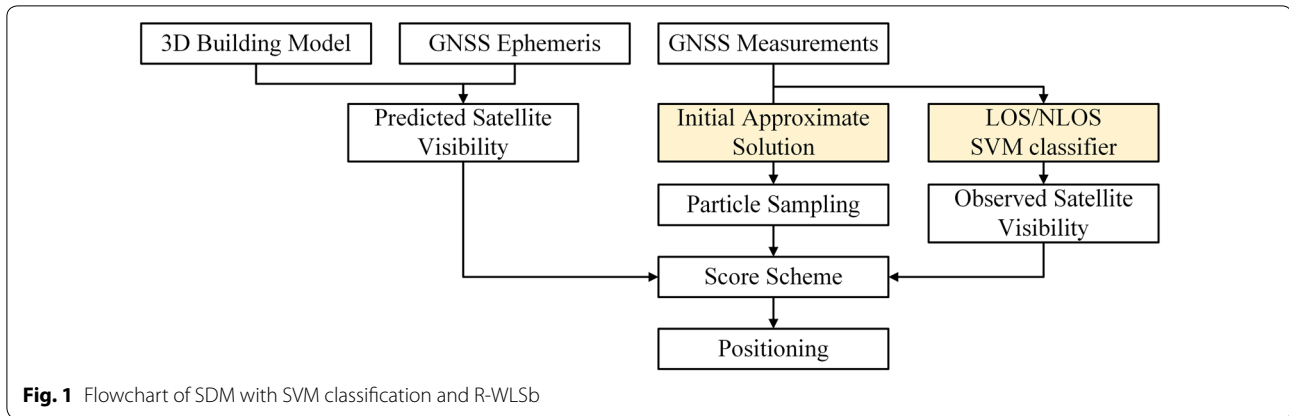
Initial approximate solution

The conventional initial position guesses of shadow matching are weighted least square solution, NLOS probabilities based weighted least square solution (Adjrad and Groves 2017) and the previous solution of shadow matching. (Adjrad and Groves 2018) For this paper, the initial position guess is based on a weighted robust estimator solution introduced in Sect. 4.

Particle sampling

The particle sampling is a method to scatter Gaussian distributed particles as the shadow matching position candidates. In detail, a search area, centered at the initial approximate solution firstly, is considered. All sampled particles have the elevation angle of the surrounding building boundary at every azimuth angle. The benefits of this sampling include dropping the number of position candidates and reducing the running time of algorithm. The demonstration of particle sampling is shown in Fig. 2.

Observed satellite visibility

In urban areas, several signals from satellites are blocked by surrounding tall buildings. With the help of 3D building models, the satellite visibility of each particle is

Fig. 1 Flowchart of SDM with SVM classification and R-WLSb

Fig. 2 Demonstration of particle sampling

predicted by comparing the elevation angles of satellites and building boundaries at the same azimuth angle.

Predicted satellite visibility

The conventional LOS/NLOS classifiers are simply based on SNR measurement and elevation angle of satellite using a decision tree. The essence of decision tree classifier is still a threshold based on the specific conditions. These attempts still face the issue of SNR unreliability and of its changing behavior according to the scenario. For these reasons, a multi-featured classification approach is necessary to supply SVM classification. The detail of the proposed SVM classifier will be given in Sect. 3.

Score scheme

After resolved for the observed and predicted satellite visibility, a scoring scheme for candidate locations is used. If a satellite, expected to be unblocked by buildings, is classified as the LOS signal, the candidate will gain a score. Conversely, for satellite predicted as blocked, only by meeting the labeled condition of NLOS or not measured, the satellite would add a score for this candidate position as shown in Table 1.

Table 1 A score scheme for shadow matching

Score scheme		Measured and Classified		Not measured
		LOS	NLOS	\
Predicated	Blocked	0	1	0
	Unblocked	1	0	1

Positioning

The last step to generate a positioning solution is employing scores of every candidate particle. The probability of ground truth position solution would be derived from the score of the candidate particle. For particle *i*, the possibility is calculated as following:

$$\rho_i = \frac{s_i - s_{min}}{s_{max} - s_{min}}, \tag{1}$$

where s_{min} and s_{max} are the maximum and minimum scores of all particles. For the current positioning, a threshold of 85% is set to pick the candidate particles over possibility threshold. Afterward, for the final position solution, the average of the picked candidate particles is computed as following:

$$P_{sm} = \frac{\sum_{i=1}^{n} \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}}{n}, \tag{2}$$

where $\begin{bmatrix} x_i & y_i & z_i \end{bmatrix}$ are the coordinates of the picked candidate particles in earth-centered, earth-fixed (ECEF) frame; n is the number of picked candidate particles.

LOS/NLOS classifier based on machine learning

This section provides a brief introduction to the principles of the SVM LOS/NLOS classifier. Moreover, the features for LOS/NLOS classification are discussed.
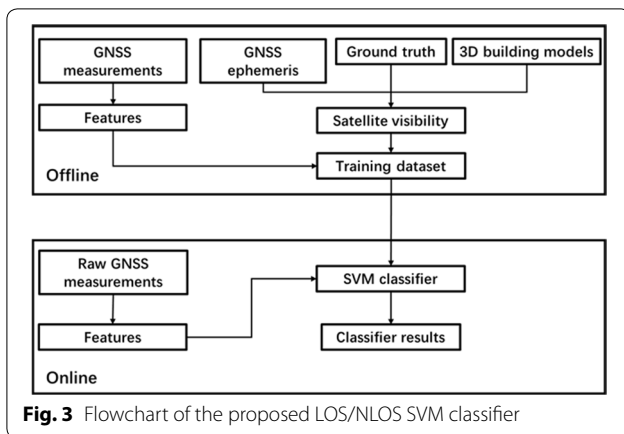
Xu *et al. Satell Navig*   (2020) 1:15

Page 4 of 12

Fig. 3 Flowchart of the proposed LOS/NLOS SVM classifier

The architecture of classifier

The architecture of SVM classifier contains two stages: offline and online, as shown in Fig. 3. For the offline stage, the raw GNSS measurements are used for extracting features of machine learning approach, and the features are labeled using the 3D building models, ground truth and satellite positions calculated by GNSS ephemeris. The elevation and azimuth angles of satellites could be calculated by the satellite position from GNSS ephemeris and the ground truth position, then the elevation of satellites is compared with the elevation angle of building edges at the same azimuth angle. For the LOS satellites, the elevation is higher than the maximum elevation angle of the buildings at the same azimuth angle, and vice versa. Finally, an offline labeled dataset is created to train a linear SVM classifier. For a linear SVM classifier, the score of classification is calculated by:

$$Score(x) = (x/s)^T \beta + b, \qquad (3)$$

whereЇ x is the machine learning feature vectors, and s, β, b donate the kernel scale, the vector of fitted linear coefficients and bias from linear SVM classifier, respectively. The predicted LOS/NLOS label is calculated by:

$$Label = \begin{cases} LOS & Score \geq 0 \\ NLOS & Score < 0 \end{cases}, \qquad (4)$$

For the online stage, the feature vector from the raw GNSS measurement is put into the SVM score formula to obtain the predicted satellite visibility.

Features of machine model

According to our preliminary result (Xu et al. 2018), there are differences between LOS and NLOS signals existing in features as follows:

Signal noise ratio (SNR)

The SNR is a conventional variable to predict satellite visibility, because the reflection and refraction of the NLOS signal transmission decrease the signal strength for most cases. The signal strength of each received signal could be obtained from the raw GNSS measurements in receiver independent exchange format (RINEX) data. To present the real SNR measurement, a dataset of about 20 min is collected in urban scenario, as shown in Fig. 4. It is evident that there are some SNR regions where the LOS and NLOS signals coexist at the same time, demonstrating that the simple SNR threshold classification might not work perfectly in urban environments.

Normalized pseudorange residual (NPR)

The pseudorange residual is also a useful feature related with satellite visibility (Hsu et al. 2017). The pseudorange residual is computed by the least square approach, which is a conventional approach to estimate user position. The least square approach is computed by:

$$X = \left(H^T H\right)^{-1} H^T \rho, \qquad (5)$$

where X is a vector with the estimated receiver position and clock bias, H is a matrix with unit LOS vectors pointing from the receiver to satellites. ρ denotes pseudorange measurements. After iterations, the pseudorange residual of each satellite is expressed as:

$$Pr = \rho - H \cdot X, \qquad (6)$$

However, the estimated position in urban area always contains a large error, so the pseudorange residual could not indicate the difference between LOS and NLOS signals clearly. For that reason, the pseudorange residuals of each epoch are normalized as:
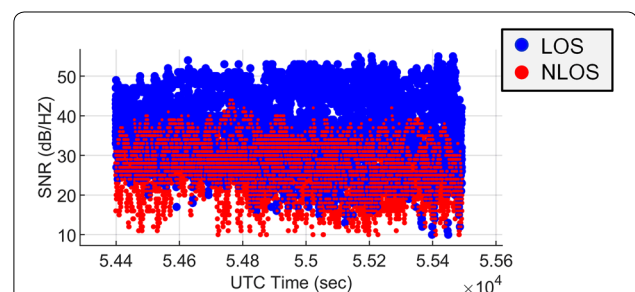
$$NPR = \frac{Pr_i - Pr_{min}}{Pr_{max} - Pr_{min}}, \qquad (7)$$

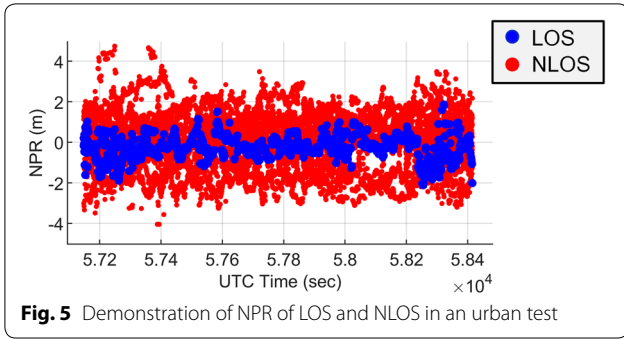**Fig. 4** Demonstration of SNR of LOS and NLOS in an urban test

Fig. 5 Demonstration of NPR of LOS and NLOS in an urban test

where Pr_{max} and Pr_{min} are maximum and minimum pseudorange residual of each epoch. A demonstration of normalized pseudorange residual is shown in Fig. 5. With an accurate position estimation, the normalized pseudorange residual of LOS signal is closer to zero than that of NLOS signal, since the NLOS signal have additional propagation path in pseudorange.

Elevation angle (EA)

The elevation angle of satellite has relationship with the satellite visibility. The main reason is that the higher elevation angle signal is less possible to be blocked by the surrounding building. The existing classification algorithm also applied the elevation angle into LOS/NLOS classification (Yozevitch et al. 2016).

Pseudorange rate consistency (PRC)

The pseudorange rate is the changing rate of pseudorange measurement between two epochs and expressed as:

$$\Delta P_t^i = \rho_t^i - \rho_{t-1}^i, \tag{8}$$

where ρ_t^i and ρ_{t-1}^i is the pseudorange measurement of satellite i at epoch t and t-1. The pseudorange measurement of raw data comes from the receiver code tracking loop. Meanwhile, the Doppler shift of signal is estimated from the receiver frequency tracking loop, and the pseudorange rate could be related with Doppler shift by:

$$\rho_t^i = \left(-\lambda_i \cdot f_{d.i}\right)\Delta t, \tag{9}$$

where λ_i is the negative of carrier wavelength and $f_{d.i}$ is Doppler shift measurement for satellite i. Comparing with receiver code tracking loop, the multipath and reflection path have less impact on frequency tracking loop, which shows the consistency between the pseudorange rate from pseudorange measurement and Doppler shift could reveal the influence from NLOS signal. The pseudorange measurement consistency is expressed by:

$$PRC = \rho_t^i - \Delta P_t^i, \tag{10}$$

where $\hat{P}_t^i$ and ΔP_t^i are the pseudorange rate from Doppler shift and pseudorange measurement respectively. The pseudorange rate of LOS signal have a more stable and smaller absolute value than that of NLOS signal, as shown in Fig. 6.

After generated from the GNSS measurements, the four features are used in the proposed SVM classifier.

Weighted robust estimation

The most common mode in GNSS navigation is the SPP, where the used estimation technique is the LS (or WLS) method. LS optimization criterion is the minimization of the sum of the squared residuals; it is very popular due to its simplicity, being the LS estimation computable explicitly from the measurements as shown in Sect. 3. The main drawback of LS (and WLS) is its sensibility to anomalous measurements (in literature also called outliers or blunders) (Rousseeuw and Leroy 1987). In general, two different strategies could be carried out to tackle the outlier issue:

- a diagnostic approach
- a robust approach.

The diagnostic approach consists of identifying and rejecting the outliers by checking the consistency of redundant measurements; in GNSS context, RAIM techniques follow this way (Brown 1993; Castaldo et al. 2014; Kuusniemi et al. 2004). On the other hand, the robust approach is carried out by the robust estimators which are inherently resistant to outliers. Several classes of robust estimators exist, e.g. L-estimators, M-estimators, R-estimators, differing each other for the optimization criterion. The robust approach has been applied to GNSS in Gaglione et al. (2017), Knight and Wang (2009). In particular, in Gaglione et al. (2017) it has been shows the effectiveness of Huber M-estimator for processing
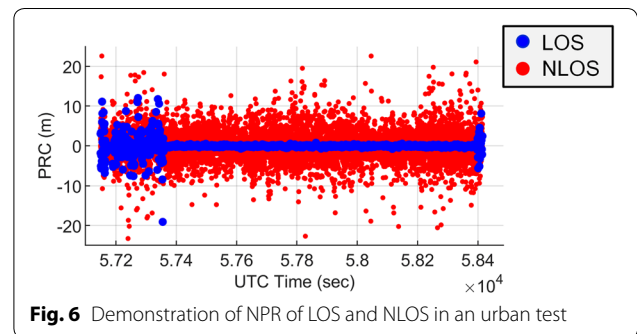

Fig. 6 Demonstration of NPR of LOS and NLOS in an urban test

Xu *et al. Satell Navig*   (2020) 1:15

Page 6 of 12

GNSS measurements in urban scenario and it has been demonstrated the importance of using a suitable weighting scheme for improve the performance of the robust estimators. In this work, a similar approach is followed and Huber M-estimator, with weighting scheme based on signal-to-noise ratio and satellite elevation, is used to provide the initial solution to shadow matching algorithm. The implemented technique is shortly indicated a weighted robust estimator.

In general, in M-estimators, the solution is obtained applying iteratively the WLS, with the weights depending on the residuals.

In Huber M-estimator, the i-th diagonal element of the weighting matrix at each iteration j, is obtained as follows:

$$(w_i)_j = \begin{cases} (w_i)_{j-1}, & \frac{|(r_i)_{j-1}|}{\hat{\sigma}_0} \le k \\ (w_i)_{j-1} \frac{k}{|(r_i)_{j-1}/\hat{\sigma}_0|}, & \frac{|(r_i)_{j-1}|}{\hat{\sigma}_0} > k \end{cases} \qquad (11)$$

where $(r_i)_{j-1}$ is the residual of the i-th measurement at the j-th iteration, k is a constant set to 1.345, $\hat{\sigma}_0$ is the standard deviation of the residuals.

The initial weighting matrix is defined according to the following pseudorange variance model σ_{PR}^2, whose effectiveness has been demonstrated in Angrisano et al. (2018) and Tay and Marais (2013):

$$\sigma_{PR}^2 = A \frac{10^{-\frac{SNR}{10}}}{\sin^2(El)}, \qquad (12)$$

where A is a parameter empirically defined, El is the satellite elevation in degrees, SNR is the signal-to-noise ratio of the carrier in dB, and the bandwidth of receiver is 1 Hz.

The initial i-th weight is obtained as $(w_i)_0 = 1/(\sigma_{PR}^2)_i$.

Experimental results and discussion
Experiment setup
The static data were collected at serval different locations in in Hung Hom, Hong Kong with a UBLOX NEO M8T receiver.

To test all the possible blockage geometries of the urban environment, the experimental locations are selected as shown in Fig. 7 and their sky plots with building boundaries are shown in Fig. 8. About 15–20 min of data were collected in each point.

After collected the raw data from above locations, features for training approach are derived as Signal Noise Ratio (SNR), Normalized Pseudorange Residual (NPR), Elevation Angle (EA) and Pseudorange Rate Consistency (PRC). For the training stage, these features of all locations are sorted into a random ordering, labeled by the surrounding 3D building model and the ground truth,
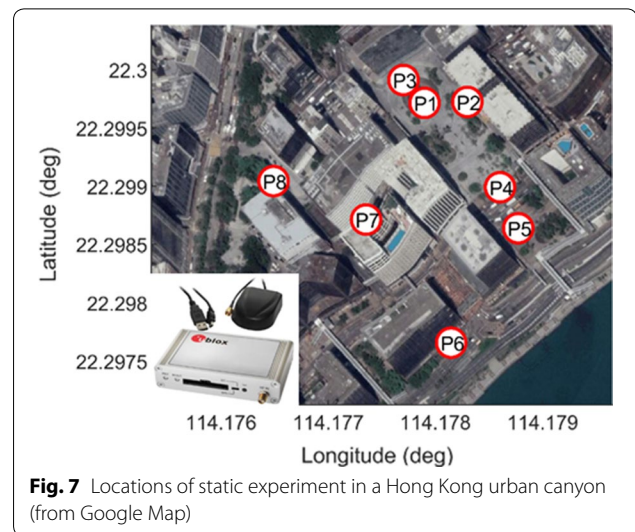

Fig. 7 Locations of static experiment in a Hong Kong urban canyon (from Google Map)

and dropped in to the linear SVM model to produce a machine learning model. Furthermore, the data of above locations is also used for testing the classification rate of the linear SVM model and the positioning performance of our improved shadow matching algorithm. Additionally, the model of classification is based on the given training data, and different kind of GNSS receiver has its unique setting like bandwidth, antenna gain and satellite constellation, which make features of the same signal variously. Therefore, the model of classification should be trained for each receiver.

Classification results
Considering the collected data as training dataset and testing dataset for SVM classifier, the classification rate of SVM classification and simple SNR classification are shown in Table 2. The simple SNR classification is express as follows:

$$\text{Label} = \begin{cases} LOS & SNR \ge 35\,\text{dB} \\ NLOS & SNR < 35\,\text{dB} \end{cases}, \qquad (13)$$

From Table 2, it could be noted that the SVM classification has improvements for urban cases. The SVM classification rate of P2, P3, P4 and P8 has a better performance than the SNR classification. Moreover, a confusion statistic between LOS and NLOS satellite is shown in Table 3.

From Table 3, most of SVM classification for LOS satellite is good in urban environment such as P1, P2, P3 and P7. Meanwhile, the SVM classification of two-side blockage building geometry (P4 and P5) is not good for LOS classification, but still useful for NLOS classification. In such environment, there are many satellites near to the building edges. For P6 test, the SVM classification failed totally in both of LOS and NLOS
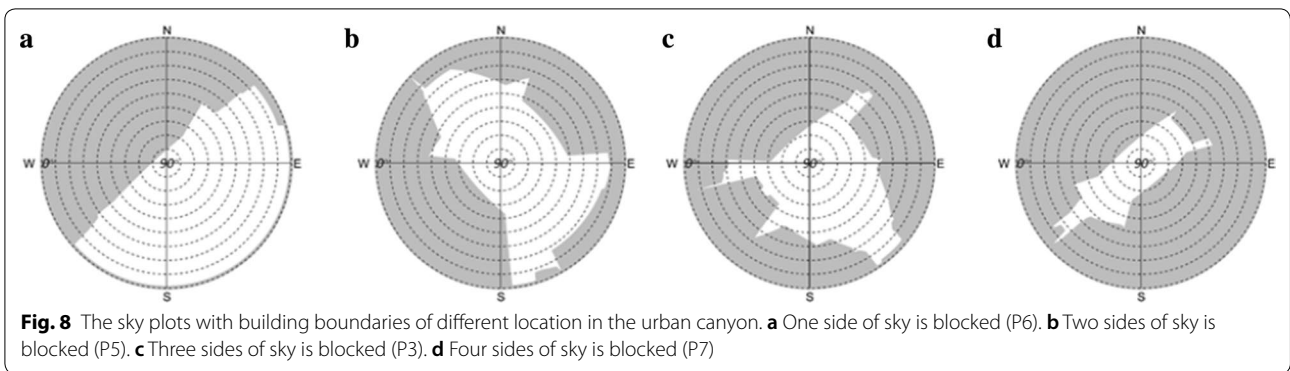
Xu *et al. Satell Navig*   (2020) 1:15

Page 7 of 12

Fig. 8 The sky plots with building boundaries of different location in the urban canyon. **a** One side of sky is blocked (P6). **b** Two sides of sky is blocked (P5). **c** Three sides of sky is blocked (P3). **d** Four sides of sky is blocked (P7)

Table 2 The SNR and SVM Classification rate of the experiment using U-BLOX receiver at the Hong Kong urban area

Location	P1	P2	P3	P4	P5	P6	P7	P8
SNR	95.8%	83.0%	71.0%	56.5%	74.3%	81.2%	85.4%	71.7%
SVM	84.6%	91.5%	73.2%	63.9%	63.0%	45.4%	67.6%	85.0%

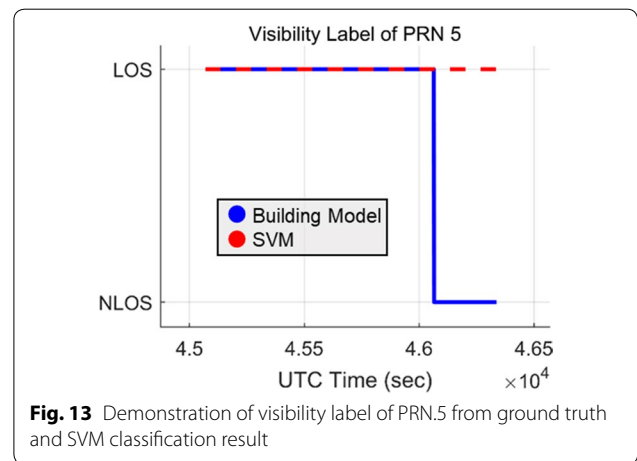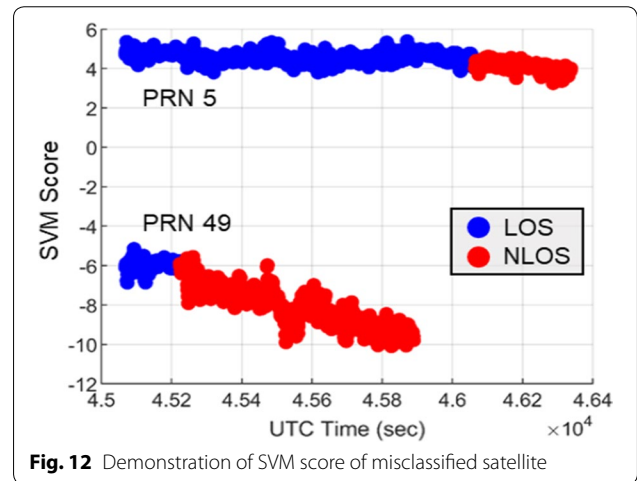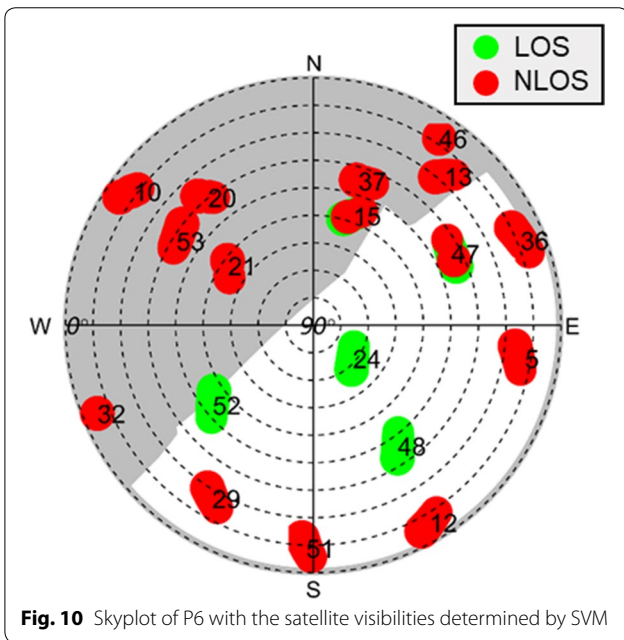**Table 3 Confusion statistic matrix of the experiment using U-BLOX receiver at the Hong Kong urban area**
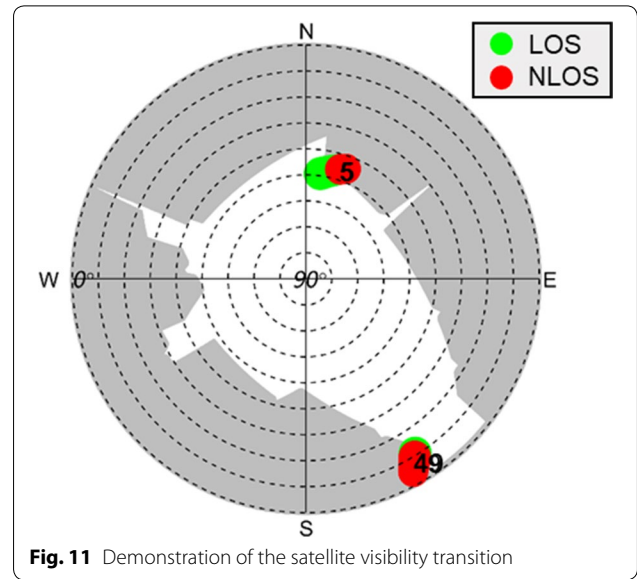
SVM-predicted visibility								
Location	P1		P2		P3		P4	
Actual Visibility	LOS	NLOS	LOS	NLOS	LOS	NLOS	LOS	NLOS
LOS	98.0%	32.1%	84.8%	2.2%	75.1%	24.8%	54.8%	18.3%
NLOS	2.0%	67.9%	15.2%	97.8%	24.9%	75.2%	45.2%	81.7%
Location	P5		P6		P7		P8	
Actual Visibility	LOS	NLOS	LOS	NLOS	LOS	NLOS	LOS	NLOS
LOS	49.4%	23.1%	44.3%	47.2%	95.1%	32.7%	72.9%	14.6%
NLOS	50.6%	76.9%	55.7%	52.8%	4.9%	67.3%	27.1%	85.4%

classification, and the SNR classification still maintain a classification rate with 81.2%. Thus, the SNR classifier could be integrated with the SVM classification to provide a basic classification in the future. For the experiment at P7, it is a deep urban environment with a 95.1% LOS classification rate and 67.3% for NLOS satellites. The main reason is that the amount of LOS satellite is very limited while satellites' elevation angles are very high. At the same time, there are some NLOS satellites with high elevation angle, which are likely to be misclassified as LOS satellites. By considering the SVM's performance with respect to the elevation angle, the SVM model is always taking the high elevation angle satellite as LOS satellite and taking the low elevation angle satellite as NLOS, as shown in Figs. 9 and 10. From the eight locations' dataset, the effected range in elevation angle is nearly from 0 to 30 degrees in low elevation angle and from 60 to 90 degrees in high

elevation angle. We call this case as elevation mask angle of SVM. The main reason of this misclassification is that the most of high elevation satellites from the real measurements are LOS satellite and used for training the SVM model, similarly with the NLOS satellite.

Another case that SVM classifier always failed is the transition of satellite visibility as shown in Fig. 11. However, the classification score of SVM model is too far to reach separating hyperplane and change the classification result in a short time, as shown in Fig. 12. The LOS/NLOS labels by 3D building model and its comparison with SVM model is shown in Fig. 13.

If the two cases (elevation mask angle and visibility transition) area excluded, the classification rate of SVM classifier can be improved greatly, which is shown in Table 4. The classification rates are nearly 90% for the most of cases, which presents these issues are very important to SVM classification.

Xu *et al. Satell Navig*   (2020) 1:15

Page 8 of 12

Fig. 9 Skyplot of P6 with the satellite visibilities labelled from ground truth

Fig. 10 Skyplot of P6 with the satellite visibilities determined by SVM

Fig. 11 Demonstration of the satellite visibility transition

Fig. 12 Demonstration of SVM score of misclassified satellite

Fig. 13 Demonstration of visibility label of PRN.5 from ground truth and SVM classification result

Weighted robust estimation results

The mean error of different positioning approaches, including position solution from the commercial receiver National Marine Electronics Association (NMEA) solutions, conventional weighted least square based on open source software goGPS 0.4.3 and the proposed weighted robust estimator (WRE) are

Xu *et al. Satell Navig* (2020) 1:15

Page 9 of 12

Table 4 The classification rate of SVM classification and SVM classification without issues

Location	P1	P2	P3	P4	P5	P6	P7	P8
SVM	84.6%	91.5%	73.2%	63.9%	63.0%	45.4%	67.6%	85.0%
SVM^w	87.5%	94.3%	92.5%	89.3%	87.2%	79.6%	68.3%	90.4%

Table 5 The mean error of different positioning approaches (m)

Location	P1	P2	P3	P4	P5	P6	P7	P8
WLS	26.40	66.14	52.19	25.74	28.90	26.78	41.10	55.55
WRE	3.03	9.43	30.58	16.07	23.94	7.76	19.50	16.42
NMEA	3.48	4.27	20.23	11.57	11.34	8.71	78.89	12.40

Table 6 The standard deviation of different positioning approaches (m)

Location	P1	P2	P3	P4	P5	P6	P7	P8
WLS	14.78	24.03	31.13	10.69	12.22	26.56	17.52	19.72
WRE	1.96	8.85	21.92	11.09	21.71	4.59	1.66	11.35
NMEA	1.15	3.03	12.82	3.94	2.00	3.52	29.12	10.60

compared in Table 5. The standard deviation of different positioning approaches is also shown in Table 6.

From Tables 5 and 6, the weighted robust estimator has a great improvement in mean error over the conventional least square approach for all tests, especially in light urban scenarios (P1, P2). For the test of P1, WRE has an error of 3.03 m in mean and 1.96 m in standard deviation, while the NMEA has a mean error of 3.48 m. Similar performance can be found in P6, where WRE provides a better performance than NMEA solution. Both locations have enough LOS satellites with 7 LOS satellites in P1 and P6. Moreover, WRE has a better performance in deep urban location (P7) with 19.5 m in mean error, where the NMEA solution has an error of 78.89 m. There are only two LOS satellites at P7, which makes the conventional least square has an error of 41.1 m. In this case, the performance of WRE is related to the sky view visibility. The sky plots from ground truth locations at P1 and P7 are shown in Figs. 14 and 15, respectively.

Fig. 14 Sky plot from ground truth label at P1

Shadow matching results with the different initialization approaches

The searching area of the shadow matching is a 40 m × 40 m square with the initial positioning solution as the center, and the observed satellite visibility is based on the SVM classification. The mean error and across street error of different positioning initialization of shadow matching is shown in Tables 7 and 8, and 'WLS', 'WRE' and 'GT' means weighted least square solution,
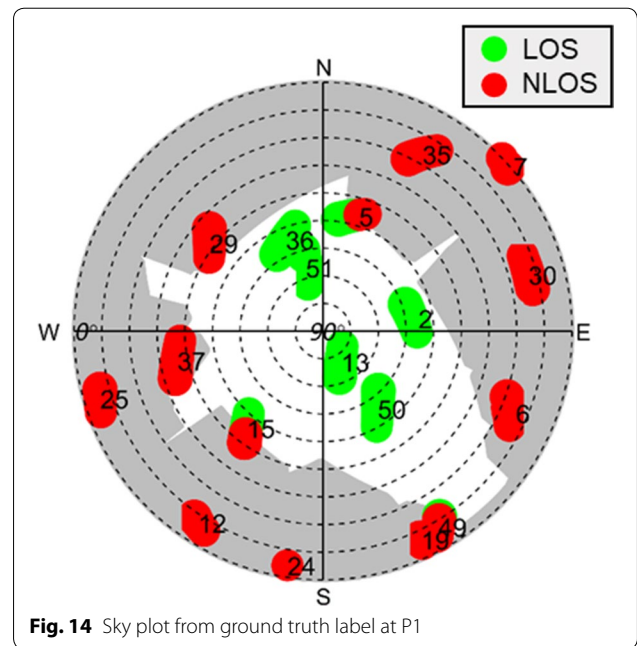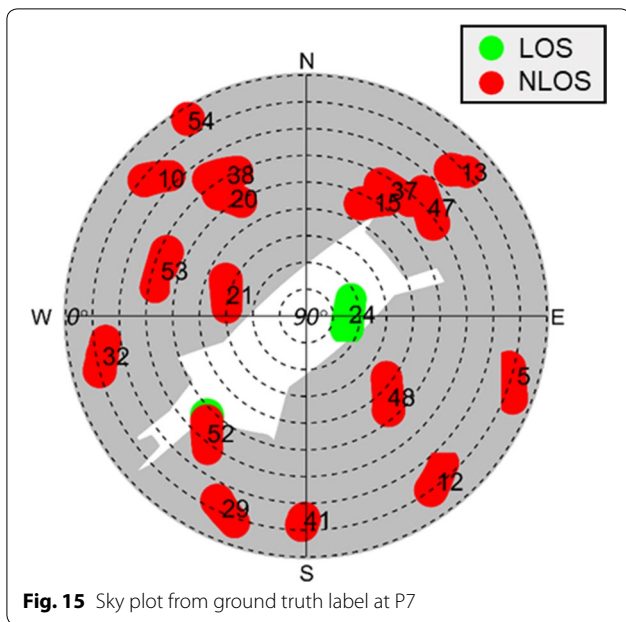
weighted robust estimation solution and the ground truth for the shadow matching initialization.

According to the shadow matching results, the robust result is close to the ground truth when the classification rate is high, and the initial solution is accurate. Therefore, the initial position guess is essential for shadow matching. At P7, the initial position from NMEA is far

Xu *et al. Satell Navig*   (2020) 1:15

Page 10 of 12

Fig. 15 Sky plot from ground truth label at P7

away from the ground truth, thereby it is impossible to get the usable results from shadow matching. At P6, the shadow matching error between ground truth and other approaches is similar, since there are similar building geometries around the initial guess. In the other words, the multi-modal issue at P6. Moreover, its SVM classification rate is around 45%, which gives the wrong score for the particles. In this case, it could reveal that the classification for shadow matching affects its results directly. Similarly, at P4 and P5, the shadow matching failed with low classification rate around 60%. In summary, a threshold of classification rate is necessary for the

shadow matching to indicate the availability of shadow matching. From Table 8, although the mean error of the shadow matching performs not well enough, the error in the across street direction decreases with an accurate initial position, which is the unique advantage of shadow matching.

On the other hand, the shadow matching performance is relied on the signal classification rate. Therefore, the comparison between different initialization approaches cannot be evaluated when the shadow matching classification rate is low. For example, at P4, mean errors of initializations obtained using NMEA and robust estimator are 16.07 m and 11.57 m, respectively. However, mean errors of shadow matching initialized by NMEA and robust estimator are 46.73 m and 47.01 m, respectively, with a classification rate of 63.9%. Therefore, it is difficult to evaluate the performance of NMEA and robust estimator in this case. The shadow matching also applied to different initial positioning solutions with real satellite visibility. The mean error and across street error of different positioning initialization of shadow matching are shown in Tables 9 and 10.

From Tables 9 and 10, all initialization approaches have a prefect classification accuracy on satellite visibility, thus WRE and NMEA solution have the comparative error in terms of mean error and the error in the across-street direction at P1 and P2. However, the standard deviation error of SPP from the robust solution and NMEA are 8.85 m and 3.03 m, respectively. It is because the search area of these two approaches contains the similar high possibility particles of shadow matching, which is illustrated in Figs. 16 and 17. Therefore, it could be stated that when the initial position error is within a range, shadow matching shows similar performances.

Table 7 The mean error of SVM classification based shadow matching initialized by different approaches (m)

Location		P1	P2	P3	P4	P5	P6	P7	P8
Initial	WLS	49.38	83.48	53.45	40.79	40.50	20.64	41.55	41.31
guess	WRE	10.72	10.27	45.71	46.73	45.80	21.75	29.19	28.54
of	NMEA	11.50	10.00	28.03	47.01	37.94	20.70	66.87	17.31
SDM	GT	9.04	9.21	17.63	34.88	35.87	21.48	6.46	19.41

Table 8 The across error of SVM classification based shadow matching initialized by different approaches (m)
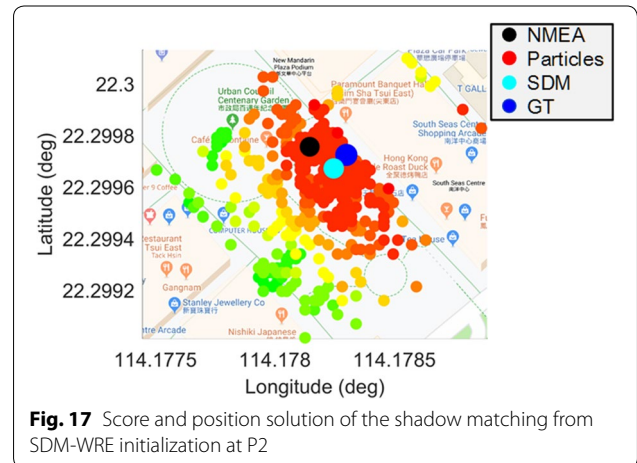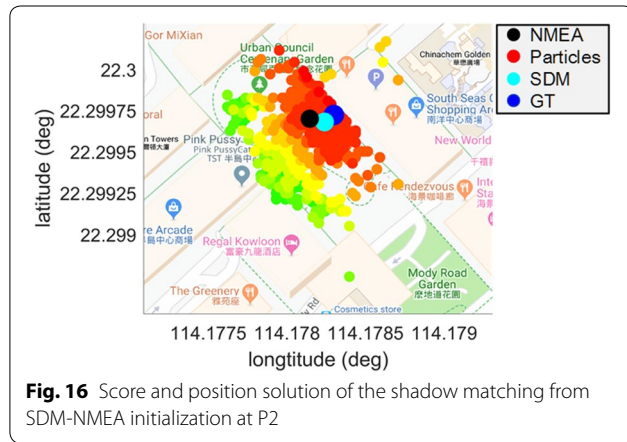
Location		P1	P2	P3	P4	P5	P6	P7	P8
Initial	WLS	14.18	12.57	26.73	7.42	10.58	6.24	13.41	16.95
guess	WRE	3.97	1.44	4.27	9.93	16.92	2.09	2.45	8.13
of	NMEA	4.47	1.91	6.12	6.73	4.43	2.18	10.85	6.09
SDM	GT	3.38	1.77	2.08	7.30	3.09	1.38	3.86	2.56

Table 9 The mean error of real satellite visibility based shadow matching initialized by different approaches (m)

Location		P1	P2	P3	P4	P5	P6	P7	P8
Initial	WLS	52.08	81.99	52.65	41.92	44.90	20.96	42.14	42.57
guess	WRE	11.03	8.16	28.76	49.58	34.84	21.25	25.96	25.62
of	NMEA	11.01	8.12	23.61	39.79	28.02	20.31	65.47	14.22
SDM	GT	11.35	7.95	17.53	24.12	23.08	20.99	23.70	18.11

Table 10 The across error of real satellite visibility based shadow matching initialized by different approaches (m)

Location		P1	P2	P3	P4	P5	P6	P7	P8
Initial	WLS	15.90	12.83	26.21	8.26	10.66	6.41	13.76	16.42
guess	WRE	3.57	1.87	0.45	12.31	20.10	2.29	2.15	13.31
of	NMEA	2.97	1.99	3.22	8.97	3.13	2.64	10.91	4.65
SDM	GT	3.60	1.75	2.36	7.13	2.39	1.93	2.76	5.50

Fig. 16 Score and position solution of the shadow matching from SDM-NMEA initialization at P2

Fig. 17 Score and position solution of the shadow matching from SDM-WRE initialization at P2

Conclusion and future work

In this paper, the shadow matching is integrated with the proposed SVM LOS/NLOS classifier and a robust SPP. For the SVM classification, it is discussed the satellite visibility-related features as SNR, normalized pseudorange residual, elevation angle and pseudorange rate consistency. Moreover, the major problems of SVM classification are identified as elevation mask angle and visibility transition. For the integrated shadow matching, the robust SPP can provide an improved initial position compared to NMEA solution. Experimental results show that the SVM classifier has achieved a classification rate of 91.5%. However, there are misclassifications in some urban scenarios, with a classification rate of 45.4%. By excluding problematic satellites, the SVM classification shows the potential to supply a stable classification rate around 90% at different urban scenarios. The robust SPP also provides a solution of 3.03 m in mean error in urban area. Additionally, the improved shadow matching provides a 10.27 m in mean error and 1.44 m error in cross street direction with the SVM classifier. With the prefect classification, the shadow matching with robust SPP provides 1–3 m accuracy in the cross-street direction, which could be useful, for instance, in e-hailing app. In the future, the SVM classifier could be integrated with other LOS/NLOS classifiers to provide the confidence coefficient of the current classified satellite visibility. Moreover, the confidence coefficient of classified result could be applied into shadow matching to mitigate the impact from the misclassification. The weighted robust estimator can be improved with the satellite visibility from the SVM classification.

Xu *et al. Satell Navig*   (2020) 1:15

Page 12 of 12

Author details
[1] The Hong Kong Polytechnic University, Kowloon, Hong Kong, People's Republic of China. [2] The University of Benevento G. Fortunato, Benevento, Italy. [3] The University of Naples Parthenope, Naples, Italy.

References
Adjrad, M., & Groves, P. D. (2017). Enhancing least squares GNSS positioning with 3D mapping without accurate prior knowledge. *Journal of The Institute of Navigation, 64*(1), 75–91. https://doi.org/10.1002/navi.178.

Adjrad, M., & Groves, P. D. (2018). Intelligent urban positioning: integration of shadow matching with 3D-mapping-aided GNSS ranging. *Journal of Navigation, 71*(1), 1–20. https://doi.org/10.1017/S0373463317000509.

Angrisano, A., Del Pizzo, S., Gaglione, S., Troisi, S., & Vultaggio, M. (2018). Using local redundancy to improve GNSS absolute positioning in harsh scenario. *Acta Imeko, 2018*(7), 16–23.

Brown, R. G. (1993). Global positioning system: Theory and applications. In Spilker Jr, J. J., Axelrad , P., Parkinson, B. W., & Enge, P. (Ed.), *Gps navigation data* (pp. 143–165). The American Institute of Aeronautics and Astronautics. https://doi.org/10.2514/4.866388.

Castaldo, G., Angrisano, A., Gaglione, S., & Troisi, S. (2014). P-RANSAC: An integrity monitoring approach for GNSS signal degraded scenario. *International Journal of Navigation and Observation, 2014*, 1–11. https://doi.org/10.1155/2014/173818.

El-Sheimy, N., & Youssef, A. (2020). Inertial sensors technologies for navigation applications: State of the art and future trends. *Satellite Navigation, 1*, 2. https://doi.org/10.1186/s43020-019-0001-5.

Gaglione, S., Innac, A., Pastore Carbone, S., Troisi, S., & Angrisano, A. (2017). Robust estimation methods applied to GPS in harsh environments. 2017 European Navigation Conference (ENC). *Lausanne, 2017*, 14–25. https://doi.org/10.1109/EURONAV.2017.7954169.

Groves, P. D. (2011). Shadow matching: a new GNSS positioning technique for urban canyons. *Journal of Navigation, 64*(3), 417–430.

Groves, P. D., Wang, L., Adjrad, M., & Ellul, C. (2015). GNSS shadow matching: The challenges ahead. In *Proceedings of the ION GNSS+ 2015, 14–18 September 2015, Tampa, Florida, The Institute of Navigation.*

Hsu, L. (2017). GNSS multipath detection using a machine learning approach. In: *IEEE 20th international conference on intelligent transportation systems (ITSC)* 2017 (pp. 1–6), Yokohamaz. doi:10.1109/ITSC.2017.8317700.

Hsu, L., Tokura, H., Kubo, N., Gu, Y., & Kamijo, S. (2017). Multiple faulty GNSS measurement exclusion based on consistency check in urban canyons. *IEEE Sensors Journal, 17*(6), 1909–1917. https://doi.org/10.1109/JSEN.2017.2654359.

Ji, S., Chen, W., Ding, X., Chen, Y., Zhao, C., & Hu, C. (2010). Potential benefits of GPS/GLONASS/GALILEO integration in an urban Canyon—Hong Kong. *Journal of Navigation, 63*(4), 681–693. https://doi.org/10.1017/S0373463310000081.

Knight, N. L., & Wang, J. (2009). A comparison of outlier detection procedures and robust estimation methods in GPS positioning. *Journal of Navigation, 62*(4), 699–709. https://doi.org/10.1017/S0373463309990142.

Kuusniemi, H., Lachapelle, G., & Takala, J. H. (2004). Position and velocity reliability testing in degraded GPS signal environments. *GPS Solutions, 8*(4), 226–237. https://doi.org/10.1007/s10291-004-0113-7.

Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection.* New York: Wiley.

Tay, S., & Marais, J. (2013). Weighting models for GPS Pseudorange observations for land transportation in urban canyons. In: *Proceedings of the 6th European workshop on GNSS signals and signal processing,* Munich, Germany, pp. 1–6.

Tsakiri, M., Stewart, M., Forward, T., Sandison, D., & Walker, J. (1998). Urban fleet monitoring with GPS and GLONASS. *Journal of Navigation, 51*(3), 382–393. https://doi.org/10.1017/S0373463398007929.

Viandier, N., Nahimana, D. F., Marais, J., & Duflos, E. (2008). Gnss per-formance enhancement in urban environment based on pseudo-rangeerror model. In: *Proceedings of IEEE/ION position, location navigation symposium, Monterey, CA,* pp. 377–382. doi: 10.1109/PLANS.2008.4570093.

Wang, L. (2014). Kinematic GNSS shadow matching using a particle filter. In: *Proceedings of the 27th international technical meeting of the satellite division of the institute of navigation (ION GNSS+ 2014), Tampa, Florida,* pp. 1907–1919.

Wang, L., Groves, P. D., & Ziebart, M. K. (2015). Smartphone shadow matching for better cross-street GNSS positioning in urban environments. *Journal of Navigation, 68*(3), 411–433. https://doi.org/10.1017/S0373463314000836.

Xu, H., Zhang, G., Xu, B & Hsu, L. T. (2018). GNSS shadow matching based on intelligent LOS/NLOS Classifier. *Paper presented at the The 16th world congress of the international association of institutes of navigation (IAIN) Chiba, Japan.*

Yozevitch, R., & Moshe, B. B. (2015). A robust shadow matching algorithm for GNSS positioning. *Journal of Navigation, 62*(2), 95–109. https://doi.org/10.1002/navi.85.

Yang, Y.X., Mao, Y., & Sun, B.J. (2020). Basic performance and future developments of BeiDou global navigation satellite system. *Satellite Navigation, 1*, 1. https://doi.org/10.1186/s43020-019-0006-0.

Yozevitch, R., Moshe, B. B., & Weissman, A. (2016). A robust GNSS LOS/NLOS signal classifier. *Journal of Navigation, 63*(4), 429–442. https://doi.org/10.1002/navi.166.