

Research Article

A Unified Framework for GPS Code and Carrier-Phase Multipath Mitigation Using Support Vector Regression

Quoc-Huy Phan,¹ Su-Lim Tan,² Ian McLoughlin,³ and Duc-Lung Vu¹

¹ University of Information Technology, Km 20, Ha Noi Highway, Linh Trung Ward, Thu Duc, HCMC 70000, Vietnam

² Singapore Institute of Technology, 25 North Bridge Road, Singapore 179104

³ School of Information Science and Technology, University of Science and Technology of China, No. 443 Huangshan Road, Hefei, Anhui 230027, China

Correspondence should be addressed to Quoc-Huy Phan; huyppq@uit.edu.vn

Received 1 October 2012; Accepted 14 January 2013

Academic Editor: Paolo Gastaldo

Copyright © 2013 Quoc-Huy Phan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multipath mitigation is a long-standing problem in global positioning system (GPS) research and is essential for improving the accuracy and precision of positioning solutions. In this work, we consider multipath error estimation as a regression problem and propose a unified framework for both code and carrier-phase multipath mitigation for ground fixed GPS stations. We use the kernel support vector machine to predict multipath errors, since it is known to potentially offer better-performance traditional models, such as neural networks. The predicted multipath error is then used to correct GPS measurements. We empirically show that the proposed method can reduce the code multipath error standard deviation up to 79% on average, which significantly outperforms other approaches in the literature. A comparative analysis of reduction of double-differential carrier-phase multipath error reveals that a 57% reduction is also achieved. Furthermore, by simulation, we also show that this method is robust to coexisting signals of phenomena (e.g., seismic signals) we wish to preserve.

1. Introduction

Multipath is defined as one or more indirect replicas of the line-of-sight (LOS) signal from satellites arriving at a receiver's antenna from a satellite. It normally occurs due to reflection from objects in the vicinity of the receiver and constitutes a major error source that contaminates receivers' measurements, resulting in performance degradation of GPS positioning solutions. The errors induced by multipath are typically up to 15 meters for C/A code [1] and a few centimeters for carrier-phase measurements [2]. Multipath mitigation is hence important for a variety of applications which utilize this data, such as ionospheric monitoring [3], geodesy [4, 5], and navigation [6].

On the one hand, multipath mitigation is a very challenging task. As multipath is site-dependent, differencing measurements among multiple short-baseline receivers (DGPS) are unlikely to help. Furthermore, aggressively removing multipath error may harm wanted coexisting information and perturbations such as seismic signals induced by an

earthquake, as their frequency spectra likely overlap with that of the contaminating multipath error.

Various mitigation approaches have been proposed in the literature, classified into either frequency-domain or time-domain processing. The former is based on spectral analysis of multipath error in the frequency domain using fast Fourier transform (FFT) [7], or wavelet decomposition [8, 9]. However, they unintentionally rule out other coexisting signals. To overcome this issue, signal-to-noise (SNR) measurements [10, 11] can be used as alternative for analysis. Unfortunately, this suffers from unavailability and inconsistency in units from different types of receivers. Time-domain methods range from the popular carrier smoothing filter (CSF) [12, 13], band-pass finite impulse response (FIR) filter [14, 15] to stacking [16–18]. These methods also tend to filter out coexisting signals of interest and require high-rate data to boost their performance [16].

In previous works [19, 20], we propose a regression model, which integrates kernel support vector regression (SVR) with geometrical features to deal with the code multipath error

prediction on ground fixed GPS stations. To the best of the authors' knowledge, this is the first work using machine learning to address GPS multipath error estimation. This paper extends our previous work to define a unified framework for both code and carrier-phase multipath mitigation. The contribution of this paper is threefold: (1) deriving geometric models for code and carrier-phase multipath errors, (2) formulating multipath error estimation as a regression problem with geometrical features, and (3) unifying the framework for code and carrier-phase multipath estimation using support vector regression.

The rest of this paper is organized as follows. Section 2 briefly reviews the mathematical models of GPS measurements and derives the geometrical models of multipath errors. By posing multipath estimation as a nonlinear regression problem, Section 3 defines the framework for multipath mitigation. Experimental results and discussions for code and carrier-phase multipath mitigation will be presented in Sections 4 and 5, respectively. The conclusion will follow in Section 6.

2. GPS Measurements and Code Multipath Extraction

In this section, we briefly review the GPS measurement data as generated by GPS receivers. Following this, geometrical models will be derived for both code and carrier-phase multipath errors.

2.1. GPS Measurements. The code measurement ρ_1 and carrier-phase measurement ϕ_1 for channel L1 are given as in (1) and (2), respectively [13, 21]. The measurements for L2 are similar,

$$\rho_1 = r + c(\delta_u - \delta_s) + I_1 + T + M_1^p + \varepsilon_1^p, \quad (1)$$

$$\phi_1 \lambda_1 = r + c(\delta_u - \delta_s) - I_1 + T + N_1 \lambda_1 + M_1^\phi + \varepsilon_1^\phi, \quad (2)$$

where r represents the true range from a satellite to a receiver, δ is clock bias, the subscripts u and s refer to the user (receiver) and the satellite, respectively, c is the speed of light, T , I , M^p , ε^p , M_1^ϕ , and ε_1^ϕ denote tropospheric delay, ionospheric delay, code multipath error, random receiver noise on code, carrier-phase multipath error, and random receiver noise on carrier-phase, respectively, and the symbols λ_1 denote wavelengths of L1. The term N_1 is the *ambiguous integer* of L1. The opposite signs of the ionospheric delays in (1) and (2) are due to the fact that the ionosphere affects code and carrier measurements equally but in opposite directions when the signals travel through the dispersive ionospheric layer in the atmosphere [2].

2.2. Geometrical Model of Code Multipath. Ideally, in a multipath-free environment, only one direct signal is received by the antenna from each satellite. However, no environment is completely multipath-free in practice. A receiver antenna receives one or more replicas of the direct signal reflected from objects near the LOS path, particularly those in the

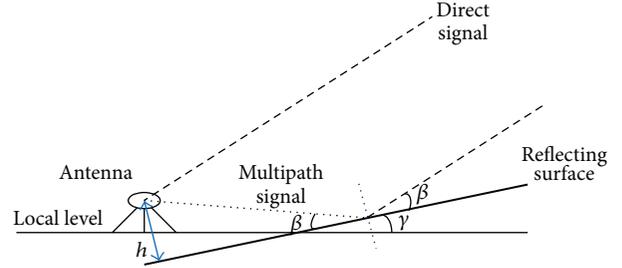


FIGURE 1: The direct signal and one multipath signal with simplified geometry.

vicinity of the receiver. As a result, the receiver will track a composite signal that is a combination of the direct path and the multipath replicas.

For clarity, let us consider the simplified case of one multipath signal. Let A_d and A_m denote the amplitudes of the direct signal and the multipath signal, respectively, δ the path delay, ψ the multipath relative phase in radians, $\alpha = A_m/A_d \leq 1$ the ratio of the multipath and direct amplitudes, and θ and μ the azimuth and elevation angles, respectively. The position of any reflecting object is described as a planar surface tilted relative to the local level with a tilt angle γ at a distance h from the antenna centre as illustrated in Figure 1.

Let β denote the reflection angle relative to the reflecting surface; the induced code multipath error is given as [16]:

$$M^p = \frac{\alpha \delta \cos \psi}{1 + \alpha \cos \psi}. \quad (3)$$

In order to obtain the geometrical model, we need to relate multipath error with geometrical parameters, that is, azimuth and elevation angles. Assuming that the satellite, antenna, and normal vector to the reflecting surface are coplanar, multipath reflections fall into two categories: *forward-scatter* and *backscatter* [22] as in Figure 2.

In both forward-scatter and backscatter scenarios, it is easily to obtain

$$\delta = 2h \sin \beta, \quad (4)$$

$$\psi = \frac{2\pi}{\lambda} \delta = \frac{2\pi}{\lambda} 2h \sin \beta. \quad (5)$$

Furthermore, using the convention of angles measured anti-clockwise over the interval $[0^\circ 180^\circ]$, $\beta + \gamma = \theta$ for forward-scatter whereas $\beta + \theta = \gamma$ for backscatter. To generalize,

$$\beta = |\theta - \gamma|. \quad (6)$$

Substituting (4), (5), and (6) into (3), the code multipath equations corresponding to one reflecting signal can be rewritten as

$$M^p = \frac{\alpha 2h \sin(|\theta - \gamma|) \cos((4\pi h/\lambda) \sin(|\theta - \gamma|))}{1 + \alpha \cos((4\pi h/\lambda) \sin(|\theta - \gamma|))}. \quad (7)$$

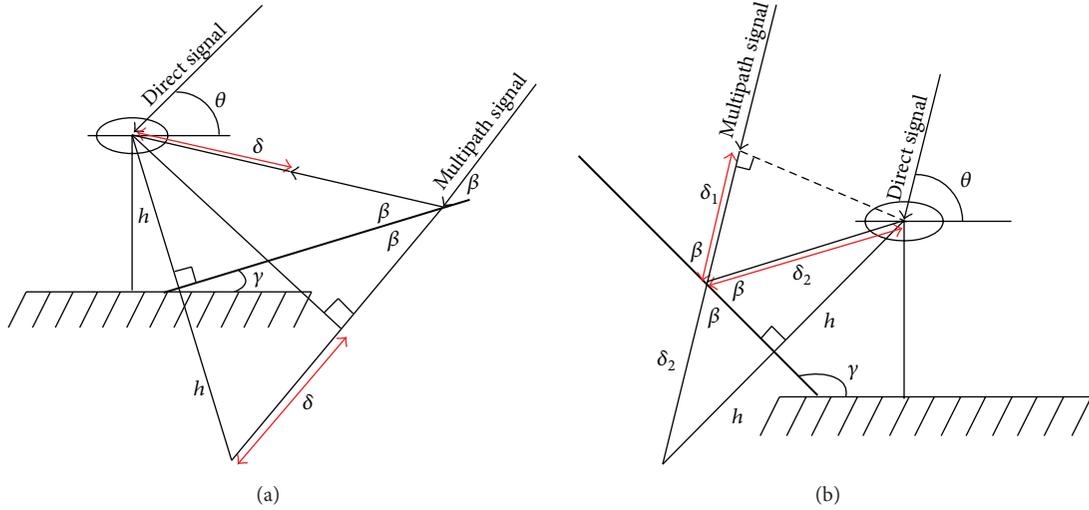


FIGURE 2: Multipath reflections: (a) forward-scatter and (b) backscatter.

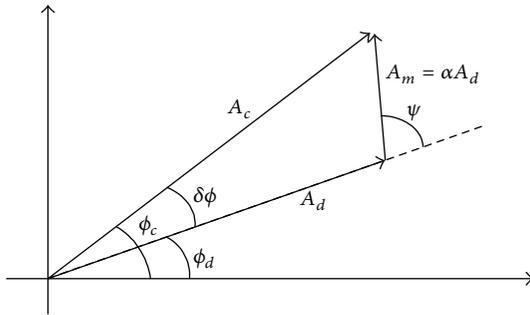


FIGURE 3: Effect of multipath to carrier-phase measurement.

In the general case of m reflecting signals, the total code multipath is the sum of the individual code multipath:

$$M^p = \frac{\sum_{i=1}^m \alpha_i 2h_i \sin(|\theta - \gamma_i|) \cos((4\pi h_i/\lambda) \sin(|\theta - \gamma_i|))}{1 + \sum_{i=1}^m \alpha_i \cos((4\pi h_i/\lambda) \sin(|\theta - \gamma_i|))}. \quad (8)$$

2.3. Geometrical Model of Carrier-Phase Multipath. The effect of multipath on carrier-phase measurement can be demonstrated by the phasor diagram in Figure 3. Additionally, let A_c denote the amplitude of the composite signal which is the combination of the multipath signal and the direct signal. ϕ_d , ϕ_c , and ψ are the direct signal's phase, the composite signal's phase and the multipath relative phase with respect to the direct signal, respectively. δ_ϕ denotes the phase error due to multipath.

The phase error due to multipath is easily derived in terms of multipath parameters as follows:

$$\tan(\delta_\phi) = \frac{A_m \sin(\psi + \phi_0)}{A_d + A_m \cos(\psi + \phi_0)} = \frac{\alpha \sin(\psi + \phi_0)}{1 + \alpha \cos(\psi + \phi_0)}, \quad (9)$$

where ϕ_0 is a possible phase offset at time 0. Therefore,

$$\delta_\phi = \arctan\left(\frac{\alpha \sin(\psi + \phi_0)}{1 + \alpha \cos(\psi + \phi_0)}\right). \quad (10)$$

Substituting (5) into (10), we obtain the phase error caused by one multipath signal:

$$\delta_\phi = \arctan\left(\frac{\alpha \sin((4\pi h/\lambda) \sin(|\theta - \gamma|) + \phi_0)}{1 + \alpha \cos((4\pi h/\lambda) \sin(|\theta - \gamma|) + \phi_0)}\right). \quad (11)$$

In general case of m multipath signals,

$$\delta_\phi = \arctan\left(\frac{\sum_{i=1}^m \alpha \sin((4\pi h_i/\lambda) \sin(|\theta - \gamma_i|) + \phi_0^i)}{1 + \sum_{i=1}^m \alpha \cos((4\pi h_i/\lambda) \sin(|\theta - \gamma_i|) + \phi_0^i)}\right). \quad (12)$$

3. Multipath Mitigation Framework Using Nonlinear Support Vector Regression

In this section, we will define the unified framework for code and carrier-phase multipath mitigation followed by the formulation to solve multipath estimation as nonlinear regression using support vector regression algorithm.

3.1. Multipath Mitigation Framework. Under assumption that the multipath environment around a ground fixed receiver is held fixed, then quantities α_i , h_i , γ_i in (8) and (12) will remain constant. As a result, code multipath error M^p and carrier-phase multipath error δ_ϕ are complicated functions of satellite-relative elevation angle θ and azimuth angle μ in the general case which characterizes the geometry of the satellite with respect to the receiver. Give these functions, in order to compute multipath errors, it is necessary to somehow evaluate the functions in (8) and (12) given as an azimuth/elevation pair. Viewed as a regression problem, these functions will be approximated by learning from the historical multipath

data and satellite orbital information. Specifically, estimating code and carrier-phase multipath errors is a 2-dimensional regression setting:

$$\text{multipath} = f(\text{azimuth}, \text{elevation}). \quad (13)$$

The rationale behind this approach is the observation that a fixed receiver experiences highly sidereal day-to-day correlation of satellite-receiver geometry and multipath error. It is well known that the GPS satellite orbits were selected to have a period of half a sidereal day (23 hours 56 minutes 4 seconds) with a daily repeating ground track [23, 24]. Because of this, satellite visibility from any point on earth is the same from day to day, with the satellites appearing in their positions approximately 4 minutes or 236 seconds earlier each day due to the time difference between the sidereal and solar day. Furthermore, for a fixed station, its surroundings and the antenna usually do not change significantly across consecutive days. Therefore, GPS multipath signals are expected to largely repeat over the same time period. Our proposed approach leverages this observation with a regression approach to approximate the function of the repeatable multipath error.

Code and carrier-phase multipath errors can be estimated independently, but the procedure for training the estimators is the same. Satellite-specific multipath estimator will be trained using ε -SVR [25], a well-known support vector regression algorithm. After being trained, the code and carrier-phase multipath estimators will be used to provide estimate of multipath errors. Afterwards, these estimates will be used to correct code and carrier-phase measurements to achieve multipath mitigation.

3.2. Nonlinear Regression for Multipath Estimation. Among various kinds of regression models, SVRs [25, 26] are well known to have more potential advantages than traditional models, such as neural networks [27]. They are based on the strong statistical learning theory and have been shown to be less prone to overfitting as well as being independent of the dimensionality of the input space. We employ ε -SVR [25], which is most commonly used for regression problem. In the hard margin loss setting for ε -SVR, the estimate has at most ε deviation from the actual target for all training data, meaning that one does not care about errors as long as they are less than ε , but will not accept any deviation larger than this. Nevertheless, owing to noise in data, we usually use soft margin loss setting to allow for some errors by introducing slack variables in the formulation of ε -SVR.

We denote input vector as $x \in \mathfrak{R}^n$ and denote a value of multipath error as $y \in \mathfrak{R}$. Given the multipath model in (8) and (12), our goal is to learn a function $f: \mathfrak{R}^n \mapsto \mathfrak{R}$ mapping from an observation vector x to an estimate of multipath error \hat{y} . Formally, this can be accomplished by first choosing a set of N training samples $\{(x_1, y_1), \dots, (x_N, y_N)\} \in \mathfrak{R}^n \times \mathfrak{R}$. Due to noise in the training data, it is unlikely that $f(x_i)$ will be equal to y_i for all x_i , so a loss function $L(f(x), y)$ must also be chosen to quantify the penalty for $f(x_i)$ differing from y_i . The estimator f can be found by minimizing the total loss over the training data.

For each satellite, the multipath estimator is trained using ε -SVR [25]. We denote the regression function $f(x) = \langle w, \varphi(x) \rangle + b$, where w is the weight vector in the kernel feature space, $b \in \mathfrak{R}$ is a bias term, $\langle \cdot, \cdot \rangle$ denotes the dot product, and φ is the kernel feature map of data point x . The ε -insensitive loss function given by (14) is chosen so that the function $f(x)$ is found to have at most ε deviation from the targets y_i for all training samples:

$$L(f(x), y) = \begin{cases} 0, & \text{if } |f(x) - y| < \varepsilon, \\ |f(x) - y| - \varepsilon, & \text{otherwise,} \end{cases} \quad (14)$$

$f(x)$ can be solved through the following optimization problem [25]:

$$\begin{aligned} & \text{minimize} && \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \\ & \text{subject to} && \begin{aligned} & y_i - \langle w, \varphi(x_i) \rangle - b \leq \varepsilon + \xi_i \\ & \langle w, \varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* > 0, \quad i = 1, \dots, N, \end{aligned} \end{aligned} \quad (15)$$

where $\varepsilon > 0$ is the parameter of the ε -insensitive loss function that controls the accuracy of the regressor. The constant $C > 0$ adjusts the tradeoff between the regression error and the regularization on f . $\xi = \{\xi_1, \dots, \xi_N\} \in \mathfrak{R}^N$ and $\xi^* = \{\xi_1^*, \dots, \xi_N^*\} \in \mathfrak{R}^N$ are slack variables allowing errors around the regression function. After solving the optimization problem, the form of the estimator is

$$\begin{aligned} f(x) &= \sum_{i=1}^{N_{SV}} \omega_i \langle \varphi(x), \varphi(p_i) \rangle + b \\ &= \sum_{i=1}^{N_{SV}} \omega_i \kappa(x, p_i) + b, \end{aligned} \quad (16)$$

where $\omega_1, \dots, \omega_{N_{SV}}$ are scalar coefficients, $p_1, \dots, p_{N_{SV}}$ are support vectors, and $\kappa: \mathfrak{R}^n \times \mathfrak{R}^n \mapsto \mathfrak{R}$ is a kernel function. $f(x)$ depends only on the training samples having nonzero coefficients (support vectors) through the representation of the kernel function κ . The Gaussian kernel given by (17) is reasonably chosen due to its ability to handle nonlinearity:

$$\kappa(x_i, x_j) = \exp\left(-\eta \|x_i - x_j\|^2\right), \quad (17)$$

where η is kernel bandwidth. Solving the regression problem by ε -SVR, learning from training data, code multipath error can be estimated for each visible satellite.

4. Experiment I: Code Multipath Mitigation

In this section, we will describe experiments conducted to train multipath estimators and subsequently use them for multipath mitigation. We demonstrate that our approach outperforms state-of-the-art results in code multipath mitigation in terms of standard deviation. The advantages of the exploited methods will be also discussed.



FIGURE 4: The observation site on the rooftop of N2 building (NTU campus).

4.1. Experimental Data Set. The data set used for evaluation was recorded at a sampling rate of 0.1Hz from a GPS monitoring station equipped with a Trimble NetRS receiver on the rooftop of the N2 building in Nanyang Technological University (NTU) campus (Singapore) during five consecutive days: from the day of year (Day) 306 to 310 of 2010. The nominal position of the observation site is $(-1507932.6167, 6195587.6757, 148897.9990)$ in the earth-centered, earth-fixed (ECEF) Cartesian coordinate system [13]. Rooftops are usually bad multipath environments since there are often many vents and other reflective objects within the GPS antenna field of view. In the photographs shown in Figure 4, it can be seen that the observation site is surrounded by many buildings and reflectors which make multipath potentially more severe. During the evaluation period there were 31 visible satellites ranging from PRN 2 to PRN 32 observed at this site.

To illustrate the repeatability of GPS satellite geometries, Figure 5 plots the geometries of 4 visible satellites with respect to the observation site during 4 consecutive days from Day 306 to Day 309. For the sake of simplicity and clarity, only 4 of 31 in-view satellites whose full arcs completely fall in each day period are plotted. As seen from these plots, the footprints of the day-to-day repeated geometries of the satellites are obviously exposed.

Likewise, in order to illustrate the day-to-day repeatability of code multipath error, Figures 6(a) and 6(b) plotted the code multipath sequences extracted from the recorded data set during four observation days from Day 306 to Day 309. The code multipath error can be extracted in batch mode from code measurements using code-minus-carrier (CmC) combination on a whole arc [9, 13]. The correlation is more clearly revealed after we smoothed the sequences with a CSF [9] having a 50-second window to largely remove high-frequency noise. The original multipath sequences are plotted in blue and CSF-smoothed multipath sequences are plotted in red. Day-to-day correlation of the two multipath sequences is numerically evaluated by their normalized

TABLE 1: Normalized cross-correlation of PRN 12's multipath sequences for original data shown in blue and for CSF-smoothed data shown in red.

	Day 306	Day 307	Day 308	Day 309
Day 306	N/A	0.6668	0.5215	0.5215
Day 307	0.9055	N/A	0.6513	0.5254
Day 308	0.8775	0.8988	N/A	0.6260
Day 309	0.8649	0.8630	0.8851	N/A

cross-correlation. Pair-wise normalized cross-correlation values of the multipath sequences are tabulated in Table 1 for PRN 12 with blue and red values representing the original and smoothed multipath sequences, respectively. The day-to-day correlation is around 89% between two consecutive days and slowly degrades with time. This is understandable since cumulative environmental changes become noticeable as the time span increases.

4.2. Training Code Multipath Estimators. For each satellite, the training data is prepared using 4-day data from Day 306 to Day 309 extracted from the experimental data set, which we have found to be redundant enough to capture distribution of multipath sequences. Azimuth and elevation angles of the satellites in degrees with respect to the receiver, which are inputs for training, are computed from the broadcast navigation data [13]. For the desired multipath outputs, after being detached from observation data, the CmC sequences containing multipath errors are filtered with CSF to remove high-frequency noise. The effect of this smoothing on the multipath error is negligible as long as the smoothing window is shorter than the highest rate multipath. The smoothing window is set to 50 seconds (equivalent to 5 epochs) which is only a fraction of the shortest anticipated multipath fading period of 200 seconds [9, 28]. Thus, the receiver noise was significantly reduced without removing the multipath which was to be quantified. This smoothing operation helps to

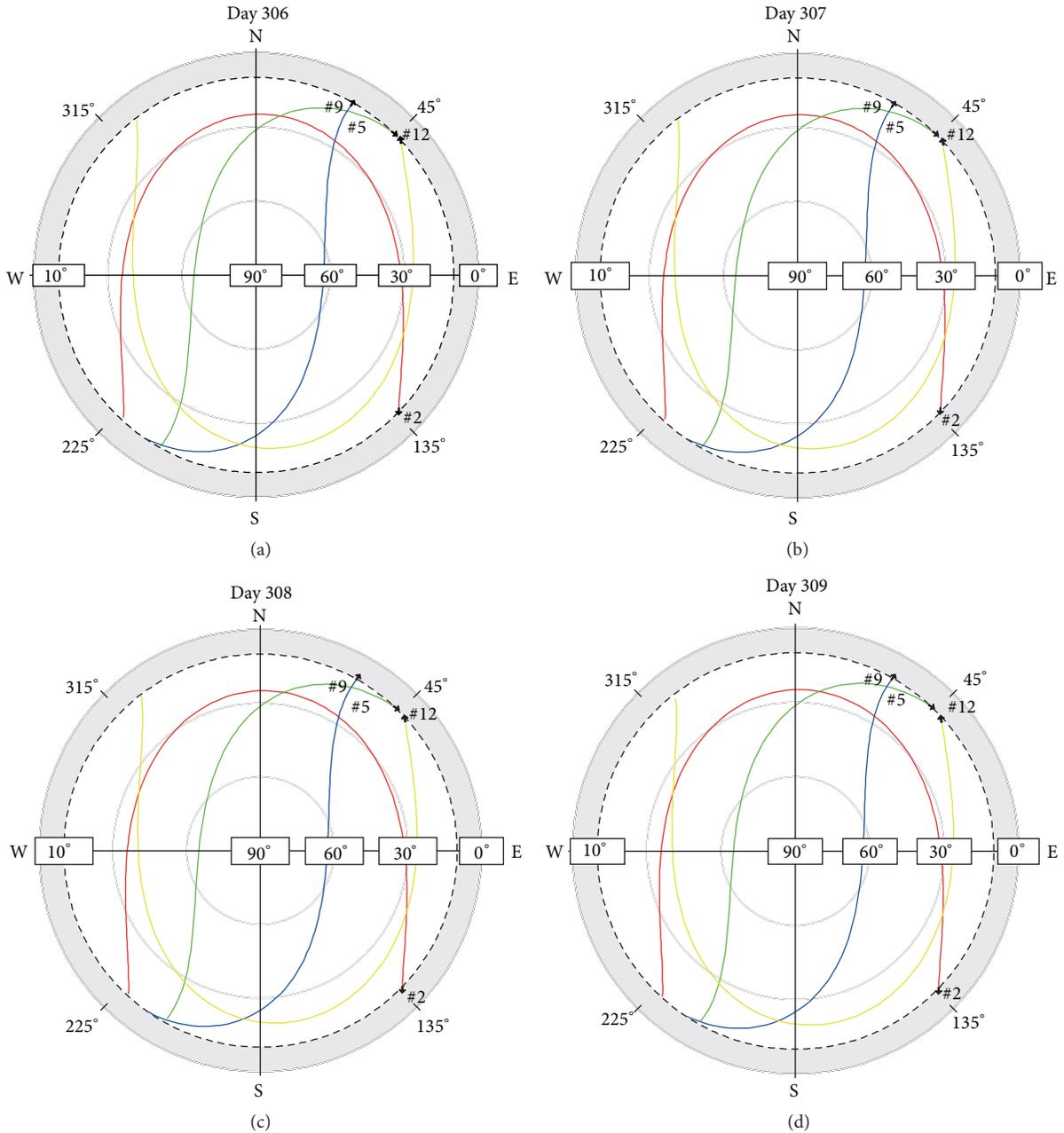


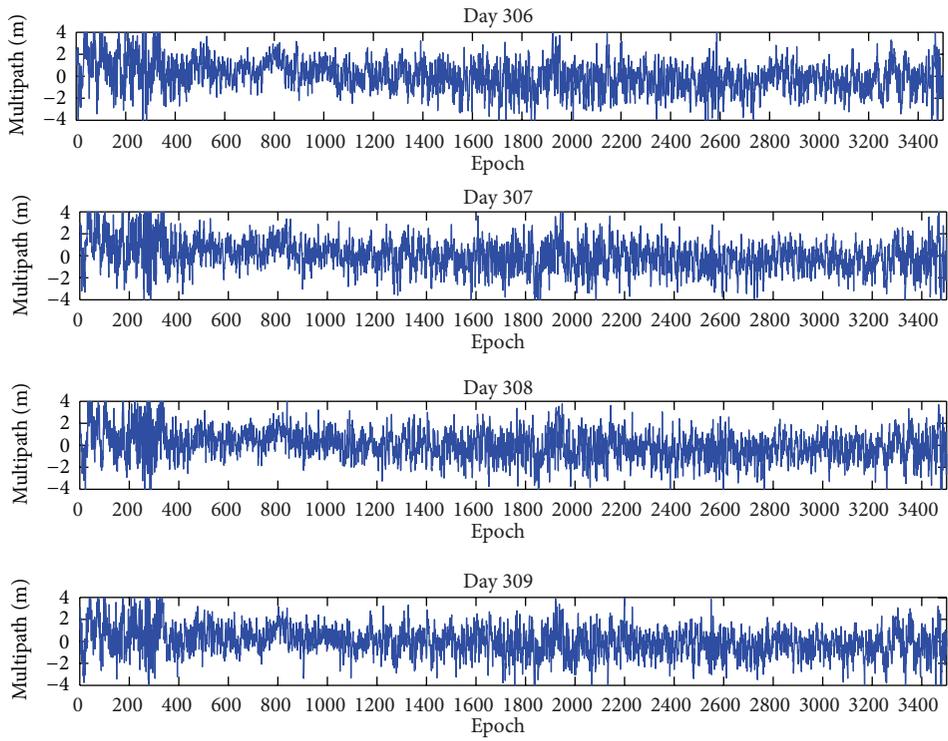
FIGURE 5: Geometry footprints of 4 satellites PRN 02, 05, 09, and 12 in 4 consecutive sidereal days.

clearly expose multipath patterns and, as a result, to enhance the estimators' generalization.

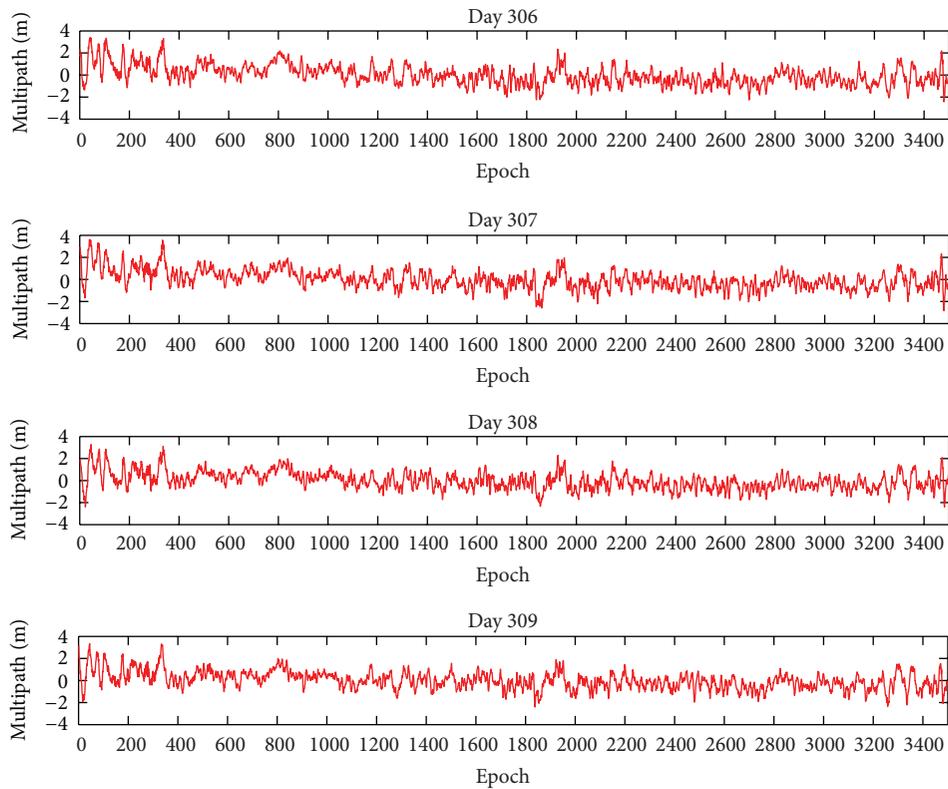
Scaling is applied to the training data before use. Azimuth angles, elevation angles, and code multipath values are scaled to the range $[-1; +1]$. The main advantage of scaling is not only to avoid numerical difficulties during the calculation but also to prevent domination of values in greater numeric ranges over those in smaller numeric ranges.

The *libSVM* package (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>), which implements ε -SVR, was used to find the support vectors and coefficients of each satellite-specific code multipath estimator. ε , the kernel parameter η and the penalty

parameter of the error term C , must be chosen a priori. It is not known beforehand which parameter values are best for a given problem; consequently, some kind of model selection (parameter search) must be done. For η and C , grid search and cross-validation were applied for parameter search. Following the recommended *libSVM* approach, a coarse grid search was firstly performed on exponentially growing sequences of $\eta = 2^{-15}, 2^{-13}, \dots, 2^3$ and $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ followed by cross-validation for each pair of (η, C) with a fixed $\varepsilon = 0.01$. The result with the best 6-fold cross-validation accuracy was picked. With this strategy, better region on the grid can be identified and finer grid-search



(a) Original multipath sequences for PRN 12 over four consecutive days



(b) 50-second CSF-smoothed sequences corresponding to the data in (a)

FIGURE 6: Multipath sequences of PRN 12 in four sidereal days from Day 306 to Day 309 of 2010.

across those regions. ε was searched separately in the range (0.005–0.1) with step size of 0.005 after the best pair (η, C) has been chosen.

Learning from the training data set, the support vectors and coefficients in (16) are found for multipath estimation of each satellite. Each separately trained estimator should estimate multipath error when presented with a new observation of azimuth/elevation angles thereafter.

4.3. Experimental Results. In order to evaluate the performance of code multipath estimators, the proper multipath correction is directly applied to code measurements for Day 310. Note that the inputs must be scaled as they were during the training phase, and the estimated multipath values subsequently need to be descaled.

For the sake of demonstration, Figure 7 presents pseudorange multipath errors and responses of the multipath estimator corresponding to PRN 12 in the Day. As observed, standard deviation of the multipath sequence is significantly reduced after being corrected with the responses of the SVR estimator.

Performance of all multipath estimators corresponding to the visible satellites is tabulated in Table 2 where multipath reduction is measured in terms of the reduced standard deviation of multipath errors. The percentages of reduction range from 68% to 91%. On average, calibrating the data with CSF followed by the SVR estimators gains improvement from 36.68% to 78.99%. With the assumption about unchanged surroundings of this approach, performance of the multipath estimators would depend upon how fast the reflecting surfaces along the propagation direction change. The environmental changes are expected to be different for different propagation directions of the GPS satellites. Therefore, the variation in performance of the estimators as seen in Table 2 is expected.

The goodness of the corrections can also be illustrated in the positional domain. Figure 8 shows the variation of the solved positions from the nominal position of the receiver. Weighted least mean square single-point positioning [13] with broadcast navigation data was applied to the data of Day 310. The measurements were only multipath corrected while other noises and biases (e.g., atmosphere delays, etc.) were not calibrated. The plot reveals noticeably higher centralization of the solution on the data corrected with SVR estimators over those obtained from the original data and the 50-second CSF-smoothed data. The reduction of standard deviation of coordinate time series North, East, and Up is tabulated separately in Table 3.

5. Discussions

For comparison, Table 4 tabulates numerical performance of different methods in terms of percentage of reduced multipath error. The performance of CSF is reported with a 100-second smoothing window [13]. The performance range of the frequency analysis method using fast fourier transform (FFT) [7] is reported with block sizes of 256 and 512, respectively, whilst the block size of the frequency analysis method using wavelet analysis [9] is 100 seconds. In particular, the

TABLE 2: Standard deviation (m) of noise before and after correction applied with CSF and SVR estimators.

PRN	Original	CSF corrected	SVR corrected
02	1.4927	0.9992 (33.06%)	0.3671 (75.41%)
03	1.1479	0.6741 (41.28%)	0.2478 (78.41%)
04	1.4009	0.8913 (36.38%)	0.2201 (84.29%)
05	1.2474	0.7722 (38.10%)	0.2325 (81.36%)
06	1.3194	0.6957 (47.28%)	0.2946 (77.67%)
07	1.2561	0.8768 (30.20%)	0.2711 (78.42%)
08	1.2806	0.8680 (32.22%)	0.2409 (81.19%)
09	1.3406	0.7955 (40.66%)	0.3599 (73.15%)
10	1.2084	0.7954 (34.18%)	0.2193 (81.85%)
11	1.4290	0.9993 (30.07%)	0.2995 (79.04%)
12	1.4813	0.9943 (32.87%)	0.2509 (83.06%)
13	1.5451	0.9960 (35.54%)	0.4969 (67.84%)
14	1.4514	0.8575 (40.92%)	0.2880 (80.16%)
15	1.3496	0.9539 (29.32%)	0.3034 (77.52%)
16	1.2200	0.7730 (36.64%)	0.3902 (68.01%)
17	1.4946	0.9515 (36.34%)	0.2763 (81.51%)
18	1.7889	1.1555 (35.40%)	0.3425 (80.85%)
19	1.1336	0.6832 (39.73%)	0.2506 (77.90%)
20	1.3501	0.8263 (38.80%)	0.2020 (85.04%)
21	1.1886	0.7329 (38.34%)	0.3492 (70.62%)
22	1.2568	0.8065 (35.83%)	0.3058 (75.67%)
23	1.2730	0.8483 (33.36%)	0.3378 (73.47%)
24	1.4779	0.9000 (39.11%)	0.1706 (88.46%)
25	1.4986	0.9260 (38.21%)	0.2914 (80.55%)
26	1.4292	0.9721 (31.99%)	0.2940 (79.43%)
27	1.9437	1.1705 (39.78%)	0.2504 (87.12%)
28	1.2519	0.8063 (35.59%)	0.2662 (78.74%)
29	1.6434	1.0105 (38.51%)	0.5116 (68.87%)
30	1.3446	0.8137 (39.48%)	0.1207 (91.02%)
31	1.3725	0.8317 (39.40%)	0.2569 (81.28%)
32	1.3985	0.8607 (38.46%)	0.2675 (80.87%)
Average reduction		36.68%	78.99%

TABLE 3: Standard deviation (m) of coordinate time series.

	Original	CSF corrected	SVR corrected
North	0.9136	0.7223 (20.94%)	0.5902 (35.40%)
East	1.2180	0.9997 (17.92%)	0.9033 (25.94%)
Up	2.6069	2.1715 (16.70%)	1.9496 (25.21%)

performance of the FIR filter method [14] is unreliable as only one satellite (PRN 9) was used for analysis. It is clear that SVR estimators significantly outperform the other reported methods. The state-of-the-art performance of SVR estimators on pseudorange multipath reduction emphasizes the efficiency of the proposed method. However, as the accumulative environmental changes become more and more severe over time, the performance of the estimators would temporally degrade. Therefore, the multipath estimators need to be equipped with adaptability, which has not been addressed so far.

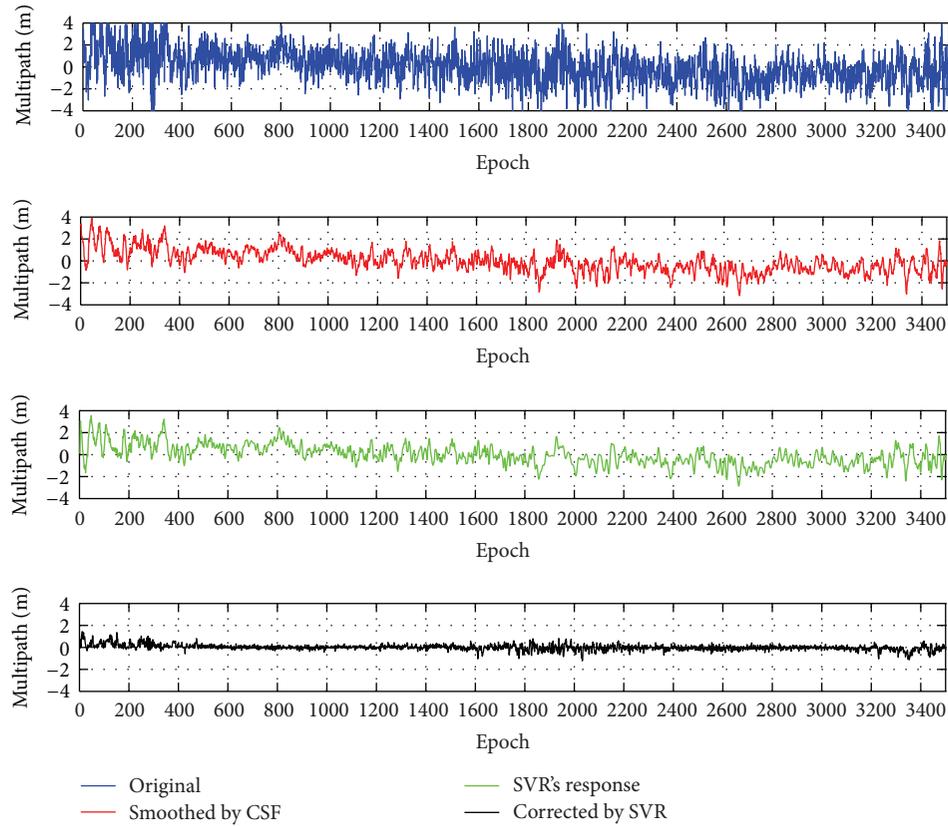


FIGURE 7: Original pseudorange multipath, CSF-smoothed pseudorange multipath, SVR-estimator response, and SVR-corrected pseudorange multipath of PRN 12 of the orbital plane B in Day 310 of 2010.

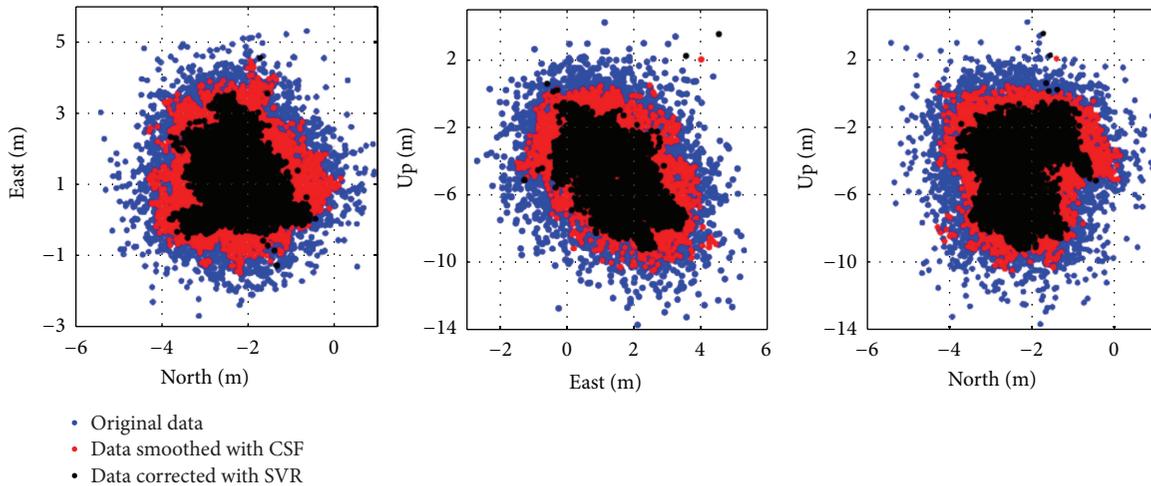


FIGURE 8: Positioning solution on original data, CSF-smoothed data, and SVR-corrected data of Day 310 of 2010.

Unlike the multipath stacking-based approaches [16, 17, 29], modeling multipath errors as functions of continuous variables (i.e., azimuth and elevation angles) does not experience the difficulty in the determination of the time-shifting period. In addition, the interpolation ability of the trained estimators makes them applicable for different data rates provided that training data is adequate to capture the

underlying distribution of multipath errors. Furthermore, with the nature of sparsity, the multipath estimators just count on a subset of training data, being simpler while requiring less storage. All of these imply better scalability.

Another distinct advantage of the proposed approach is that it is able to preserve other signals of studied phenomena such as deformation caused by earthquakes. This is achieved

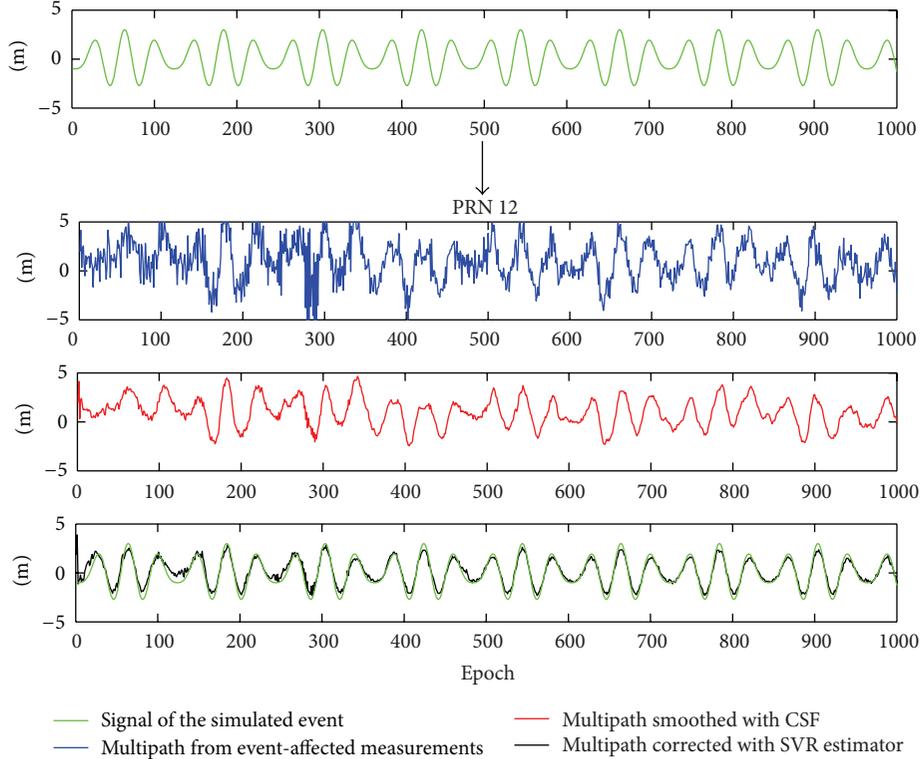


FIGURE 9: Simulation of event signal added to pseudorange measurements of PRN 12. It can be seen that the event signal is indeed left intact from correction of the PRN 12’s SVR estimator.

TABLE 4: Performance comparison of code multipath mitigation methods.

CSF [13]	58%
FIR Filter [14]	75%
FFT [7]	50%–70%
Wavelet [9]	55%–65%
SVR estimator	68%–91%

by training the models with data on normal days without displacement before using them to correct data on the subsequent days where a phenomenon occurs. For the purpose of demonstration, we simulated an earthquake-like event by adding the signal given in (18) to PRN 12’s pseudorange measurements for Day 310:

$$e(t) = 2 \cos\left(\frac{\pi}{10}t + \pi\right) + \cos\left(\frac{\pi}{15}t\right). \quad (18)$$

Since signals of phenomena are usually low frequency [14, 23], the frequencies of the simulated event were chosen to exhibit diminishing effects of CSF, which is a low-pass filter [9, 12]. The PRN 12’s pseudorange multipath sequence was smoothed by CSF with a 50-second smoothing window and then corrected by the trained PRN 12’s SVR estimator. As shown in Figure 9, the corrected multipath sequence aligns very well with the event signal; that is, the event signal is not affected significantly.

Although carrier-phase multipath error is more sensitive to environmental variation, it also saw correlation on a side-real daily basis [15, 29]. Therefore, the proposed method could be extended to apply to carrier-phase multipath mitigation.

6. Experiment 2: Carrier-Phase Multipath Mitigation

6.1. Double-Differential Carrier-Phase Multipath. Unlike code multipath error, carrier-phase multipath error cannot be isolated using combinations of code and carrier-phase measurements of one receiver alone. In order to do that, double differential (DD) combination of two short-baseline receivers (<10 km) is used. In practice, high-precision applications rely on this combination for relative positioning. The DD combination of two satellites a and b observed at two receivers A and B is given by

$$\phi_{AB}^{ab} = \Delta T^{ab} - \Delta I^{ab} + \Delta M^{\phi,ab} + \lambda \Delta N^{ab} + \Delta \varepsilon^{\phi,ab}. \quad (19)$$

In (19), clock error terms are eliminated through the differencing process. Furthermore, for a short baseline, atmospheric effects are approximately equal, leading to $\Delta T^{ab} \approx 0$ and $\Delta M^{ab} \approx 0$. The integer ambiguity term ΔN^{ab} can be computed by debiasing over the entire DD sequence as long as there are no cycle slips. The multipath effect, hence, is



FIGURE 10: IGS observation stations: (a) KIRU and (b) KIRO.

dominant in the DD measurements which can now be written as

$$\phi_{AB}^{ab} = \Delta M^{\phi,ab} + \Delta \varepsilon^{\phi,ab}. \quad (20)$$

The DD multipath error is the composition of four multipath errors of the four carrier-phase components. Denote θ_j^i and φ_j^i the elevation and azimuth angles of the satellite i with respect to the receiver j , DD multipath error actually depends on eight geometrical variables:

$$\Delta M^{\phi,ab} = f_{\phi}(\theta_A^a, \varphi_A^a, \theta_B^a, \varphi_B^a, \theta_A^b, \varphi_A^b, \theta_B^b, \varphi_B^b). \quad (21)$$

Fortunately, at each epoch, eight variable parameters can be calculated given the orbital information of the satellites broadcast to the receivers and nominal positions of the receivers which are known beforehand. Therefore, the DD multipath function f_{ϕ} can be learned from historical data.

6.2. Experimental Data Set. In this experiment, two short-baseline stations from the IGS monitoring network were chosen. Their continuous 1-second data during two consecutive days (Day 050 and 051 of 2011) were downloaded from IGS data archive (<http://igsb.jpl.nasa.gov/components/prods.html>). The stations named KIRO and KIRU whose corresponding geodetic coordinates are (21.0602°, 67.8776°, 497.9000 m) and (20.9684°, 67.8573°, 391.1000 m) (20.9684°, 67.8573°, 391.1000 m) in (longitude, latitude, and height) triplet or equivalent to (2242.624 km, 5516.729 km, 2277.795 km) and (2245.915 km, 5519.218 km, 2268.268 km) in ECEF coordinate system. They are located in Kiruna, Sweden, with a separation distance of approximately 10 km. Figure 10 shows photographs of the locations and surroundings of the two stations. As seen in Figure 10, the stations' vicinity is covered by snow; therefore, the multipath environment is expected to change quickly even in a short time span as the weather changes. If so, the performance of the multipath estimator would degrade.

The station KIRU is selected as the reference station. With predetermined nominal positions of the stations, the geometrical parameters of the satellites with respect to the stations can be computed easily at every epoch given the

broadcast navigation data. In order to extract DD multipath as described earlier, a satellite with high elevation is usually selected as a reference to form the DD combination [1, 13]. The pair of satellites PRN 8 and PRN 18 with the longest DD sequences are thus selected for analysis. PRN 8, with higher elevation peak, is used as the reference satellite. Their geometries with respect to the two stations are illustrated in Figure 11.

6.3. Training Multipath Estimator. In order to train a DD multipath estimator for the satellite pair, the steps that have been done previously for code multipath estimation will again be performed. The geometrical data and multipath errors in Day 050 are scaled to $[-1; +1]$ before feeding to the training program. The *libSVM* package that implements ε -SVR algorithm is employed to train the multipath estimators. The values of ε , the Gaussian kernel parameter η , and the penalty parameter of the error term C are again chosen using grid search and cross-validation with the same strategy as described previously for code multipath mitigation.

Learning from the training data set, the support vectors and coefficients of the DD multipath estimator are found. Finally, the multipath estimator f_{ϕ} should estimate DD multipath error when presented with a new observation of azimuth/elevation angles thereafter. The estimated multipath error will be used to correct carrier-phase DD of the satellite pair in the successive days.

6.4. Experimental Results. In order to demonstrate the ability of the proposed method, the trained estimator is used to estimate the carrier-phase multipath error of the following Day 051. From top to bottom, Figure 12 shows the original multipath sequence, the response of the trained estimator, and the corrected multipath sequence with the response. The standard deviations of the original multipath sequence and the corrected multipath sequence are 2.15 cm and 0.92 cm, respectively. In other words, the multipath error has been reduced by 57.2% using the trained multipath estimator. Although only two data sets for one satellite pair have been presented here, this preliminary result provides a strong indication that the proposed technique can be applied to the carrier-phase multipath mitigation problem.

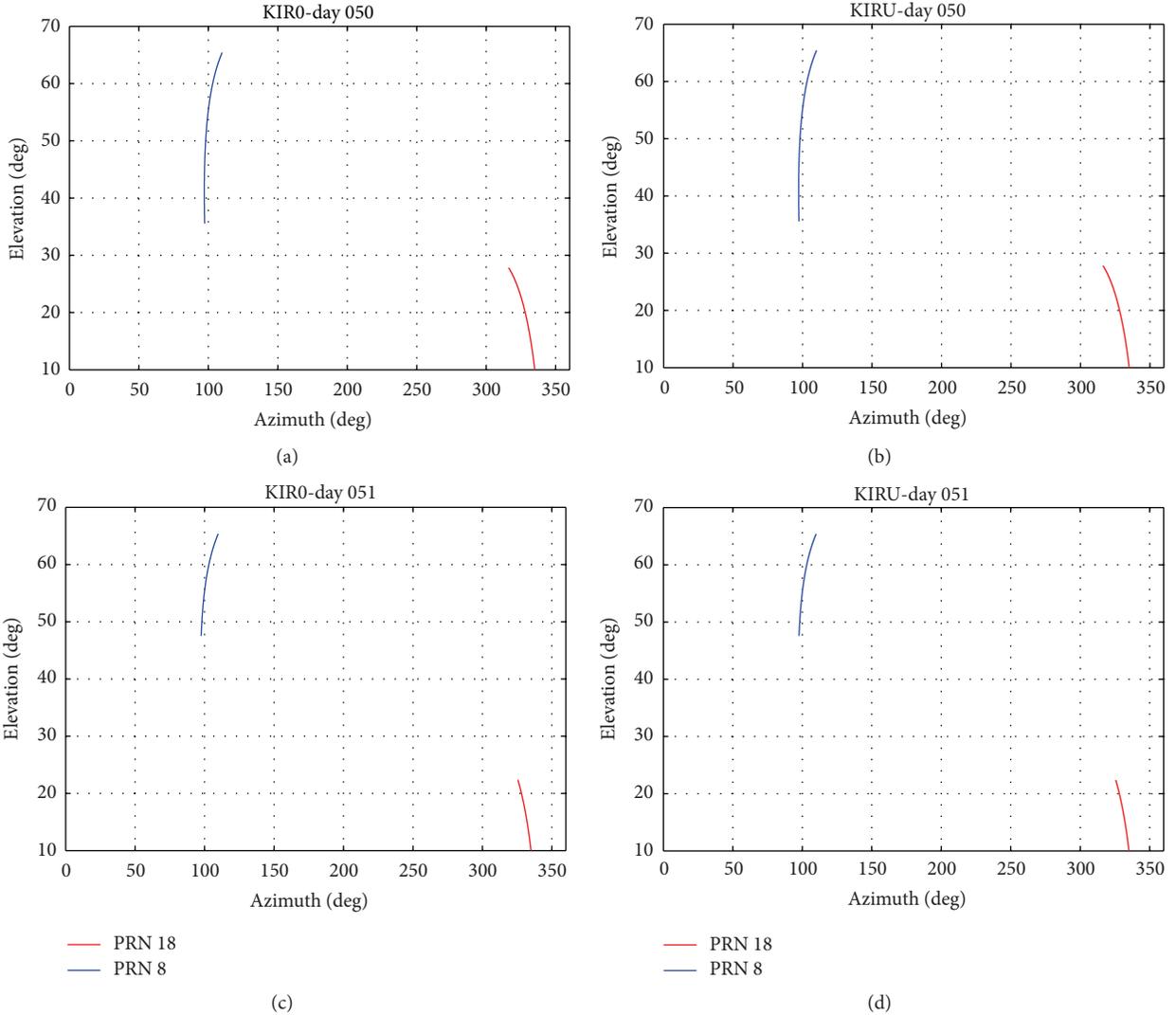


FIGURE 11: Geometries of PRN 8 and PRN 18 with respect to KIRO and KIRU in Day 050 and Day 051.

7. Discussion

As expected, the response of the multipath estimator can capture the dominant trend of the multipath sequence which is the dominant component of the multipath error [8, 9]. However, the result of the proposed method when applied to carrier-phase multipath mitigation seems to be lower than what was achieved for code multipath mitigation. It is explainable as there is no equivalent smoothing algorithm like CSF to smooth the multipath sequence to attenuate high-frequency noise during training and correction. Although it is unwise to conclude firmly since the experimental result is only for one pair of satellites, for the sake of comparison, the proposed method outperforms the FIR filter method (39.8%–56.1%) [15], and the frequency-domain processing methods relying on analysis of SNR measurements (20%) [10, 22]. It provides a motivation for further exploration of the proposed method on the carrier-phase multipath mitigation problem.

However, this result is incomparable to that of the direct frequency-domain processing methods using wavelets such

as those of Elhabiby et al. [8], which can reduce carrier-phase DD multipath up to 84%. It is because these methods can attenuate both errors from low to high frequencies to achieve better mitigation effects at the cost of filtering out other signals if their frequency content overlaps with multipath's frequency content. Although it has not been proven and more comprehensive experiments need to be conducted in future work, the proposed method should share similar advantages of code multipath estimators that it, intuitively, does not affect other phenomena signals. This could be fulfilled if data without displacement is used for training before estimating multipath error for data with displacement. Therefore, for applications where additional important signals such as seismic signals exist, the proposed method is promising.

As carrier-phase multipath is much more sensitive to environmental changes than pseudorange multipath, the data training over a short time span is more suitable for training SVR estimators under the constant environment assumption. It leads to higher data rates being required for training in

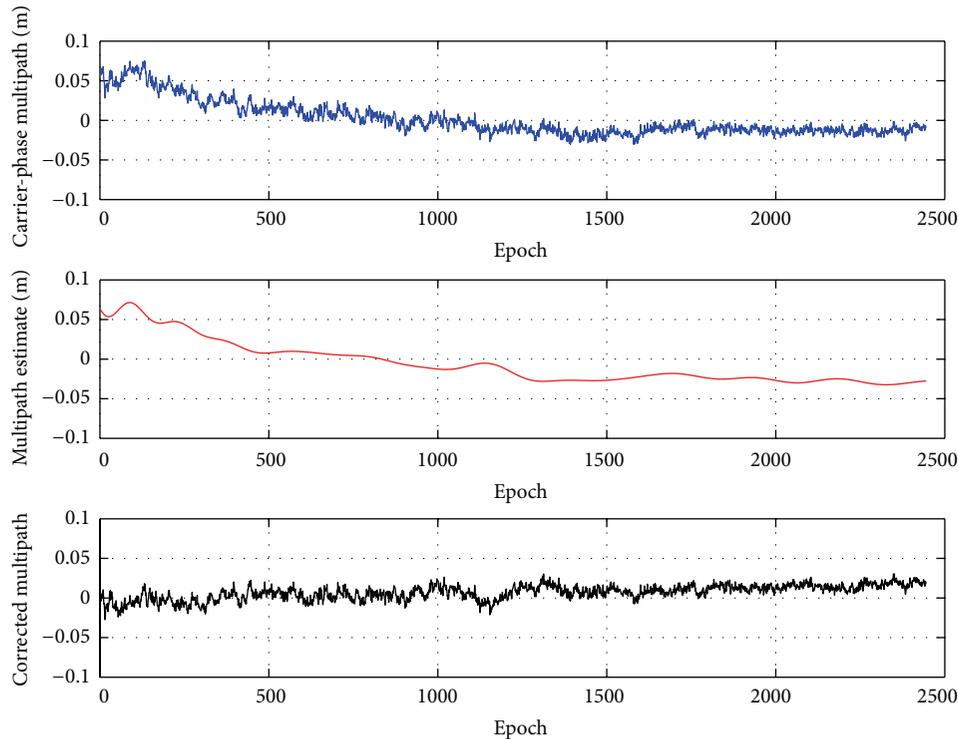


FIGURE 12: Original carrier-phase multipath (blue), response of the trained multipath estimator (red), and corrected carrier-phase multipath (black) sequences of the satellite pair PRN 18–PRN 8.

order to retain data density. However, it does not strictly require high-rate data (≥ 1 Hz) as stacking-based approaches do [17, 29]; only data rates that can capture the underlying distribution of the multipath error are required. Fortunately, the proposed method is also scalable with data rate due to its inherent sparsity. In other words, only a subset of training data which are support vectors are kept and involved in the computation. However, if the number of support vectors is significantly large, computational complexity may still be a concern for real-time operation. This issue needs further exploration.

8. Conclusion

This paper has presented a unified framework based on nonlinear support vector regression to address the GPS code and carrier-phase multipath mitigation problems for ground fixed GPS stations. Based on analysis of the geometry of multipath signal reflections, geometrical models of multipath errors are developed. More specifically, multipath errors corresponding to a satellite are mathematically formulated as functions of the satellite's geometry with respect to a receiver, which is parameterized by azimuth and elevation angles. As a result, the problem of multipath error estimation amounts to regression problem where the multipath functions are approximated by learning from training data using ε -SVM algorithm. Finally, the trained multipath estimators are employed to correct measurements for successive days. The proposed method demonstrates good performance and is

scalable to data rate. Furthermore, the multipath estimators do not affect the simulated signal of phenomena.

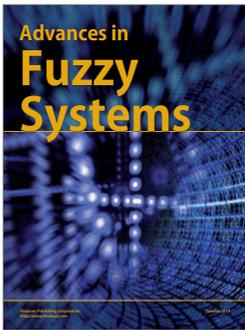
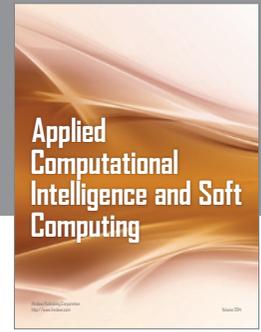
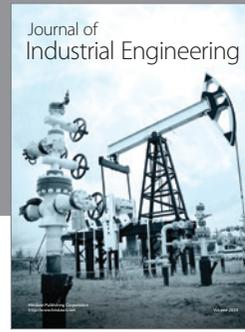
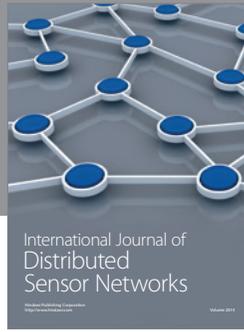
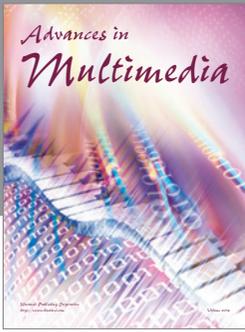
Acknowledgment

This work was done while Q.-H. Phan was a graduate research student at School of Computer Engineering, Nanyang Technological University, Singapore.

References

- [1] B. Hoffmann-Wellenhof, H. Lichtenegger, and J. Collins, *GPS: Theory and Practice*, Springer, Wien, Austria, 2001.
- [2] A. Leick, *GPS Satellite Surveying*, 3rd John Wiley & Sons, New York, NY, USA, 2004.
- [3] G. J. Bishop, D. S. Coco, P. H. Kappler, and E. A. Holland, "Studies and performance of a new technique for mitigation of pseudorange multipath effects in GPS ground stations," in *Proceedings of the 1994 National Technical Meeting of The Institute of Navigation*, pp. 231–242, San Diego, Calif, USA, January 1994.
- [4] N. Kubo and A. Yasuda, "How multipath error influences on ambiguity resolution," in *Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS/GNSS '03)*, pp. 2142–2150, Portland, Ore, USA, 2003.
- [5] Y. Zhang, *High performance differential global positioning system for long baseline application [Ph.D. thesis]*, Ohio University, 2005.

- [6] B. Parkinson and P. Enge, "Differential GPS," in *Global Positioning System*, B. Parkinson, J. Spilker, P. Axelrad, and P. Enge, Eds., vol. 2, pp. 3–49, American Institute of Aeronautics and Astronautics, Washington, DC, USA, 1996.
- [7] Y. Zhang and C. Bartone, "Multipath mitigation in the frequency domain," in *Proceedings of the IEEE/ION Position Location and Navigation Symposium (PLANS '04)*, pp. 486–495, Athens, Ga, USA, April 2004.
- [8] M. Elhabiby, A. El-Ghazouly, and N. El-Sheimy, "A new wavelet-based multipath mitigation technique," in *Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '08)*, pp. 625–631, Savannah, Ga, USA, September 2008.
- [9] Y. Zhang and C. Bartone, "Real-time multipath mitigation with WaveSmooth \hat{U} technique using wavelets," in *Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '04)*, pp. 1181–1194, Long Beach, Calif, USA, September 2004.
- [10] A. Bilich, K. M. Larson, and P. Axelrad, "Modeling GPS phase multipath with SNR: case study from the Salar de Uyuni, Bolivia," *Journal of Geophysical Research B*, vol. 113, no. 4, Article ID B04401, 2008.
- [11] C. Rost and L. Wanninger, "Carrier phase multipath mitigation based on GNSS signal quality measurements," *Journal of Applied Geodesy*, vol. 3, no. 2, pp. 81–87, 2009.
- [12] P. Y. Hwang, G. A. McGraw, and J. R. Bader, "Enhanced Differential GPS carrier-smoothed code processing using dual-frequency measurements," *Navigation*, vol. 46, no. 2, pp. 127–137, 1999.
- [13] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, IGanga-Jamuna Press, Lincoln, Mass, USA, 2nd edition, 2006.
- [14] L. Ge, S. Han, and C. Rizos, "Multipath mitigation of continuous GPS measurements using an adaptive filter," *GPS Solutions*, vol. 4, pp. 19–30, 2000.
- [15] H. Liu, X. Li, L. Ge, C. Rizos, and F. Wang, "Variable length LMS adaptive filter for carrier phase multipath mitigation," *GPS Solutions*, vol. 15, no. 1, pp. 29–38, 2011.
- [16] P. Axelrad, K. Larson, and B. Jones, "Use of the correct satellite repeat period to characterize and reduce site-specific multipath errors," in *Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS '05)*, pp. 2638–2648, Long Beach, Calif, USA, September 2005.
- [17] K. M. Larson, A. Bilich, and P. Axelrad, "Improving the precision of high-rate GPS," *Journal of Geophysical Research*, vol. 112, no. B5, 2007.
- [18] P. Zhong, X. Ding, L. Yuan, Y. Xu, K. Kwok, and Y. Chen, "Sidereal filtering based on single differences for mitigating GPS multipath effects on short baselines," *Journal of Geodesy*, vol. 84, no. 2, pp. 145–158, 2010.
- [19] Q. H. Phan and S. L. Tan, "Mitigation of GPS periodic multipath using nonlinear regression," in *Proceedings of 19th European Signal Processing Conference (EUSIPCO '11)*, pp. 1795–1799, Barcelona, Spain, 2011.
- [20] Q. H. Phan, S. L. Tan, and I. McLoughlin, "GPS multipath mitigation: a nonlinear regression approach," *GPS Solutions*, 2012.
- [21] B. Parkinson, "Introduction and heritage of NAVSTAR, the Global Positioning System," in *Global Positioning System*, B. Parkinson, J. Spilker, P. Axelrad, and P. Enge, Eds., vol. 1, pp. 3–28, American Institute of Aeronautics and Astronautics, Washington, DC, USA, 1996.
- [22] A. Bilich and K. M. Larson, "Mapping the GPS multipath environment using the signal-to-noise ratio (SNR)," *Radio Science*, vol. 42, no. 6, 2007.
- [23] Y. Bock, R. M. Nikolaidis, P. J. de Jonge, and M. Bevis, "Instantaneous geodetic positioning at medium distances with the Global Positioning System," *Journal of Geophysical Research B*, vol. 105, no. 12, pp. 28223–28253, 2000.
- [24] J. F. Genrich and Y. Bock, "Rapid resolution of crustal motion at short ranges with the Global Positioning System," *Journal of Geophysical Research*, vol. 97, no. 3, pp. 3261–3269, 1992.
- [25] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [26] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [27] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK, 1996.
- [28] J. Dickman, C. Bartone, Y. Zhang, and B. Thornburg, "Characterization and performance of a prototype wideband airport pseudolite multipath limiting antenna for the local area augmentation system," in *Proceedings of the 2003 National Technical Meeting of The Institute of Navigation*, pp. 783–793, Anaheim, Calif, USA, 2003.
- [29] K. Choi, A. Bilich, K. M. Larson, and P. Axelrad, "Modified sidereal filtering: implications for high-rate GPS positioning," *Geophysical Research Letters*, vol. 31, no. 22, pp. 1–4, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

